

Narmax model identification of a Variable Geometry Turbocharged diesel engine*

Gianluca Zito†, Ioan Doré Landau
Laboratoire d'Automatique de Grenoble
ENSIEG, BP 46 38402 Saint Martin d'Herès,
France

gianluca.zito@lag.ensieg.inpg.fr,
landau@lag.ensieg.inpg.fr

Abstract—A nonlinear system identification procedure, based on a polynomial NARMAX representation, is applied to a variable geometry turbocharged diesel engine. The relation between the variable geometry turbine (VGT) command and the intake manifold air pressure is described by a nonlinear model, directly identified from raw data. The intent of the paper is to explore the advantages of such a modeling procedure in automotive applications in terms of efficiency and complexity, in view of the related controller design and tuning problem. Simulation results on a HDI diesel engine model illustrate the whole procedure.

Keywords: nonlinear system identification, NARMAX models, VGT diesel engines, automotive applications.

I. INTRODUCTION

Modeling and control design of diesel engines play an important role in the development of the new generation of internal combustion engines. Improvements in overall performances, beside their inherently superior fuel economy, make diesel engines competitive with spark-ignited engines in the class of high performances vehicles.

The tuning of a turbocharged diesel engine is a challenging task for engineers. A standard procedure used in the practice is based on long time spending experimental tests in order to map all possible operative conditions of the engine. As a consequence, controller design and tuning are often developed in an empirical way, as a result of several experiences and of a “try and error” approach on the real system.

The use of simple and efficient models, on the basis of which the control design phase could be easier and faster, is crucial in engine developing, especially for automotive manufacturers. Model-based controller design allows for shorter development times: performances and robustness of control schemes can be evaluated rapidly on a model, drastically reducing the number of calibrations needed on the engine.

Examples of control oriented models of diesel engines, derived from physics principles, are presented in [1], [2], [3]. An adequate seven states mathematical model can be found in [4] with EGR/VGT control perspectives and has

been used in [5]. One should note some fundamental issues to be considered as high nonlinearities present in the engine dynamics, interactions between controlled variables and difficulties to set correct values for the physical parameters. As a consequence, simple control strategies, as PI controllers with parameters depending on the operating points, are widely used in practice to reduce the complexity of the controller tuning problem. Thus, it is often necessary to identify input-output models in order to obtain satisfying results in the controller design phase.

A turbocharger is often used to enhance acceleration performances in diesel engines. Variable geometry turbochargers (VGT) are employed to achieve good boost at all speed conditions, with no lose in terms of efficiency and transient performances.

In this paper a procedure to provide the nonlinear (discrete time) model of the dynamics between the VGT actuator command and the boost pressure in a turbocharged diesel engine directly from raw data is presented. Data are obtained from a complex Simulink model simulating a high pressure direct injection (HDI) diesel engine in which the EGR vane is kept closed (see section II for more details on EGR), as the primary objective of this paper is to analyze the feasibility of the procedure in automotive applications. The effect of the EGR vane will be considered in future works when the procedure will be applied to the full engine model.

Black-box modeling is an attractive alternative to models derived from physics, since it directly provides from data an input-output model to be used for control design and controller tuning. A class of nonlinear models is required for the identification of complex and highly nonlinear systems. A polynomial NARMAX model is chosen to be used in the identification algorithm (model estimation and validation), together with techniques for structure selection which preserve from over-parametrization.

Emphasis must be done to the fact that the model is derived with control purposes, that is, its structure has been conceived for an efficient and high performing diesel engine control design, as a nonlinear pole-placement (see [6]).

* This paper is submitted as regular paper to ACC05.

† Corresponding author.

