15

The Applications of Artificial Neural Networks to Engines

Deng, Jiamei, Stobart, Richard and Maass, Bastian
Loughborough University
UK

1. Introduction

Artificial Neural Networks (ANN) provide a broad spectrum of functions which are required in the field of engine applications (modelling, especially for controller design, on-board testing and diagnostics). Exhaust emissions laws are becoming progressively more stringent, while the pressure on fuel economy has been intensifying significantly in the last few years. For diesel engines, a large number of technologies, such as, multi-pulse injection and variable valve actuation, show significant promise to both improve fuel economy and reduce exhaust emissions.

Such technologies lead to high degree of freedom systems. Therefore, the engine management system has to handle this increased complexity. The traditional orthogonal grid look up tables will increase exponentially as the degrees of freedom increase. This will increase the complexity and cost of the mapping and calibration. The electronic control unit (ECU) memory consumption will increase in parallel. Use of non-linear functions and in particular neural networks is offering one important route to managing the data tables and achieving the overall goal of reducing the emissions and improving fuel economy. The need for speed and accuracy in the modelling process tends to militate against phenomenological methods

Moreover, in the general control system design, variables, such as exhaust temperature and exhaust manifold pressure, are the usual feedback signals. The brake specific fuel-consumption (BSFC) and emissions (concentration or specific) are the objective variables to which the controller set points are set in order to achieve minimum values. All of these variables can potentially be represented by black-box models. Brahma et al. proposes a dynamic model as the basis for a fuel path control system (Brahma et al., 2004). Wu et al. demonstrated a neural network approach to represent air flow rate (Wu et al., 2004), Maass et al presented a NO_x prediction neural network model (Maass et al., June 2009) and Maass et al presented a smoke prediction neural network model (Maass et al., November 2009].

Real-time operation and the mapping of complex, highly non-linear and dynamic patterns in engine behaviour are challenges that have to be met in modern combustion engines. Neural networks can handle single-input single-output up to multiple-input multiple-output problems, classification tasks and also function approximation. Their generalisation to unforeseen situations enables a wide application if the design of input data captures all the dynamics of the system. In addition, architectures and combinations of networks have a considerable impact on the performance level. We will address these challenging areas.

Firstly, this chapter will address some data collection procedures, from the design of the experiment to neural network identification. The data acquisition for network development is crucial and the design of experiments has a significant impact on the model performance and data collection length, especially for engine systems. We will explain how to choose data perturbation signal, design of experiment to achieve minimum data. We will use practical engine examples to demonstrate these issues. For the application to engines, the relation should be explainable through the chosen inputs and the choice is influenced by the understanding of relations between inputs and outputs. Acquisition of data needs to be done accurately. It needs to be determined if transient behaviour or steady-state operation provide sufficient features for training and validation. The more features the training data covers, the better the network is trained for generalisation of engine behaviour.

Secondly, this chapter addresses architectures and combinations of networks, the application of ANN and combination of those in engine diagnostics and controller development. Combinations of ANN into groups are described achieving improved overall model behaviour. Here, task distribution into special subtask or error reduction through model redundancy can lead to the best possible result. The combination of ANN includes specialised networks trained for subtasks combined with others resulting in a superior task solution. Task distribution helps in overcoming generalisation problems by including redundant networks whose best result is chosen for solution of a specific task.

Thirdly, practical application examples are shown in the domain of emission modelling and estimation of on-board diagnostics of NO_x and PM for heavy- and medium-duty diesel engines (Maass et al., 2009; Maass et al., 2009). It will also cover Non-linear autoregressive exogenous input (NLARX) neural networks to represent intake manifold pressure, exhaust manifold temperature, exhaust manifold pressure to support control system development (Deng et al., 2010). Neural networks are chosen due to their capability to represent complex and highly nonlinear input/ output relationships and can be used to represent the plant during control simulation, and the behaviour of nonlinear control methods.

2. Architecture choices of neural networks

2.1 Introduction of architectures

The choice of network architecture is dependent on the problem. Classification, linear or non-linear problems, with or without underlying system dynamics guides the choices of network composition and the topology. In general it can be distinguished between three types of networks:

- Single-Feedforward Networks (SLFN)
- Multi-Layer Feedforward Networks (MLFN)
- Recurrent Networks (RNN).

Where the single feedforward network describes a simple mapping network it can be used in classification or for mapping of simple input output functionality. It is defined through a single layer of neurons. Hence, the knowledge storage capacity is restricted and only simple logic relations can be mapped. An extension of this is the multi-layer feedforward network, also found as multi-layer perceptron. This network architecture is defined through a minimum of one hidden layer of neurons. The number of hidden layers can be increased dependent on the problem. However, literature states (reference) that a multi-layer perceptron with three hidden layers is sufficient to map every continuous function by adding a certain number of neurons to meet required complexity. However, big growing networks can be ill-posed for overtraining and be difficult to implement in real-time applications. Therefore, recurrent structures of networks are in place that will accommodate the underlying output dynamics, a feature that is of particular interest with engine applications. In turbocharged combustion engines intake and exhaust shows related dynamics through the turbine and compressor connection. Those dynamics can be taken into consideration with output recurrent network structures.

The automotive sector has applied neural networks models in several different cases. Their main implementation is seen in control design in the area of engine operation. Hence, in engine development neural networks are used for control problems such as fuel injection, output performance or speed (Hafner et al., 2000; Ouladsine et al., 2004). In addition, advanced control strategies as variable turbine geometry (VGT), exhaust gas recirculation (EGR) or variable valve timing (VVT) have been in the focus of ANN modelling (Thompson et al., 2000). Nevertheless, the application is also used for virtual sensing such as emissions (Hanzevack, 1997; Atkinson, 2002) or as described in Prokhorov (Prokhorov, 2005) for misfire detection, torque monitoring or tyre pressure change detection.

The combustion process itself has been investigated and parameters been modelled with neural networks by different authors (Potenza et al., 2007; He et al., 2004). Potenza et al. developed a model estimating Air-to-Fuel Ratio (AFR) or in-cylinder pressure and temperature on the basis of crankshaft kinematics and its vibrations. In the work of He et al. combustion parameters and emissions are modelled under the consideration of boost pressure and EGR.

Typical network structures in these investigations have been the NLARX as has also presented in the example application in the previous section. The NLARX structure can accommodate the dynamics of the system by feeding previous network outputs back into the input layer. It also enables the user to define how many previous output and input time steps are required for representing the systems dynamics best. Other network structures include the radial-basis function networks or single layer feedforward networks for classification problems such as misfire indication or component failure detection.

This section describes the commonly applied architecture of the NLARX model. In addition recent investigations on combinations of artificial neural networks for more efficient applications are presented in a practical example for smoke emission output prediction.

2.2 The NLARX architecture

Amongst several architecture styles the NLARX model structure is a commonly used structure and is presented here. For further topologies the literature shows many examples as can be found in Haykin or Hagan (Haykin, 2001; Hagan, 1999).

A typical structure of a NLARX model is illustrated in Figure 1. The inputs are represented by $u(n)$ and the outputs are described by $y(n)$. The inputs are represented by $u(n)$ and the outputs are described by $y(n)$. The formulation of this NLARX model can be described as:

$$y(n) = F(y(n-1), ..., y(n-ny), u(n), ..., u(n-nu+1)) \tag{1}$$

where ny is number of past output terms used to predict the current output, nu is the number of input terms used to predict the current output.

Each output of an NLARX model is a function of regressors that are transformations of past inputs and past outputs. Usually this function has a linear block and a nonlinear block. The model output is the sum of the outputs of the two blocks. Typical regressors are simply delayed input or output variables. More advanced regressors are in the form of arbitrary user-defined functions of delayed input and output variables.

