Transient Modeling of Diesel Engine Emissions

Nobuyuki Sakushima 1) Wolf Baumann 2) Karsten Röpke 3) Mirko Knaak 4)

1)・4) IAV Co., Ltd.
101-0041 Kanda Sudacho 1-28, Chiyoda-ku, Tokyo, Japan
(E-mail: nobuyuki.sakushima@iav.jp, mirko.knaak@iav.jp)

2)・3) IAV GmbH
Carnotstraße 1, 10587 Berlin, Germany
(E-mail: wolf.baumann@iav.de, karsten.roepke@iav.de)

Received on July 12, 2010
Presented at the JSAE Annual Congress on May 22, 2010

ABSTRACT: Dynamic engine emission modeling has been attracting a lot of attention over the last years. Applications of dynamic engine modeling include model based calibration or rapid measurement, i.e. methods for saving measurement time. Whereas physical models usually show a high complexity, data driven models are estimated with significantly less effort. In this paper, we show the use of a multichannel sinusoidal excitation sequence for a nonlinear dynamic emission model. This training sequence is used for modeling transient emissions and exhaust temperature. As validation, a measured trace from a new European driving cycle and a FTP cycle is used.

KEY WORDS: Heat engine, Diesel engine, experiment, optimum design, statistical method, Diesel emission, dynamic modeling, transient modeling [A1]

1. Introduction

The future calibration process will become more and more model-based, as this makes the application of systematic optimization techniques feasible. For this purpose reliable engine models for the prediction of quantities of interest are necessary. In this context the simulation of engine out and also of tailpipe emissions plays a key role.

For steady state calibrations, the method of Design of Experiment (DoE) has nowadays become a standard. It relies on the optimal placement of test points, where the optimality is usually derived from the uncertainty of parameter estimates. This allows for a significant reduction of the overall number of necessary measurements. Subsequently, the measured points are used for fitting the model coefficients (model training) whereas the model quality is verified against validation measurements, which have not been in the training dataset. The optimization can then be carried out based on model predictions, and the optimal ECU maps are found.

The steady state approach clearly has its restrictions in case of transient engine operation, like e.g. in driving cycle simulation and optimization. For these applications the use of nonlinear dynamic models may be more appropriate. They usually are purely data driven and have only minor physical relevance. Such a model type is given by Volterra series models which have successfully been applied in several projects by IAV, e.g. for cold start simulation.

In this paper, we give a short introduction to nonlinear dynamic modeling approaches with an explicit explanation of parametric Volterra series. Subsequently, the design of dynamic experiments and its realization at the test bench are discussed. We show the modeling results based on chirp signal test design and its validation against the driving cycle.

2. Nonlinear Dynamic Modeling

2.1. General procedure

From a very general perspective, the procedure for nonlinear dynamic modeling 1) can be summarized as shown in Figure 1. There exists a huge amount of different model classes and corresponding algorithms for parameter estimation.

![Fig. 1 Necessary steps in dynamic modeling](image)

As a first step it is necessary to define the quantities to be modeled and the corresponding inputs. From a physical background or expert knowledge, i.e. prior knowledge it is possible to derive the test design. In the dynamic case, this is a sequence of time variant excitation signals. A good test design is the prerequisite for a successful model fit, especially for multichannel input systems in a nonlinear dynamic setting.

The process of model parameter estimation includes a number of sub-problems, like the selection of significant terms or states (model simplification), the check of stability, or the definition of the error norms. The validation sequences again are a time varying signals, like e.g. the driving cycle of interest.
2.2. Parametric Volterra Series

Within IAV, models of type parametric Volterra series have often been used for modeling of nonlinear dynamic systems (see Figure 2). During the last years we have applied it successfully to different kinds of problems, like e.g. Rapid Measurement driving cycle prediction from cold start and catalytic converter modeling.