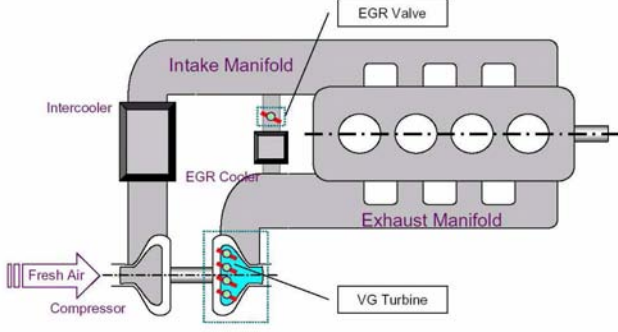


Fig. 1. The VGT/EGR diesel engine.

II. THE VGT TURBOCHARGED DIESEL ENGINE

Common diesel engines are usually turbocharged in order to increase their low power density. A turbine is driven by the exhaust gas from the engine and drives the compressor which supplies the airflow into the engine as in Fig.1. A Variable Geometry Turbocharger (VGT) is used to obtain high transient responses at low engine speeds and to avoid excessive airflow at high engine speeds. A pressure surge in exhaust manifold, in fact, has a detrimental effect for the engine acceleration performances.

The effective flow area of the turbine can be varied by changing the position of the inlet guide vanes on the turbine stator, thereby affecting the compressor mass airflow in the intake manifold. VGT can also act as an emission control mechanism: it affects the pressure drop across the exhaust gas recirculation (EGR) vane (which connects the intake manifold and the exhaust manifold) increasing the exhaust gas recirculation rate. The gas recirculated back into the engine through the EGR vane lowers the flame temperature and avoids the NO_x (oxides of nitrogen) formation.

III. NARMAX SYSTEM IDENTIFICATION

A. NARMAX representation

The NARMAX model formulation was introduced in [7] as an extension for nonlinear systems of the linear ARMAX model, and is defined as

$$y(t) = F(y(t-1), \dots, y(t-n_y), u(t-1), \dots, u(t-n_u), e(t-1), \dots, e(t-n_e)) + e(t) \quad (1)$$

where $y(t)$, $u(t)$ and $e(t)$ represent the output, the input and the system noise signals respectively; n_y , n_u and n_e are the associate maximum lags and $F(\cdot)$ is a nonlinear function.

The NARMAX representation is a well-known tool for nonlinear modeling which includes several other nonlinear representations such as block-structured models and Volterra series. This class of models has the appealing feature to be linear-in-the-parameters, so that a straight implementation of least-squares techniques can be applied.

Expanding $F(\cdot)$ in (1) as a polynomial of degree L (where L is the degree of the nonlinearity) the expression

of a polynomial NARMAX model is obtained as follows

$$y(t) = \sum_{i=1}^n \theta_i x_i(t) + e(t) \quad (2)$$

where

$$n = \sum_{i=0}^L n_i, \quad n_0 = 1 \quad (3)$$

$$n_i = n_{i-1} \frac{(n_y + n_u + n_e + i - 1)}{i}, \quad i = 1 \dots L$$

and

θ_i = i th model parameter

$$x_1(t) = 1$$

$$x_i(t) = \prod_{j=1}^p y(t - n_{y_j}) \prod_{k=1}^q u(t - n_{u_k}) \prod_{m=1}^r e(t - n_{e_m}) \quad (4)$$

$$i = 2, \dots, n, \quad p, q, r \geq 0, \quad 1 \leq p + q + r \leq L \quad (5)$$

$$1 \leq n_{y_j} \leq n_y, \quad 1 \leq n_{u_k} \leq n_u, \quad 1 \leq n_{e_m} \leq n_e \quad (6)$$

The choice of a polynomial expression for the regressor is based on the possibility to derive nonlinear control algorithms for a nonlinear polynomial model as a direct extension of classic linear pole-placement control problem.

B. Input signal design

Input signal design is a very important step for nonlinear system identification. As for the linear case, the input signal should be persistently exciting. All the frequencies of interest for the system should be excited, and the input signal should cover the whole range of operation. A simple and effective implementation is realized by means of a concatenated set of small-signal tests. Small amplitude perturbing signals may be superposed to the different operating levels, exciting all dynamic modes of the system. Increasing and decreasing level amplitudes have to be considered in order to take into account direction dependent dynamics.