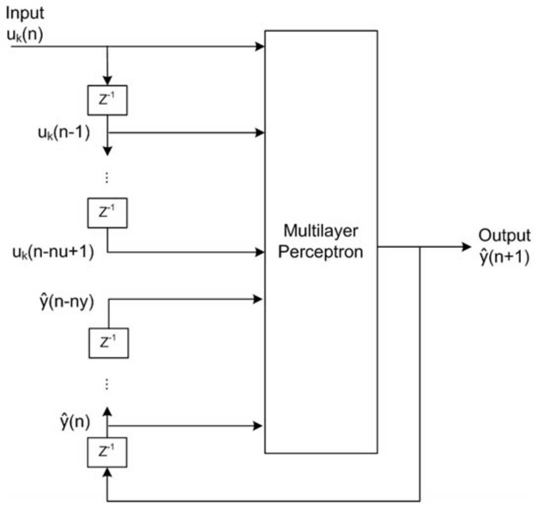


Fig. 1. Canonical representation of a NLARX model structure

The NLARX model training can be cast as a non-linear unconstrained optimization problem:

$$\min_{\theta} F_M(\theta, Z_M) = \frac{1}{2M}\sum_{t=1}^{M} ||y(k) - \hat{y}(k|\theta)||^2 \quad (2)$$

where $Z_M = [y(k), u(k)]_{k=1,...,M}$ is a training data set, $y(k)$ represents the measured output which is the measured soot in the training set, $\hat{y}(k|\theta)$ is the NLARX output, $||.||^2$ is a 2-norm operation, and θ is a parameter vector, where $\theta = [\theta_1, ..., \theta_i, ..., \theta_p]$ and p is the number of parameters. The training process can be described as follows: Given a neural network described by equation 1, there is an error metric, that is referred to as performance index of equation 2. This index is to be minimised and represents the approximation of the network to some given training patterns. The task will be to modify the network parameters θ to reduce the index $F_M(\theta, Z_M)$ over the complete trajectory to achieve the minimal value.

3. Data collection

Data collection should capture as much information possible from the engine application, either through design of experiment or using perturbation signals. This section will discuss the definition of the engine test where the target of the modelling exercise is to represent gaseous emissions, using random signals as perturbation signals and design of experiment method to decide the data requirements. .

Data acquisition is a key element for successful modelling of systems behaviour. In the field of neural network modelling the training data is crucial for creating a good generalising network covering a broad range of the systems behaviour. Hence, a sufficient design of experiments is a key for a successful neural network design.

An efficient and sufficient training requires a data generation strategy that defines the least required data covering the broadest engine operation range. This data set does not necessarily need to contain all different operation states. If it contains the main system

dynamics represented in characteristic features the network would be able to generalise engine states in between recorded data. However, missing out extreme states in the operation may result in a lack of training information. Neural networks cannot extrapolate states that are not covered by the training data as shown in the subsection.
Data collection can be divided into the following categories for diesel engine applications:

1. Predefined engine tests that are used for engine calibration or meeting legislation requirements.
2. Pseudo-random signal generation for engine parameters such as fuel-rail pressure or start of injection that explore a broader range of engine performance.
3. Design of experiment, such as classical, space-filling or optimal design experiments.

This section will use the examples to cover these three aspects of the data collection.

3.1 Predefined engine tests

New emission regulations are going to take effect within the next years in Europe and North America. These implementations bring more and more stringent Emission standards. Different regions have different engine requirement tests. The Non-Road Transient Cycle (NRTC) is an engine dynamometer transient driving schedule of total duration of about 1200 seconds. The speed and torque during the NRTC test is shown in Figure 2. It is a cycle that was devised by the Environmental-Protection Agency (EPA) of the United States of America to represent the range of operating conditions of off-highway machinery. It is the standard test cycle for Tier 4 emissions standards. Normally, the motivation for this choice of cycle is twofold. Firstly, experience has shown that this is one of the most challenging cycles in terms of emissions modelling. Secondly, engine manufacturers must conform the emissions legislation of which the NRTC cycle is an integral part. The current trend is to design engines that pass legislative emission tests by a small margin, but where that margin must be provably robust against deterioration in engine systems. For this the data generated by this cycle is of critical importance.

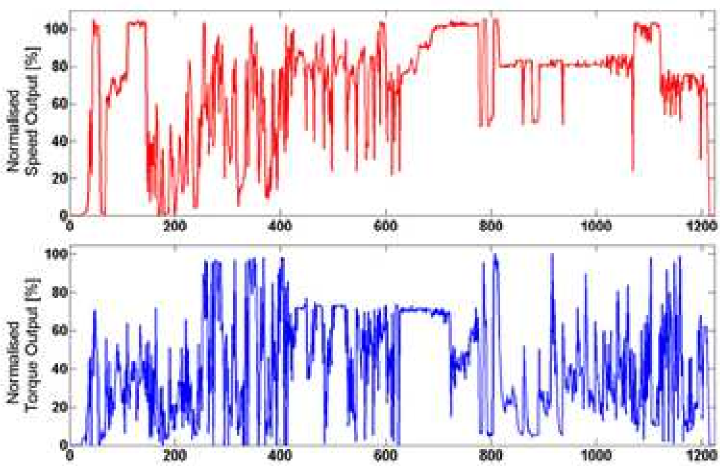


Fig. 2. Non-Road-Transient-Cycle (NRTC) displayed in normalized speed and torque characteristics – used for generation of Data set I [Dieselnet, 2009]

The data used in this section originates from two independent experiments to show the general applicability of the proposed method of prediction. The first data set is created with a NRTC as it is used for certification of non-road engines meeting EPA and EU standards. In the second test a composition of test cycles is operated also including the NRTC.

DATA SET I – The first data set consists of 12 inputs and the NO_x emission output displayed in Figure 3. It is predicted on the foundation of the inputs such as: torque, boost pressure, engine speed, liquid pilot fuel quantity, final fuel injection, back pressure, intake manifold temperature, exhaust temperature, intake depression and coolant temperatures in and out. The data is sampled at a rate of 1Hz and recorded over the whole NRTC cycle range of 1200 seconds.

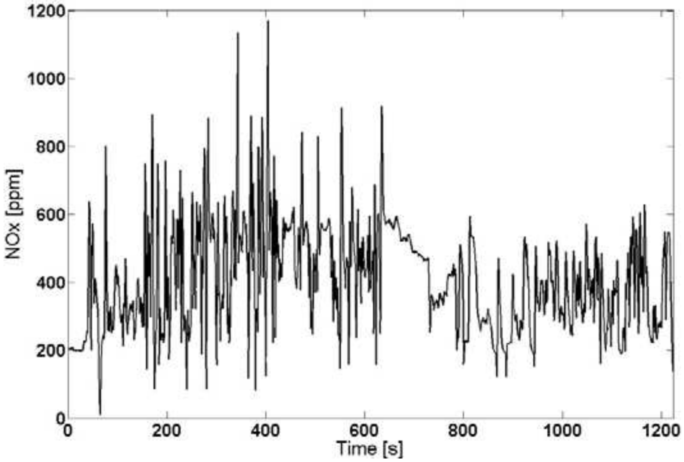


Fig. 3. Data set I - NO_x emission output generated in NRTC mode

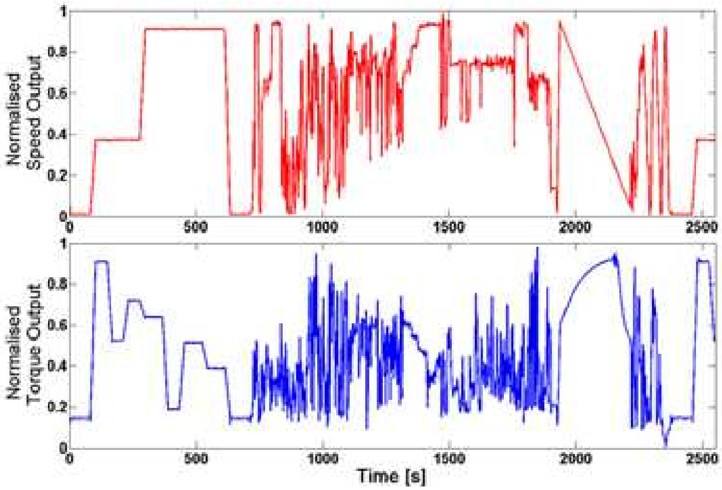


Fig. 4. Test cycle composition of NRTC, ramped modal (8 points), full load and key steady state points

DATA SET II – The second data set consists of 16 inputs to predict the NO_x emission output. The data is also sampled at 1Hz sampling frequency. The operated cycle is a composition of a NRTC, a ramped modal cycle, a full load and some key steady state points as it can be seen in Figure 4. This cycle is repeated 28 times and varied in the engine calibration maps for start of injection (SOI), fuel rail pressure (FRP) and fuel quantity.