The structure of parametric Volterra series is made up of a nonlinear transformation of the input quantities by polynomials with a subsequent finite response filter (FIR) stage. Additionally an infinite response filter (IIR) stage is used to cope with dynamic systems with large time constants. In general the model equation is given by

\[
y(t) = g(x(t), x(t-1), ...) + a_1 \cdot y(t-1) + a_2 \cdot y(t-2) + \cdots + a_N \cdot y(t-N)
\]  

(1)

with \( g(x) \) as the nonlinear transformation of the inputs.

The noteworthy properties of Volterra series are:

- the linearity of parameters,
- the high flexibility (ability to approximate Wiener and Hammerstein systems) and
- easy stability check by methods of linear system theory.

The linearity of parameters allows for estimation of a global solution on the basis of least squares or related methods, like weighted least squares (WLS) or total least squares (TLS). In case of outliers in the data there also are appropriate methods available, like e.g. robust regression.

In case of an autoregressive (AR) stage, the use of instrumental variables usually gives better parameter estimation results because the noise spectrum is not required to be shaped like the AR spectrum.

3. Test Design for Dynamic Excitation

The design of excitation signals plays a key role in the process of nonlinear dynamic modeling. From a theoretical point of view the appropriate excitation can be derived from the system or the model, respectively. That means the more prior information is available the better is the test design with respect to the system under investigation. This approach is common for the steady state case, where an optimal test design can be calculated by statistical criteria based on the Fisher information matrix.

With respect to the nonlinear dynamic case, there are some limitations of this approach. First, a general model for the simulation of transient combustion engine behavior is usually not available. Therefore only very few prior knowledge is available for the excitation design. Additionally there are strong limitations concerning the feasibility of the desired dynamic sequences on the test bench, i.e. there are hull constraints to comply with and restrictions of the gradient of adjustment. This makes the application of amplitude modulated pseudo random binary noise signals (APRBS) often difficult, where especially the big steps that may occur in the APRBS can cause an unstable engine operation. Additionally, the time dependent scaling, which is necessary to ensure that the boundaries are not violated, can lead to changing amplitude levels of the APRBS and may therefore cause deterioration of the optimality, unless the amplitude levels are designed within the engine hull.

Following the theoretical considerations, it has to be ensured that the signals show a sufficient excitation of amplitudes and frequencies and at the same time are preferably orthogonal with respect to the linear combination stage (compare Figure 2). In this regard, a sinusoidal excitation sequence performs very well, as a nonlinear polynomial transformation ensures orthogonality. Furthermore, a sinusoidal excitation has some more interesting advantages:

- Good applicability to combustion engines
- Excitation of important frequency regions
- High D-optimality value
- Easy control of the gradient of adjustment
- Control of condition number of fit via frequency and phase
- Control of input space coverage via frequency and phase.

These properties make sinusoidal sequences an interesting alternative to the APRBS approach. This is also confirmed by the simulation of a nonlinear system identification experiment with different excitation signals. We used the following four excitation signals, each with a maximum amplitude of \( \pm 1 \):

- APRBS
- Sinus
- Triangle
- Chirp

Please note that a chirp signal is a sinusoidal signal with time variant frequency, like e.g. given by the following equation:

\[
x(t) = \sin(2\pi ft^2)
\]  

(2)

The model to be identified was a single input Hammerstein system with 4th order polynomial nonlinearity and first order autoregressive filter with a pole near one. Such a filter is common e.g. for temperature models with typically large time constants. To simulate a noisy scenario a random Gaussian noise with maximum amplitude of \( \pm 0.2 \) was added to each model output.
The result in Figure 3 shows the outcome of coefficient estimation depicted as mean squared error in dB over 30 simulation runs. It supports the assumption on the superiority of sinusoidal excitations because the model coefficients can be estimated with a much higher accuracy. The error of coefficient estimation is generally less than for the APRBS sequence.