Different classes of signals can be employed for the identification process as multi-sine signals, maximum length binary sequences (MLBS) and classic pseudo-random signals. Documentation about identification signal design can be found in [8], [9].

C. Structure selection

Structure selection is a key problem in a black-box system identification. A survey of the structure identification methods is in [10], and an overview on the different approaches to nonlinear black-box modeling is in [11]. When the system to identify is nonlinear a direct estimation based on (2) generally leads to an over-parameterized model. If the values of n_y , n_u , n_e and L are increased to obtain a good accuracy, an excessively complex model will result together with a numerical ill-conditioning. A procedure is needed to select terms from the large set of candidates to provide a

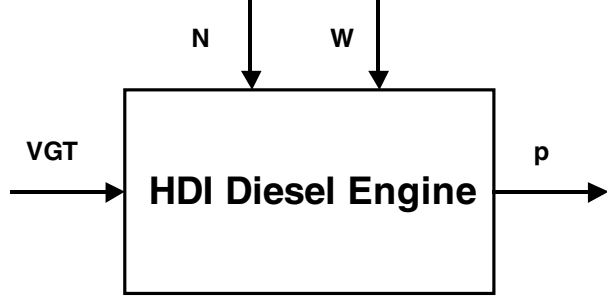


Fig. 2. Equivalent HDI diesel engine scheme for identification.

parsimonious model. A simple and effective procedure is based on error reduction ratio (ERR) defined in [12] as

$$ERR_i = \frac{g_i^2 \sum_{k=1}^N w_i^2(t)}{\sum_{k=1}^N y_i^2(t)} \quad (7)$$

where $g_i(k)$ are the parameters and $w_i(k)$ the regressors of an auxiliary model constructed to be orthogonal over the data records:

$$y(t) = \sum_{i=1}^n g_i w_i(t) + e(t) \quad (8)$$

A model is found selecting the relevant terms from the full model set following a forward-regression algorithm (for more details see [13]): at each step the parameter with the highest ERR_i is added to the current model, following the principle that a parameter which reduces the variance more than the others is more important. An information criterion, could be used to stop the procedure, as the Akaike Information Criterion [14], defined as

$$AIC = N \log_e(\sigma_\epsilon^2(\theta_p) + kp) \quad (9)$$

where σ_ϵ^2 is the variance associated to the p -terms model and k is a penalizing factor. At the end of the selection process, a recursive identification is run with the selected parameters. Several techniques have been proposed in the literature for selecting the best model structure, some of these are enhancements of the ERR algorithm or are used in conjunction with it as in [15], [16].

D. Model validation

A statistical validation of the identified NARMAX model is performed with high order correlation functions defined in [17], [18] to detect the presence of unmodelled terms in the residuals of the nonlinear model. If the identified model is adequate, the following conditions should be satisfied by the prediction errors

$$\begin{aligned} \Phi_{\epsilon\epsilon}(k) &= \delta(k) && \text{(i.e. an impulse)} \\ \Phi_{u\epsilon}(k) &= 0 && \forall k \\ \Phi_{\epsilon(\epsilon u)}(k) &= 0 && k \geq 0 \\ \Phi_{u^2 \epsilon'}(k) &= 0 && \forall k \\ \Phi_{u^2 \epsilon' \epsilon'}(k) &= 0 && \forall k \end{aligned} \quad (10)$$

TABLE I

DIESEL ENGINE OPERATING POINTS: FULL ACCELERATION.

Speed engine (rpm)	Air mass flow (mm ³ /cp)
1000	45
1250	58.2
1500	64.75
1750	68.3
2000	72.31
2250	66.92
2500	66.37
2750	67.3
3000	66.7
3250	63.11
3500	62.11
3750	61.14
4000	60.95
4250	56.53
4500	52

TABLE II

DIESEL ENGINE OPERATING POINTS: 50% ACCELERATION.