3.1.1 Data pre-processing

Both data sets require prior processing in order to ease the training process of the NLARX model. In view of the data variability the sets are normalized to reduce the range of the inputs data. Then a further step of processing is done as follows.

DATA SET I – The initial data set provides limited data in terms of different runs and variation in signal features. Consequently, the data set is re-arranged to spread features into sets of training and validation. The signal is first divided into quarters and then arranged into training sets of the first quarter & third quarter and second quarter & fourth quarter. The result can be seen in Figure 5.

The figure shows a better distribution of signal characteristics. Each set contains a part with high frequent, high amplitudes and a lower frequency section with lower amplitudes.

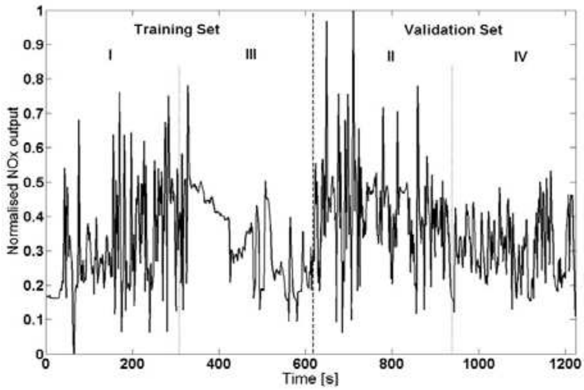


Fig. 5. Pre-processed NO_x output signal. Rearranged and composed training and validation set

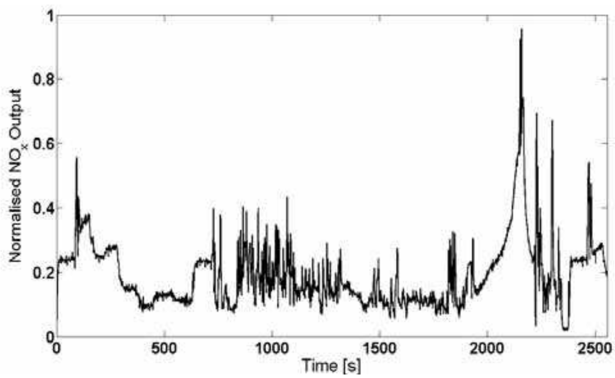


Fig. 6. Data set II training cycle of NO_x target output

DATA SET II – The second data set is split into a training set represented by the first cycle and the residual 27 cycles serve as validation sets individually. Each cycle varies slightly in its range due to the fact of cyclic variations but more importantly that different engine calibration maps are used. Start of injection (SOI), fuel-rail pressure (FRP) and fuel quantity are changed over all 28 cycles systematically. A training output can be seen in Figure 6.

3.1.2 Results

The NLARX models are "teacher forced" trained with an output target as shown before in Figure 4. and Figure 5.

DATA SET I RESULTS - The neural network is fed with the training data and trained manually. The results are promising with R^2=0.96 for the training set and R^2=0.94 for the validation set. The correlation of predicted results with the output target is realized with the correlation method coefficient of determination R^2 that is expressed through:

$$R^2 = 1 - \frac{\sum_{t=1}^{M}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{M}(y_t - \bar{y}_t)^2} \tag{3}$$

Where y_t describes the measured data, $\hat{y}_t$ the prediction and $\bar{y}_t$ the mean value of the output data. The coefficient of determination shows the explained variability of the systems output by the regression model. A result of R^2=1 means an accurate model has been found whereas with a R^2 value of 0 there is no correlation between the system and the model output.

The predicted signal shows a good correspondence with the measured signal as it can be seen in Figure 7.

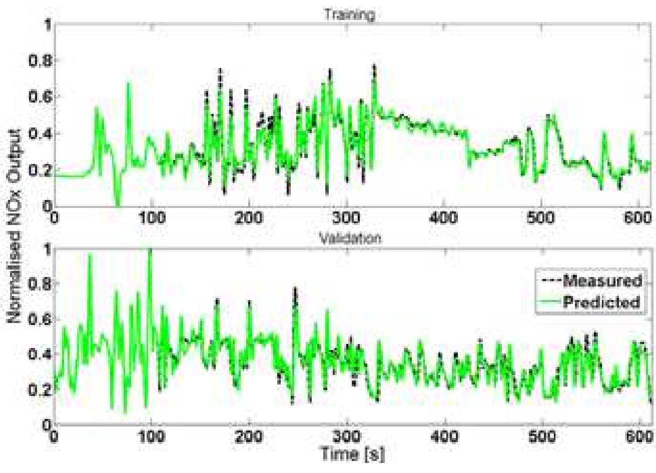


Fig. 7. Correlation of measured NO_x output with predicted neural network signal

However, the model introduces some noise in the second half of the signal. Here, the measured signal fluctuates less but the prediction is characterized with an overreaction. This is assumed to be a side effect of the good correspondence in the more oscillatory region of the test. The model is trained for a more frequent change in the signal and tends to react "nervously" on less varying patterns.

DATA SET II RESULTS – This second data set is to investigate the flexibility of the chosen network architecture. The data set stretches the signal spectrum not only by cycle variances but also with different calibration maps. For the training set a correlation of R^2 =0.95 is achieved as displayed in Figure 8.
Subsequently, this model is individually applied to the residual 27 cycles with the result displayed in Figure 9. It shows the R^2 values over the 27 validation test cycles (black line). A

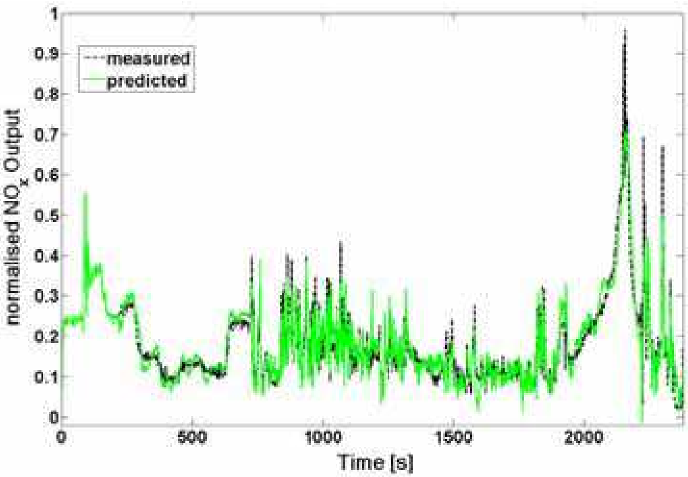


Fig. 8. Correlation between measured target output and predicted output with an R^2 = 0.95

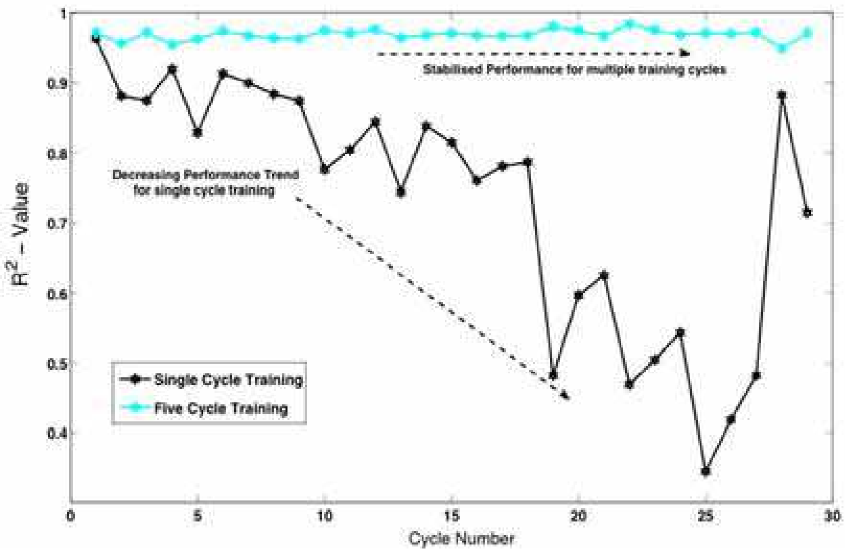


Fig. 9. Trend of prediction for 28 validation test cycles - decreasing correlation with increasing SOI timing (black line) and overcoming calibration variation with multiple training cycles (blue line)

general decreasing trend is recognized whose characteristic seems to result from the increase of SOI timing. With more advanced SOI the NO_x output increases and the signal amplitudes rise. This variance introduces an offset to the signal that cannot be handled by the present model. Hence, the calibration variance has a significant impact on the model performance. The other two calibration variables, FRP and fuel quantity show less impact on the model performance. In order to overcome this performance variance with changing engine calibration settings additional training data is required. Additional features teach the network for a broader application spectrum. The result in performance can also be seen in Figure 8. The R^2 output over all 28 cycles settles above 0.95 that is an acceptable and sufficient result (blue line). This shows that an increase of teaching features improves the knowledge area of the network and underlines the importance of sufficient engine characteristics within a predefined test cycle.