In order to realize a multichannel sinusoidal excitation in reality it is necessary to stay within the allowed operating limits of the engine, which have to be determined in advance. However, for driving cycle prediction we used a simple method of covering the range of the driving cycle of interest. Figure 4 exemplarily shows the analysis of some input parameter interactions for a new European driving cycle (NEDC) in a two dimensional projection. For simplicity we set the pilot injection to be constant.

A further analysis of the gradients of adjustment within the driving cycles, i.e. the FTP and NEDC, was carried out to estimate the maximum gradient of adjustment for each input. On this basis, the maximum frequency of the excitation chirp signal was derived.

Having estimated the constraints on the input space and on the input gradients, we designed the multi-channel excitation sequence for the six quantities of interest:

- Speed
- Indicated torque
- Boost pressure
- EGR
- SOI
- Rail pressure.

Figure 6 shows a section of the six excitation signals before and after scaling to the driving cycle limits. The dashed line shows the ideal excitation sequence, i.e. without consideration of valid input ranges, whereas the solid line shows the inputs after scaling into the operation point dependent valid range. The speed signal is designed without any constraints and the torque signal has speed dependent limitations, i.e. smaller maximum values for low speed values. The other inputs are scaled into the current speed and torque dependent limits. Note that this approach of operation point dependent scaling introduces some correlation between the inputs. Therefore, a fine-tuning of the lowest chirp frequency and the phase alignment was carried out to achieve lower inter-input correlations.

To cover also idle operation of the engine, an additional idle test plan was designed, having its own limits of the input space and the gradients of adjustment. Along with some step sequences for delay time determination, these test designs were put together in a single sequence to be measured at the test bench.
4. Results

4.1 Measurements

The execution of a dynamic multi-channel test plan is a challenge for every engine testing equipment. Within IAV for this experiment a high dynamic test bench with an up-to-date test automation system was used to vary the six selected quantities: engine speed, indicated torque, boost pressure, EGR-rate, start of injection, and rail pressure. The first two of them are directly controlled by the automation system; the latter four are transferred via a separate ECU control system to the ECU. Using a developer interface to the ECU, an adjustment frequency of 4 Hz could be reached and the six parameters are updated nearly simultaneous within one step. Figure 7 shows a section of the measured excitation sequence; the overall duration of the test plan was about 70 minutes.

In a dynamic measurement setup, it is often necessary to adjust the time alignment of the recorded channels as the different sensors usually have different delay times. Especially the delay of the exhaust gas analyzer signal with respect to the inputs has a major influence.

This delay results from the time the gas needs to flow through the engine up to the analyzer, i.e. the length of the analyzer hose has an impact as well. This delay may be identified by some step experiments. For simplicity, it is here assumed constant even though it depends on the engine mass flow. Besides the pure delay there is an additional influence of the analyzer response time, i.e. the analyzer itself has a transfer function, which may depend on the actual concentration measured.

4.2 Modeling of NOx Emissions

For modeling of the NOx emissions, we used a parametric Volterra series model that was fitted by standard least squares techniques. We set the Finite Impulse Response (FIR) order to six and the polynomial order to three. This resulted in a total number of 250 model coefficients. By means of orthogonal least squares with a threshold of –30dB, we reduced it to a total number of 54 model terms. Figure 8 shows a zoomed-in view of the fitting result where the blue dots represent the measured [ppm] values of NOx and the solid line represents the model output.

For the overall test plan, we achieved a fit RMSE (root mean squared error) of approximately 65 [ppm], which corresponds to a normalized error of 5%. As validation, we used the NEDC, starting from cold conditions and the FTP, starting from warm conditions. The simulation results are shown in Figure 9 and Figure 10, respectively.
4.3 Modeling of Exhaust Temperature

Even though the measurement of the training dataset was carried out under warm engine conditions, i.e. we did not consider any temperature influence, the over-all prediction capability of the NOx model is very good. The validation RMSE for the complete NEDC driving cycle is about 25 [ppm]. Given an overall range of approximately 300 [ppm], this corresponds to a normalized RMSE of 8.3 %. For the FTP driving cycle, we achieved a validation RMSE of 36 [ppm], i.e. approximately 4.3% normalized RMSE.