Speed engine (rpm)	Air mass flow (mm ³ /cp)
1000	23.68
1250	30.63
1500	34.3
1750	35.94
2000	37.7
2250	35.22
2500	35.8
2750	35.42
3000	35.1
3250	33.21
3500	32.69
3750	32.18
4000	32.08
4250	29.75
4500	27.37

where $\Phi_{xy(k)}$ indicates the cross-correlation function between $x(t)$ and $y(t)$, $\delta(k)$ is the Kronecker delta, $\overline{u^2(t)}$ is the mean value of $u^2(t)$ and $u^{2'}(t) = u^2(t) - \overline{u^2(t)}$. If at least one of the correlation functions is well outside the confidence limits, a new model has to be identified. It is necessary, in order to check the ability of the model to represent system dynamics, to validate the estimated model on a new set of data (*validation data*) different from the set used for the identification (*learning data*).

Model prediction ability has to be assessed, together with statistical tests, with signals that may catch system nonlinearities. Triangular or step signals of different amplitude levels are ideal input signals used for time-domain model validation.

IV. SIMULATION RESULTS

A. Simulation setup

The identification algorithm presented in the previous sections is applied to a HDI diesel engine model simulated with The MathWorks Simulink environment. The mechanical and thermodynamic interactions between the variables

TABLE III
ENGINE PARAMETERS AND VARIABLES

p	pressure (mbar)
VGT	variable geometry turbocharger signal command (%)
N	engine speed (rpm)
W	air mass flow (mm^3/cp)

describing the engine operation are modelled with algebraic and differential equations, and with lookup tables recovered by real time experiments. Thus, the model is a low level description of the system showed in Fig.1 and, providing a close approximation of the real system, the nonlinear relation between the VGT signal command and the intake manifold air pressure (MAP) can be investigated in a large set of operative conditions.

For identification purposes the system could be seen as a SISO nonlinear black-box, as shown in Fig.2. The input (VGT) to the system is the command of the actuator that adjusts the angle of guide vanes placed to vary the incoming exhaust gas flow at the entrance of the turbine. The output (p) is the air pressure measured at the intake manifold (boost pressure). N and W are the speed engine and the air mass flow, respectively: a model is identified around a operating point defined by the pair (N, W) .

The identification algorithm is feeded with input-output data sets generated from several simulations in order to find a polynomial NARMAX model of the VGT -boost pressure nonlinear relation for different pairs (N, W) , that specify the operative conditions of interest for the engine. Tables I and II resume all the different operating points for a full and 50% driver acceleration.

B. Excitation signal design

The signal used for the identification is, for all the operating points, a concatenated data set of small signals. A sequence of increasing and decreasing steps describes the different regions of the VGT command, and small amplitude (10% of the corresponding step) multisine signals are superposed as excitation signals covering a frequency range from 0 up to 2 Hz. Fig. 4, for example, shows the data set for the operating point defined by the pair $(N, W) = (3000 \text{ rpm}, 66.7 \text{ mm}^3/\text{cp})$, and a full driver acceleration. The VGT command is in the range 20%–65%, covered by a sequence of steps with an increasing/decreasing variation $\Delta = 5\%$ and superposed multi-sine signals.

C. VGT-boost pressure Model identification

The forward-regression estimation algorithm is applied to the data related to the pair $(N, W) = (3000 \text{ rpm}, 64 \text{ mm}^3/\text{cp})$. The first choice for the parameters n_y , n_u and L is based on step responses analysis to estimate dynamics and nonlinearity orders. Tests for nonlinearity detection are presented in [19].

A general inspection reveals that a linear second order system is a good representation for small variations of the

TABLE IV
NARMAX PARAMETERS.