3.1.3 Conclusion

This section shows the data collection for neural network training with a predefined engine test. It is used to create a broad spectrum of engine NO_x output response of two independent heavy-duty diesel engines.
Due to a limited stock of data in the first set the training and validation set is built from a single set of data consisting of 13 channels – 12 inputs and 1 output. As a consequence of this lack of data the available set is recomposed for a better distribution of signal characteristics. This leads through manual training of the NLARX model towards a R^2 value of 0.96 and 0.94 for training and validation set respectively.
The second data set provides a broader validation spectrum because of calibration variances in SOI, FRP and fuel quantity over 28 test cycles. The training results achieve an R^2 value of 0.97 whereas the validation value ranges between R^2=0.88 down to R^2=0.76. An increase in SOI timing causes an offset in the signal that cannot be handled by the trained model. This problem requires a broader featured training set that actually includes the peaks caused from particular input characteristics such as, for example, an increasing load demand. Hence, a training set of five cycles from data set II is created that covers different calibration settings. The correlation result improves significantly over the whole set of data with the R^2 value settling above 0.95.

3.2 Random signal for data generation

In order to capture as much dynamic information as possible, random steps are used as input signals. They are discrete time signals where steps of random magnitude may occur at sampling instants with a certain probability *p*. The input signal *r* can be expressed as follows:

$$r(k) = \begin{cases} r(k-1) & with \quad probability \quad 1-p \\ e(k) & with \quad probability \quad p \end{cases} \tag{4}$$

where k is an integer, e is a discrete time white noise process with zero mean and standard deviation. In the following a modelling approach is presented with following input signals:

- Start of injection timing
- Rail pressure

- Dwell time
- Fuel ratio (quantity ratio between two pulses).

These signals are used to predict exhaust temperature and pressure, compressor mass-air flow and the NO_x output of an engine. Figure 10 and Figure 11 show the random input signals of start of injection timing and fuel-rail pressure for both training and validation purposes. They are representative for the four generated input signals. These figures show the random frequency and amplitude changes of SOI and FRP.

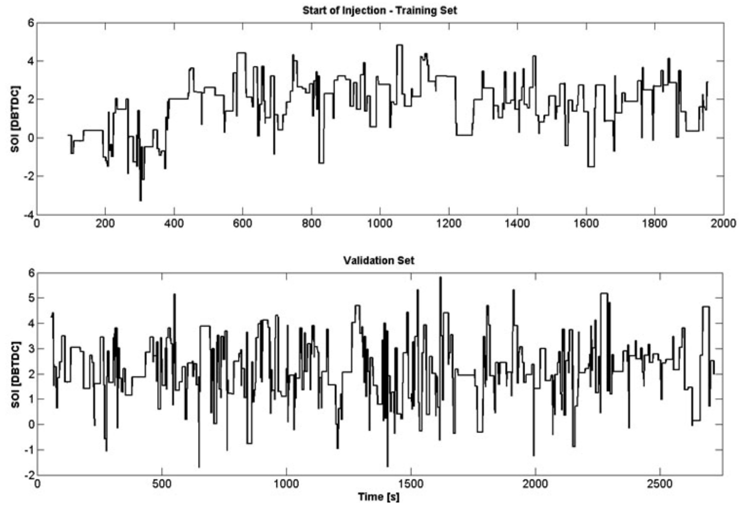


Fig. 10. Random signal of SOI for training and validation

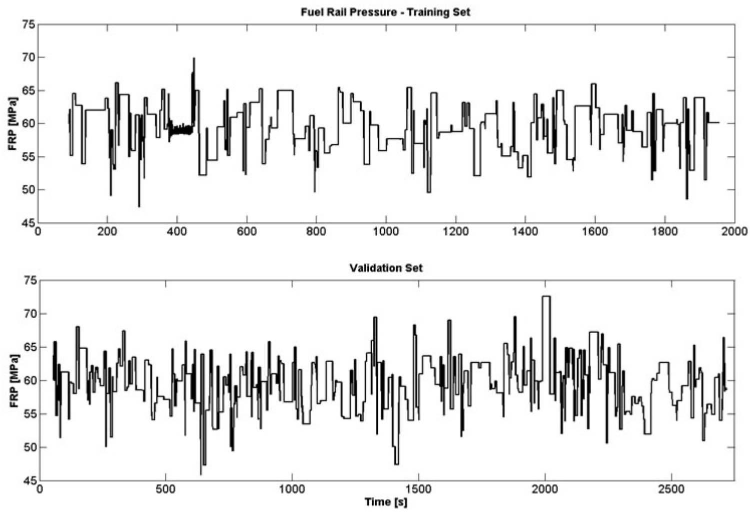


Fig. 11. Random signal of FRP for training and validation

The experiment plan was designed to cover the whole range of fuel injection space as models are effective in interpolating within the range of the training data, but not extrapolating beyond the range. With the engine running at speed of 1440 rpm and torque of 466 Nm, the injection timing spanned a range from -3 degree to 6 degree before top dead center (BTDC), rail pressure from 45 MPa to 75 MPa, dwell from 0.4 ms to 0.5 ms, fuel ratio from 0.5 to 1. Data logged for 2000 seconds was used for training purpose and data logged for a period of 2500 seconds data was used for validation.

3.2.1 Results

The results are summarized in Table 1. Four combinations of input and output are tested. Each output is predicted on the basis of all four inputs. Hence, four different models are created and trained. The correlation of the predicted results with the actual measured results is quantified using the correlation coefficient, R^2 (see (1)).

Test	**Output**	**R^2 Validation**		
		Training	**Validation**	**Fig.**
1	Exhaust manifold temperature	**0.9998**	**0.9997**	**11**
2	Compressor mass flow	**0.9998**	**0.9997**	**12**
3	Exhaust manifold pressure	**0.9957**	**0.9936**	**13**
4	NO_x	**0.9999**	**0.9999**	**14**

Table 1. Results for NLARX models for random signal training

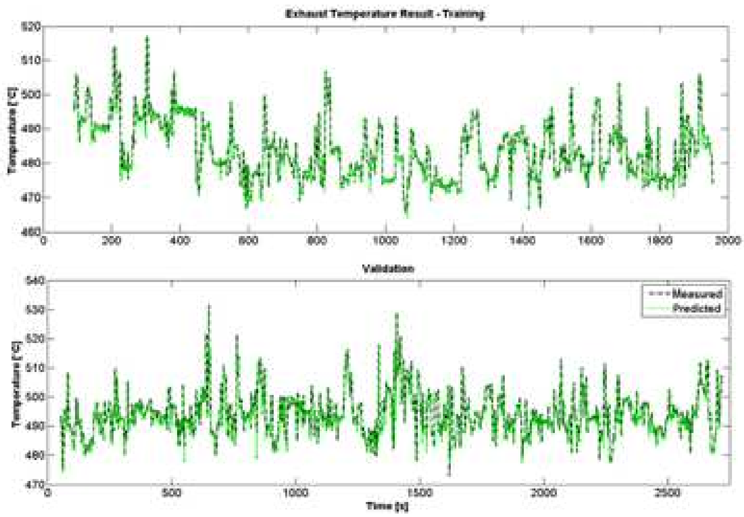


Fig. 12. Correlation of engine exhaust temperature with predicted neural network signal

The results show that the NLARX network is well able to represent the fuel path behaviour. The NLARX model has shown itself useful as a way of representing engine behaviour and that could be used as the basis for a diagnosis algorithm or as a fast measurement.