As expected, the model overestimates the amount of NOx during the first phase of the cold started driving cycle (NEDC), as we did not take any temperature influence into consideration. Therefore, the model can only predict emissions for the warm engine case and performs better for the FTP driving cycle, starting from warm conditions. Figure 11 shows the simulation result of a warm section of the NEDC. Within this segment, we achieved a validation RMSE of 14 [ppm], which corresponds to a normalized RMSE of approximately 6%. For future improvements, the model validity may be expanded by starting the sinusoidal excitation directly from cold conditions.

4.4 Modeling of Engine Mass Flow

In driving cycle simulation the use of [ppm] values is only of limited interest, because for the prediction of accumulated emissions the corresponding mass flow is needed. As the model accuracy for [ppm] values was higher than for [kg/h] values, we built a model for the NOx [ppm] data. For the estimation of emission mass flow, we need an additional model for the engine mass flow. Again, we used a Volterra series with a third order polynomial. The result of fitting and validation against the FTP is shown in Figure 13.
The initial model contained a total number of 111 terms. This was reduced by OLS to 31 terms. We achieved a very high model accuracy with a training RMSE of 2.1 [kg/h], corresponding to a relative error of 0.6 %. The normalized RMSE values of the driving cycle validation both are below 2 %.

4.5 Modeling of NOx mass flow

Having built the models for NOx and engine mass flow, we can simulate the NOx mass flow for the driving cycle of interest. The outcome of this simulation is shown in Figure 14 for the NEDC and as cumulated values in Figure 15 for the FTP cycle. The overall deviation of the simulated cumulated emissions is only 4 % for NEDC and 8 % for FTP.

5. Conclusion

Within this paper, we presented the approach of dynamic engine modeling based on parametric Volterra series and multi-channel sinusoidal excitation sequences with very promising results.

It was possible to dynamically adjust up to six engine parameters on the test bench. Furthermore, the simulation results showed that the use of parametric Volterra series in conjunction with the sinusoidal excitation offers the possibility for a simulation of driving cycles with a high accuracy. Further investigations will be done concerning the prediction of other emissions of interest, like e.g. CO and HC as well as soot.

Currently, the major limitation of the modeling approach is the missing consideration of the engine temperature. For including temperature dependencies, it will be necessary to conduct experiments with cold engine and to examine how these influences can be taken into consideration.

References

(1) Schön, T.: "Estimation of Nonlinear Dynamic Systems – Theory and Applications" Dissertation, Department of Electrical Engineering, Linköpings Universitat, 2006.
(2) Nelles, O.: "Nonlinear System Identification", Springer, Berlin, 2001.
(3) Röpke, K.; Knaak, M.; Nessler, A.; Schaum, S.: "Rapid Measurement – Grundbedatung eines Verbrennungsmotors innerhalb eines Tages?" MTZ 04/2007, Jahrgang 68.
(4) Baumann, W.; Schaum, S.; Röpke, K.; Knaak, M.: "Excitation Signals for Nonlinear Dynamic Modeling of Combustion Engines" IFAC World Conference 2008, Seoul, Korea.
(5) Ljung, Lennart; "System Identification; Theory for the user", Second edition, Prentice Hall PTR, 1999.
(6) Kötter, Hinrich; Schneider, Frank; Fang, Fang; Gußner, Thomas, Isersmarn, Rolf: "Robust Regression and Outlier-Detection for Combustion Engine Measurements”. In: DoE in Engine Development III, expert verlag. 2007

Here we used a simplified formula for the computation of the mass flow quantities with a fixed molar mass of the exhaust gas and a molar mass of 46g/mol for NOx species.