Index selected	Parameter value	Model term
1	1902.2	constant
2	-0.52096	$y(t-1)$
3	0.013717	$y(t-2)$
4	6.2607	$u(t-1)$
5	1.6462	$u(t-2)$
6	9.7052	$u(t-3)$
7	0.00019272	$y^2(t-1)$
10	0.14749	$u^2(t-1)$
12	-0.40762	$u(t-1)u(t-3)$
15	0.1361	$u^2(t-3)$

input and of the output. This means that the global nonlinear discrete time model, after a linearization, should provide a second order discrete time system. Thus, a model with $n_y = 2$, $n_u = 3$ and $L = 2$ is identified, and details about the parameters are given in table IV.

This procedure, iterated for all the pairs (N_i, W_i) , where i is the generic operating point, leads to a set of nonlinear models that describes the diesel engine boost pressure as a nonlinear discrete time difference equation of the variables VGT , N and W . Thereby, (2) can be parameterized as

$$y(t) = \sum_{i=1}^n \theta_i(N, W)x_i(t) + e(t) \quad (11)$$

Each operating point has an associated nonlinear model of low complexity: for example, model in table IV contains 10 parameters of the 21-terms full model. On the basis of this model efficient but still robust nonlinear control algorithms can be directly applied.

D. VGT-boost pressure Model validation

Statistical and time-domain validations are employed to assess the model quality. Good results for the statistical validation (10) are obtained (see Fig.3). Fig.5 and Fig.6 show model long-term prediction with *validation data* and step model validation with small and high amplitude data, respectively. In these last two cases a step-sequence is applied to the identified model to verify that, for small and large variations in the input signal, the system output is matched from the nonlinear NARMAX model output. The first step sequence is the same used to sweep input amplitude range in the identification data acquisition ($\Delta = 5\%$), in the second one a larger amplitude variation is applied ($\Delta = 15\%$). This typical engine test confirm that the model is suitable to represent system dynamics in both input direction.

V. CONCLUSIONS

Control oriented models for diesel engines are necessary for an efficient tuning of controllers. A practical solution to the nonlinear modeling problem in automotive applications is represented by a nonlinear black-box identification. Polynomial NARMAX models constitute an interesting class of

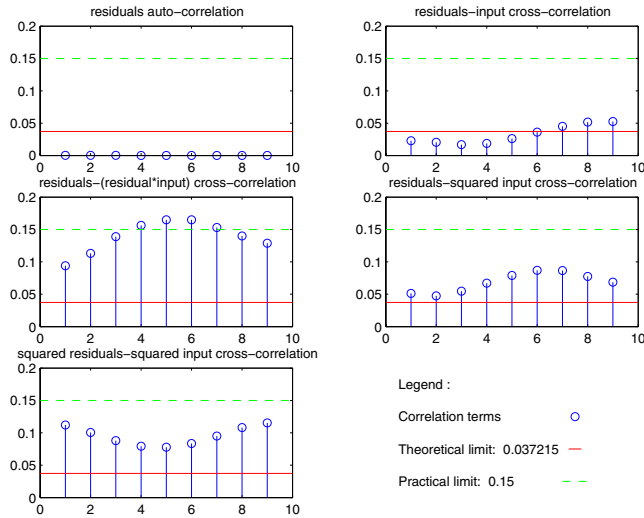


Fig. 3. Statistical validation for the Data Set corresponding to $(N,W)=(3000 \text{ rpm}, 64 \text{ mm}^3/\text{cp})$: correlation values (o), theoretical limit (value = $2.17/\sqrt{\text{Data length}}$, solid line), practical limit (value = 0.15, dashed line).