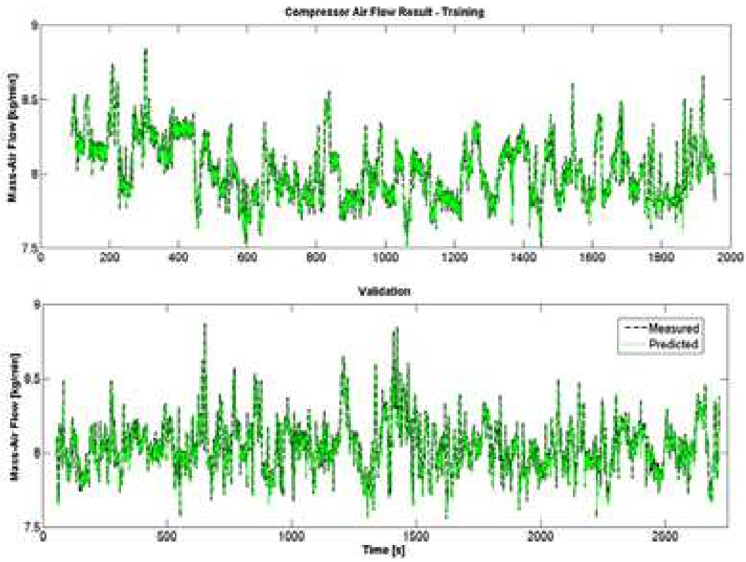


Fig. 13. Correlation of engine compressor mass-air flow with predicted neural network signal

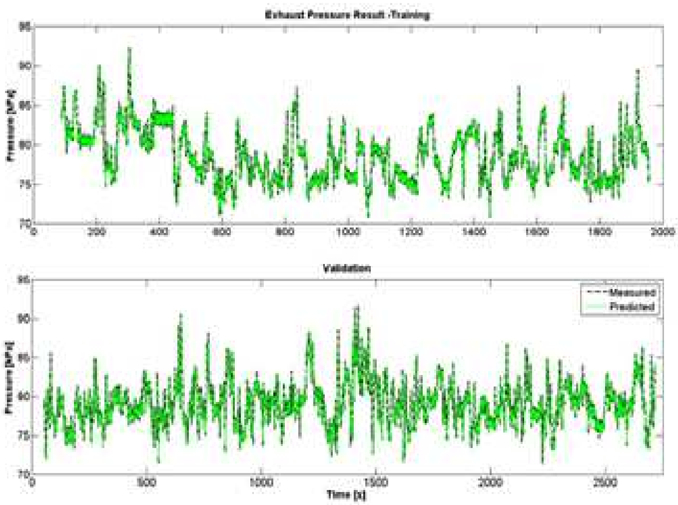


Fig. 14. Correlation of engine exhaust pressure with predicted neural network signal

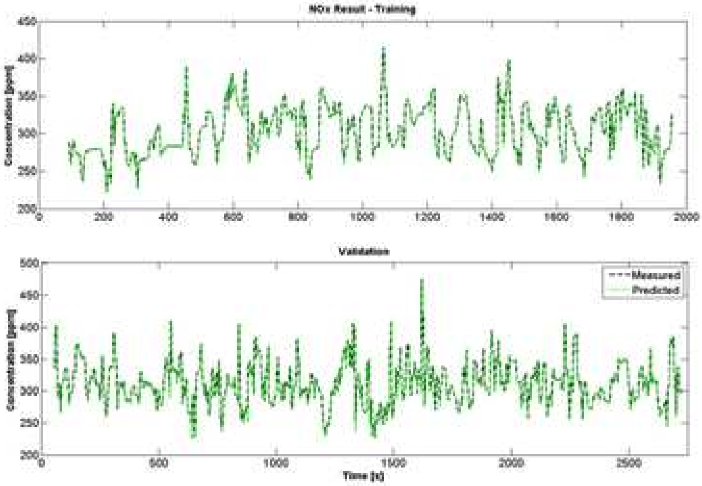


Fig. 15. Correlation of engine NOx with predicted neural network signal

3.2.2 Conclusions

The investigation of fuel path dynamics in regard to the development of a fuel path control algorithm is a novel field of study. This section has shown some initial results intended to support control systems development. The data generated for network training is created with a random signal that is used to perturb engine operation and create a variance in the engine response for exhaust manifold temperature and pressure, engine compressor mass-air flow and the NO_x output. The inputs SOI, fuel rail pressure, dwell timing between injection events and the fuel ratio are varied over a reasonable range at a fixed operation point. This can be applied for several operation points in order to create wider engine behaviour characteristic. Those points can then be used for teaching a single neural network or a combination of networks applied for specific tasks.

A single NLARX model is used for each output parameter measured: exhaust temperature, compressor mass air-flow, exhaust pressure and NO_x. The models demonstrate excellent performance at the operating conditions judged by correlation coefficients close to unity. Further work is required to evaluate the potential for the NLARX model to represent behaviour across a number of operating points. Such a non-linear model is capable of supporting diagnosis processes as well as being a fast model for controls design and evaluation.

3.3 Design of experiment for data generation

This section shows using a design of experiment method to minimise the test and collect informative data for neural networks training and validation.

Figure 16 shows the schematic diagram of a diesel engine. The original engine used for generation of neural network training and validation data is a Caterpillar C6.6 heavy-duty diesel engine with EGR, VGT and VVT function. This engine is modelled in Dynasty 9.4.1 in order to simulate cost effective the engines behaviour. Dynasty is a dynamic simulation tool designed for modelling, simulation and analysis of physical systems in both transient and steady state conditions. During the simulation study, the fuel injection timing and quantity

are held constant. The data for both neural network training and validation are extracted using the Dynasty simulation software. Figure 17 shows the intake and exhaust valve lift. Both inlet and exhaust valve profiles can be changed freely either in the transient or steady state during the simulation.
The experiment plan is designed to cover the whole operating range of the engine. The engine speed spanned a range from 660 RPM to 2000 RPM, torque from 45 Nm to 1000 Nm, EGR from 0.1 to 0.9, VGT from 0 to 1, inlet valve phase shift from 330 degrees to 360 degrees and exhaust valve phase shift from 100 degrees to 140 degrees. The experiment was designed by using the stratified Latin hypercube design method available within the Matlab R2009b Model Based Calibration Toolbox. This design method belongs to the space-filling design style that is used for modelling processes where the system understanding is rudimentary. The purpose is to cover most of the operating range. This design created a total of 196 test points for all parameters. 168 of these test points were used for training purpose and 28 test points were used for validation purpose.

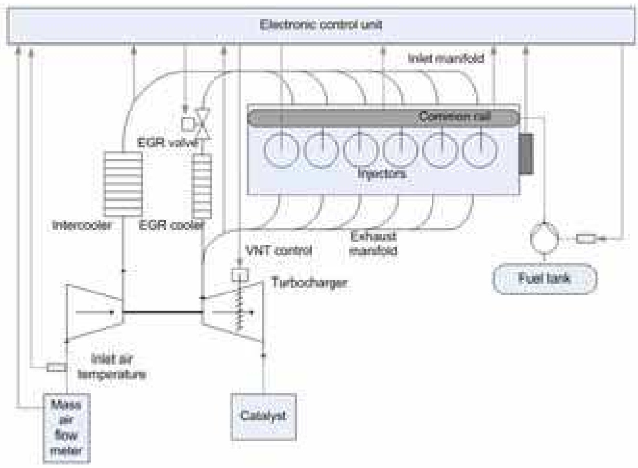


Fig. 16. Schematic drawing of a diesel engine and auxiliaries

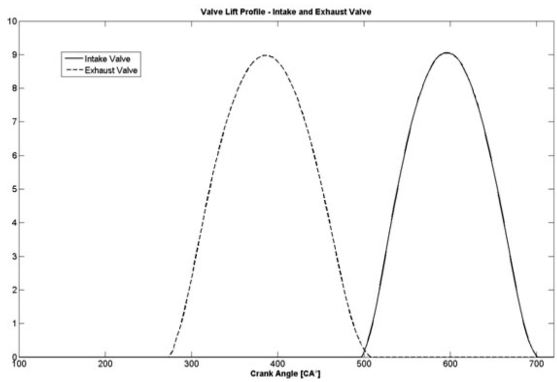


Fig. 17. Valve-Lift profile for inlet and exhaust valve

Additional designs of experiments styles are the classical approach and an optimal approach. The classical approach has been used for simple operation areas with a small number of parameters. . In case of an optimal design of experiments the system knowledge is high and the desired model type is already known. The stratified Latin hypercube design enables the definition of how many operation points per parameter are of interest and leads to an even representation of the multidimensional operation hypercube created by the six parameters in this case.

3.3.1 Results

The first neural network has one output: intake manifold pressure; and six inputs: engine speed, torque, EGR, inlet valve phase and exhaust valve phase. The results are promising with R^2=1 for the training set and R^2=0.9925 for validation set.

Figure 18 shows that the intake manifold pressure predicted from the neural network correlates closely with the generated signal of the Dynasty simulation.

The second neural network is designed to predict BSFC based on six inputs: engine speed, torque, EGR, inlet valve phase and exhaust valve phase. The results are promising with R^2=1 for the training set and R^2=0.9975 for the validation set. It can be seen in Figure 19 the predicted BSFC output of the neural network shows a good correspondence with the measured BSFC from the Dynasty model.

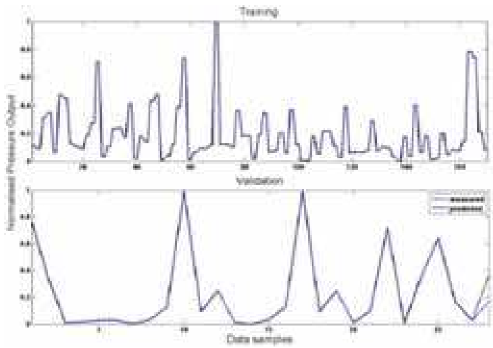


Fig. 18. Correlation of engine intake manifold pressure with predicted neural network signal

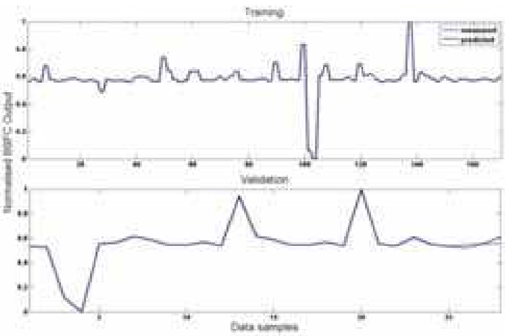


Fig. 19. Correlation of engine BSFC with predicted neural network signal

3.3.2 Conclusion

The design of experiments is a powerful tool in the optimisation of the modelling process. Progressively more complex system architectures make it difficult and eventually impossible to cover each operating point. Depending on the system knowledge different strategies for designing an experiment dictate the sampling coverage for successful modelling processes. The often small knowledge base of parameters effects on the systems response makes the space-filling design style in particular useful for neural network design. This approach allows an even distribution across the operating window and hence covering the main response characteristic for all parameters.

In this particular case for both training and validation data, the sampling points need to be increased significantly. The test points cover a minor operating range of the engine and in order to use neural networks for prediction their generalisation capability has to be increased by additional engine operation characteristics. The approach is presented for the demonstration of a design of experiment and how to use the data in teaching a NLARX that can predict intake manifold pressure and BSFC.

4. Combining Neural Networks

The complexity of today's systems makes it occasionally impossible to find a sufficiently performing single network composition, even in the case of a highly complex recurrent structure. Hence, the combination of networks has become popular where tasks are either distributed across separate networks or competitive structures with redundant networks are created (Sharkey, 1999). The literature distinguishes between modular and ensemble structures. Modular applications are defined by the fact that each network is trained for a subtask and all networks together form a superior solution. In an ensemble networks are trained differently or show different topological features but are predicting all the same output. A superior decision instance compares the results and votes for the best performance. This approach can create a more reliable performance since the optimum can be chosen from a variety of results. A third approach is the combination of modular structures and ensembles. In the following example a parallel neural network structure is composed where three individual NLARX networks are used in order to predict a superior signal that is a combination of all three. Similar to the previous NO_x example, here the smoke emissions are investigated and the behaviour is modelled by a neural network structure. The smoke signal represents in this case the solid component of particulate emissions. Smoke is assumed to be a good proxy for this emission formation.

The experiment for data generation was conducted on a heavy-duty diesel engine that is run under the conditions of an NRTC. It is applied to generate emission data for training and validating the neural network which is presented in the graph in Figure 20. The smoke output signal is predicted on the basis of 12 inputs such as torque, boost pressure, engine speed, liquid pilot fuel quantity, final fuel injection, back pressure, intake manifold temperature, exhaust temperature, intake depression and coolant temperatures for flows in and out. All parameters were used from the beginning and investigated and revised for their impact on the model.

The initial output signal shows two characteristic halves. In the first half strong fluctuations and high peaks are present, whereas the second half displays a much flatter characteristic with small oscillations. The approach of modelling and estimating the signal requires a training and validation data set. Therefore the signal is bisected. However, a training set requires preferably a broad spectrum of features provided by the signal. The signal is first divided into quarters accordingly and then newly-arranged. As a result the training and validation set cover a high oscillating part with high peaks and a flat, low oscillating part – Figure 20. Every set consists of a correspondingly split smoke output and twelve inputs. As well as the data partitioning a normalisation process is applied to the inputs and output.

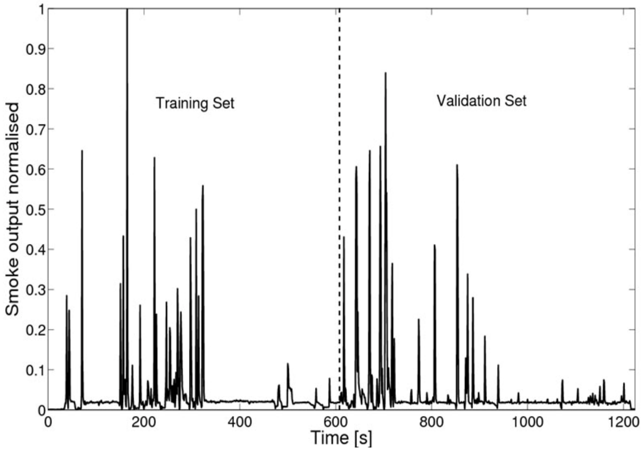


Fig. 20. Processed smoke output signal

In an initial approach of modelling the signal with an single NLARX network it was recognised that noise is introduced by the model. This occurs especially then, if the signal contains large amplitudes and high-frequencies. In Figure 21 the modelling results of a single NLARX model are plotted over the measured signal. The early phase of the signal is well predicted. However, in the second phase of the characteristic the prediction data starts oscillating in high-frequencies as well as an underlying lower frequency. The model becomes unstable. This is assumed to be forced by the training on high amplitudes in the first stage and hence the development of a hypersensitive behaviour. Other approaches are known to overcome those issues such as fuzzy logic and wavelet networks (Parasuraman & Elshorbagy, 2005). They offer a much better response to highly fluctuating signals.

Among those approaches, Guoyin et al. (Guoyin & Hongbao, 1995) introduced three classes of parallel network systems. Here, a parallel network system with multiple tasks is chosen. Lee (Lee, 1997) states that due to the approach of more than one network the risk of settling in a local minimum decreases. Additionally, the performance increases due to the fact that particular networks handle a specific subspace instead of dealing with the whole problem.

In the current work the signal is divided into different vertical layers. Consequently the amplitudes are cut and the frequency of the residual signal part is decreased. With lower frequencies the NLARX model promises satisfying results regarding performance and cost.