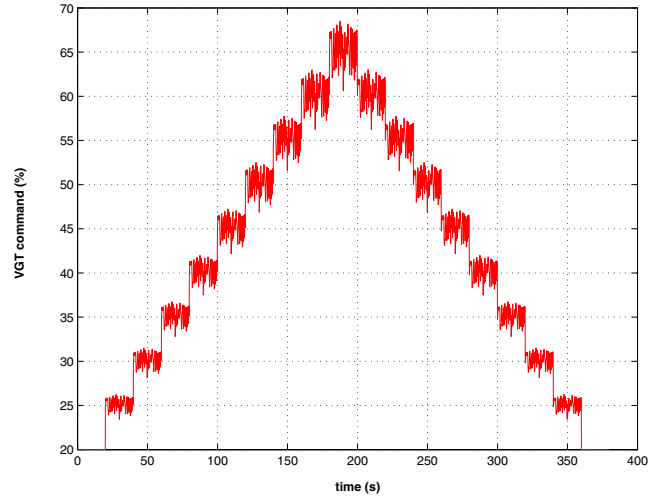
input-output models for describing a large set of nonlinear systems, as they are able to capture nonlinear dynamics and, at the same time, they can be efficiently used together with structure selection and parameters estimation procedures. This drastically reduces the time for the elaboration of a control oriented model. In this paper a practical identification procedure based on a polynomial NARMAX representation has been developed and applied to a HDI diesel engine case study. Parsimonious nonlinear models have been derived in view of nonlinear control algorithms implementation.

VI. ACKNOWLEDGMENTS

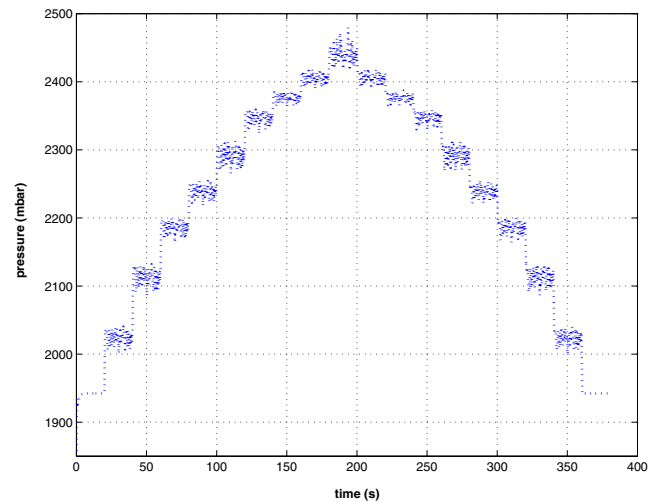
The work of Gianluca Zito is supported by a Marie Curie Industry host Fellowship of the European Community.

REFERENCES

- [1] L. Guzzella and A. Amstutz, "Control of diesel engines," *IEEE Control System Magazine*, vol. 18, pp. 53–71, Oct. 1998.
- [2] M. Jankovic, M. Jankovic, and I. Kolmanovsky, "Robust nonlinear controller for turbocharged diesel engines," *IEEE Transactions on Control System Technology*, vol. 8, Mar. 2000.
- [3] M. Kao and J. J. Moskwa, "Turbocharged diesel engine modelling for nonlinear engine control and estimation," *ASME Journal of Dynamic Systems, Measurement, and Control*, vol. 117, 1995.
- [4] I. Kolmanovsky, P. Moraal, M. van Nieuwstadt, and A. G. Stefanopoulou, "Issues in Modelling and Control of Intake Flow in Variable Geometry Turbocharged Engines," in *Proceedings of the COSY Workshop ECC 97*, Brussels, Belgium, 1997, pp. 1990–1995.
- [5] M. J. van Nieuwstadt, I. Kolmanovsky, P. Moraal, A. G. Stefanopoulou, and M. Jankovic, "EGR-VGT Control Schemes: Experimental Comparison for a High-Speed Diesel engine," *Control System Magazine*, vol. 20, no. 3, pp. 63–79, 2000.
- [6] I. D. Landau, D. Normand-Cyrot, and A. Montano, "Adaptive control of a class of nonlinear discrete time systems Application to a heat exchanger," in *Proceedings of the 26th Conference on Decision and Control*, Los Angeles, CA, 1987, pp. 1990–1995.
- [7] I. J. Leontaris and S. A. Billings, "Input-Output parametric models for non-linear systems—Part 1: Deterministic non-linear systems; Part 2: Stochastic non-linear systems," *International Journal of Control*, vol. 41, pp. 303–344, 1987.
- [8] M. R. Schroeder, "Synthesis of low peak-factor signals and binary sequences of low auto-correlation."
- [9] K. Godfrey, *Perturbation Signals for System Identification*. New York, NY: Prentice Hall, 1993.
- [10] R. Haber and H. Unbehauen, "Structure Identification of Nonlinear Dynamic Systems—A survey on Input/Output Approaches," *Automatica*, vol. 26, no. 4, pp. 651–677, 1990.