By trial and error three layers are determined as a reasonable degree of divisions. The first layer called lower layer (LL) contains the signal noise and low frequencies. The remaining part is split into a mid layer (ML) and a top layer (TL). The ML covers a part of the signal with a medium density of oscillations and peaks in the smoke value up to y=0.3. The residual signal peaks are covered by the TL. Its characteristic is marked by only a few single peaks, the occurrence of which is not distracted by noise or smaller peaks. The division borders in this approach are chosen as outlined in Table 2 and illustrated in Figure 22.

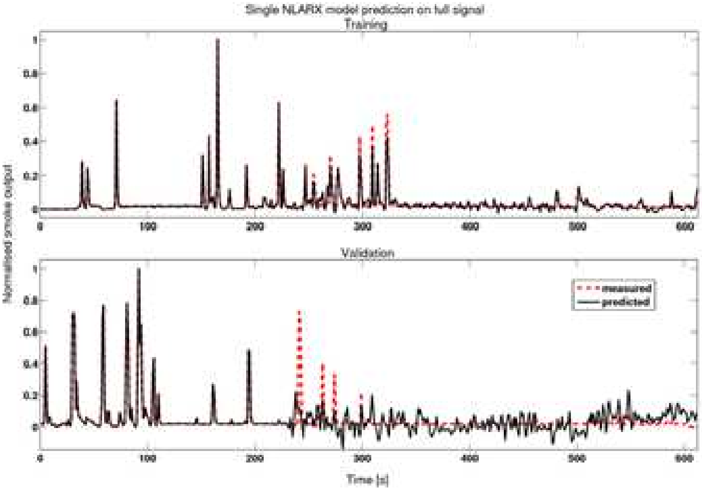


Fig. 21. Single NLARX model: measured output signal correlated versus predicted output signal

0 LL 0.035	=>	$\Delta y_{LL} = 0.035$
0.035 < ML < 0.3	=>	$\Delta y_{ML} = 0.265$
0.3 < TL < 1	=>	$\Delta y_{TL} = 0.7$

Table 2. Division borders of layer approach

Each division is processed and estimated independently. This leads to a parallel processing model structure as shown in Figure 23. The input vector U with its twelve input signals is used for all three independent layers whereas the predicted output is split into the three divisions, *top*, *mid* and *low*. After estimation they are combined to $\hat{y}_{overall}$ and compared against the overall measured output.

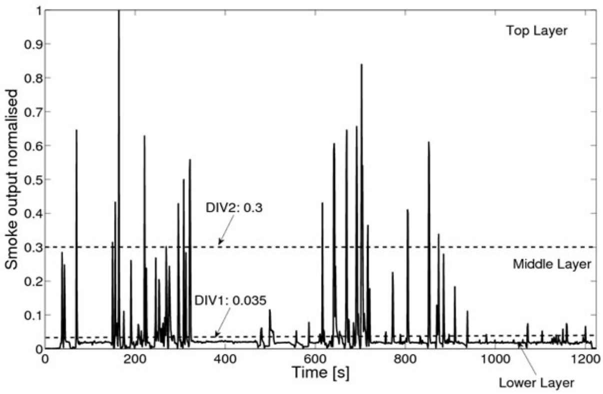


Fig. 22. Layer approach with correlating divisions

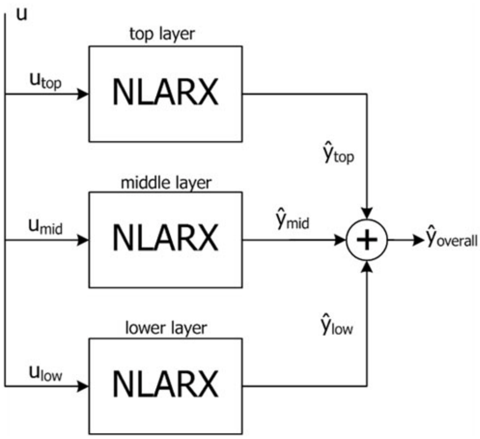


Fig. 23. Scheme of applied parallel model structure

Results - An estimation is processed by initially training and then validating an artificial neural network with the corresponding signals. Every layer is estimated independently. The NLARX-model are initialised with an arbitrarily state and taught with the corresponding training data set. Based on this data the NLARX-model is designed to estimate the desired output signal. The designing process consists of changing the design parameters in Matlab R2009b by teacher-forced learning until a satisfactory result is achieved. The design parameters are the input/ output delays.

The lower part is marked by 1) the lowest values of the higher oscillations of the signal and 2) small oscillations that are introduced by noise. By cutting off a lower part of the signal a more homogeneous distribution of the height of oscillations is created. This enables a better estimation with the chosen NLARX approach.

The training of the network generates a correlation between the measured and estimated signal of R^2=0.97. Validating the network leads to a performance of R^2=0.95 which demonstrates the practicability of the chosen design. However, the model introduces additional noise to the signal. This effect is discussed in more detail in the following sections.

The middle layer represents the central section of the high peaks and the medium peaks. The lowest values of the large signal excursions are included in the lower layer. Through training the NLARX model achieves a correlation of R^2=0.93 with the measured signal. The model's quality is confirmed by the validation set, which achieves a performance of R^2=0.9. The performance is predictably lower than in the first layer due to the higher frequencies. Higher frequencies occur because of an expanded range of y-values.

The characteristic of the graph is marked by noise in the second, low oscillating part of the signal. It is assumed that this noise is introduced as a result of the network design. There is a fast response identified by the network when managing high oscillating signals. In consequence, this leads to an oscillating estimation signal.

The top layer covers the high peaks of the signal. Consequently high frequencies are introduced and a lower correlation performance is expected. The design process achieves a result of R^2=0.83 compared to a R^2=0.92 for the validation data. Validation shows a better result because the main peaks of the validation signal are predicted well, whereas the training signal shows some missing details in the middle part for three consecutive spikes.

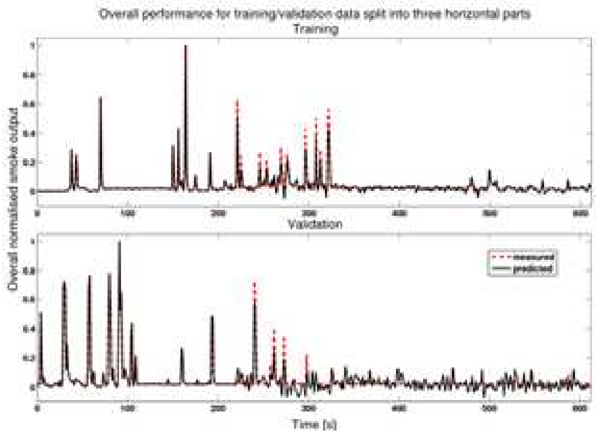


Fig. 24. Overall Performance: Training and validation sets estimation of measured and predicted signal

The overall estimation is achieved by adding the three estimated signals together and correlating it with the measured output – see Table 3. The comparison of the measured and predicted signal shows a distribution around the linear correlation in Figure 25. The reason that a cluster of points forms close to the origin is due to the fact that the most of the data samples are measured in the lower data scope. However, the results of overall correlations of the smoke output signal are R^2=0.97 and R^2=0.96 for training and validation set respectively as illustrated in Figure 24. It can be seen that parts intially classified as difficult due to big amplitude differences and high frequencies are modelled well. The patterns of high peaks and high density of oscillations show appropriate correlations. However, the flatter parts are marked by the introduction of noise through the model design as mentioned earlier.

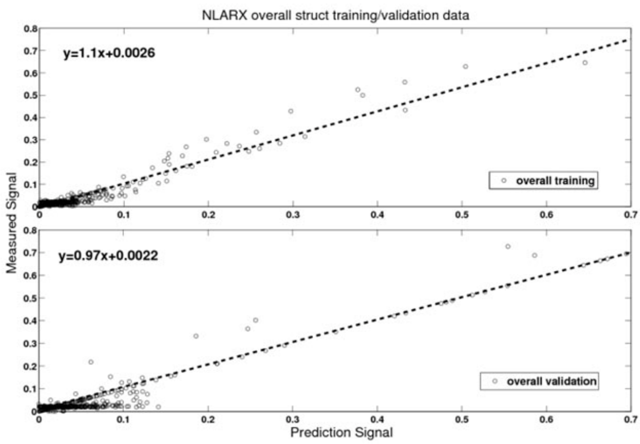


Fig. 25. Overall Performance: Training and validation sets estimation in correlation to measured data

Layer	NLARX-Model	
	train	valid
TL	0.8330	0.9163
ML	0.9328	0.9006
LL	0.9743	0.9164
Total	0.9706	0.9616

Table 3. RMS performance indication

Conclusion – Here a parallel neural network structure is presented to predict the smoke output of a diesel engine based on NLARX-models. The model is chosen due to its good generalisation. Its weakness of not being capable of high-frequency signals which is shown with a single NLARX model for comparison, is overcome by an approach of frequency filtering.
The NLARX architecture is cut into different layers to reduce the frequency bands resulting in a better overall performance. A lower layer for the signal scope that covers noise and the base of higher peaks, followed by a middle layer for medium density of oscillations and a top layer for the peak tops. This approach demonstrates the application of network combination. Three independent networks trained for different tasks can predict if combined the overall signal at a sufficient performance.
Another approach is the training of three networks with the same task just slightly different training data and in case of the prediction performance the network with the best performance wins. Here, the networks are redundant and a competitive approach is used to find the optimum output.

5. Summary

The chapter presents the application of artificial neural networks on engine applications. Several practical examples show the applicability of artificial neural networks in the domain of virtual sensing and control development support.
A critical part of for a successful modelling of engine behaviour is the generation of comprehensive system characteristic. The better the training data describes the system dynamics the better the generalisation capability of the model. A crucial part is the planning of data generation. Here the chapter presents three differen approaches:

- Predefined engine tests such as the NRTC for off-highway diesel applications
- Random control signal generation for engine response measurements
- Systematic design of experiment approach.

Another crucial part is the choice of the right model structure for the problem at hand. The chapter presents a recurrent network structure that is applicable for highly non-linear and dynamic systems. The NLARX network is presented in several different applications. A successful implementation can be seen in the virtual sensing of diesel engine emissions. However, the network has also been implemented for combustion modelling.
A last part describes the investigation of combinations of networks. Increasing complexity of systems leads to difficulties of finding cost effective network structures in view of training and operation costs. An approach is presented where a superior task, the prediciton of smoke emissions of a diesel engine is split into three individual tasks solved by independent network compounds. Other approaches can be implemented with competetive structures of redundant networks whose results are competing against each other.

Artificial neural networks can be a powerful tool for monitoring of engine operation or the design of controller applications. However, the correct training data and the optimal design are crucial for a successful modelling process.

6. References

Atkinson, C.; Traver, M.; Long, T. W. and Hanzevack, E. (2002), Predicting Smoke, *InTech*, 2002, pp. 32-35

Brahma, A.; Upadhyay, D.; Serrani, A. And Rizzoni, G. (2004). Modeling, identification and state estimation of diesel engine torque and NO_x dynamics in response to fuel quantity and timing excitations, *Proceedings of the 2004 American Control Conference*, Vol. 3, 2004, pp. 2166-2171

Deng, J; Maass, B.; Stobart, R. K. (2009), Using Artificial Neural Networks for Representing the Brake Specific-Fuel Consumption and Intake Manifold pressure of a Diesel Engine, *Asia Pacific Power and Energy Engineering Conference 2009*, pp. 1-4

Deng, J; Winward, E.; Stobart, R. and Desai, P. R. (2010), Modeling Techniques to Support Fuel Path Control in Medium Duty Diesel Engines, *SAE World Congress 2010*, SAE paper no. 2010-01-0332, April 2010

Guoyin, W. (1995). Parallel Neural Network Architectures, *Proceedings of the IEEE International Conference on Neural Networks 1995*, Vol. 3, 1995, pp. 1234-1239

Hafner, M.; Schüler, M.; Nelles, O. and Isermann, R. (2000). Fast Neural Networks for Diesel Engine Control Design, *Control Engineering Practice*, Vol. 8, 2000, pp. 1211-1221

Hagan, M. T.; Demuth, H. B. and Beale, M. (1999), *Neural Network Design*, PWS Publishing Company

Hanzevack, E. (1997). Virtual Sensors for Spark Ignition Engines using Neural Networks, *Proceedings of the 1997 American Control Conference*, Vol. 1, 1997, pp. 669-673

Haykin, S. (1999), Neural Networks – A Comprehensive Foundation, Prentice Hall

Lee, B. (1997). Parallel Neural Networks for Speech Recognition, *Proceedings of the IEEE International Conference on Neural Networks 1997*, pp. 2093-2097

Maass, B.; Stobart, R. K. and Deng, J. (2009). Diesel Engine Emissions Prediction Using Parallel Neural Networks, *American Control Conference*, June 2009.

Maass, B.; Stobart, R. K. and Deng, J. (2009). Prediction of NOx emissions of a heavy duty diesel engine with a NLARX model, *SAE Powertrains, Fuels and Lubricants Meeting*, SAE paper no. 2009-01-2796, November 2009

Ouladsine, M.; Bloch, G. and Dovifaaz, X. (2004). Neural Modelling and Control of Diesel Engine Pollution Constraints, *Journal of Intelligent and Robotic Systems*, Vol. 41, 2004, pp. 157-171

Parasuraman, K. and Elshorbagy, A. (2005). Wavelet Networks: An Alternative to Classical Neural Networks, *Proceedings of the IEEE International Conference on Neural Networks 2005*, Vol. 5, 2005, pp. 2674-2679

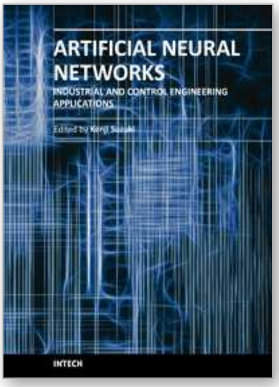
Potenza, R.; Dunne, J. F.; Vulli, S.; Richardson, D. And King, P. (2007). Multicylinder Engine Pressure Reconstruction using NARX Neural Networks and Crank Kinematics, *International Journal of Engine Research*, Vol. 8, 2007, pp. 499-518

Prokhorov, D. (2005), Virtual Sensors and Their Automotive Applications, Intelligent Sensors, *Proceedings of the 2005 International Conference on Sensor Networks and Information Processing*, 2005, pp. 411-416

Sharkey, A. J. C. (1999), Combining Artificial Neural Nets, ensemble and modular multi net systems, Springer, London

Thompson, G.; Atkinson, C.; Clark, N.; Long, T. and Hanzevack, E. (2000), Neural Network Modelling of the Emissions and Performance of a Heavy-Duty Diesel Engine, *Proceedings of the Institute of Mechanical Engineers*, Vol. 214, Part D, No. D04499, 2000

Wu, B.; Filipi, Z.; Assanis, D.; Kramer, D. M.; Ohl, G. L.; Prucka, M. J. and Divalentin, E. (2004). Using artificial neural networks for representing the air flow rate through a 2.4 litre VVT engine, SAE paper no. 2004-01-3054



Artificial Neural Networks - Industrial and Control Engineering Applications
Edited by Prof. Kenji Suzuki

ISBN 978-953-307-220-3
Hard cover, 478 pages
Publisher InTech
Published online 04, April, 2011
Published in print edition April, 2011

Artificial neural networks may probably be the single most successful technology in the last two decades which has been widely used in a large variety of applications. The purpose of this book is to provide recent advances of artificial neural networks in industrial and control engineering applications. The book begins with a review of applications of artificial neural networks in textile industries. Particular applications in textile industries follow. Parts continue with applications in materials science and industry such as material identification, and estimation of material property and state, food industry such as meat, electric and power industry such as batteries and power systems, mechanical engineering such as engines and machines, and control and robotic engineering such as system control and identification, fault diagnosis systems, and robot manipulation. Thus, this book will be a fundamental source of recent advances and applications of artificial neural networks in industrial and control engineering areas. The target audience includes professors and students in engineering schools, and researchers and engineers in industries.

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821