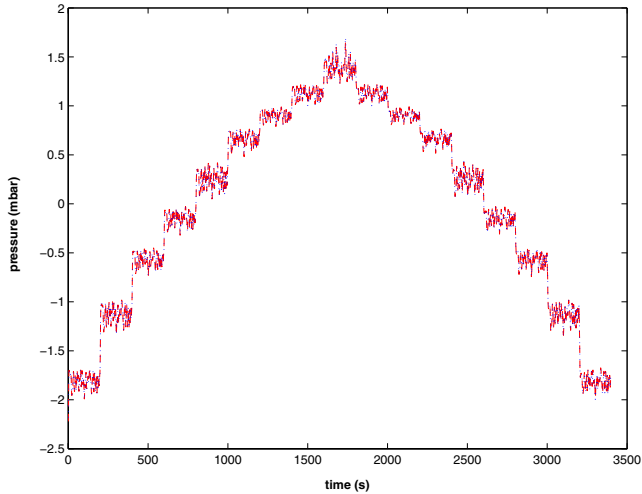


a)

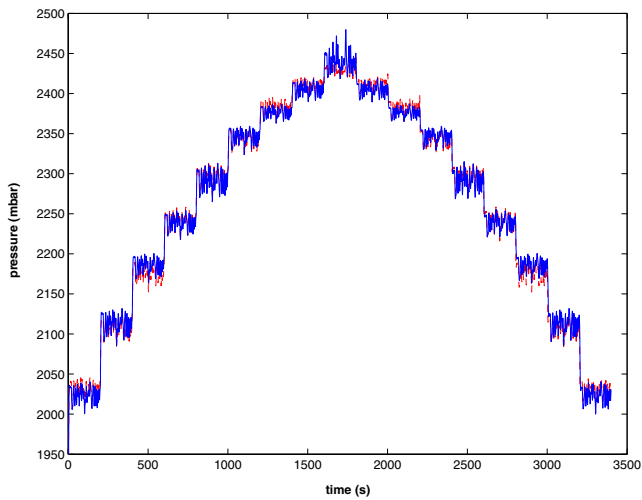


b)

Fig. 4. Data Set for $(N,W)=(3000 \text{ rpm}, 64 \text{ mm}^3/\text{cp})$: a) VGT command, b) boost pressure.

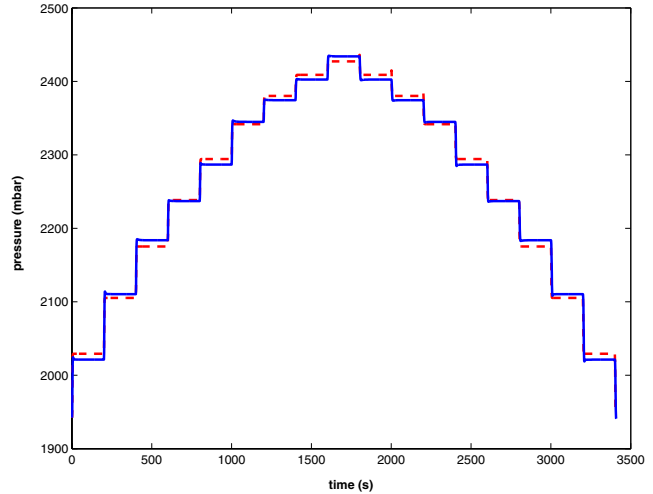


a)

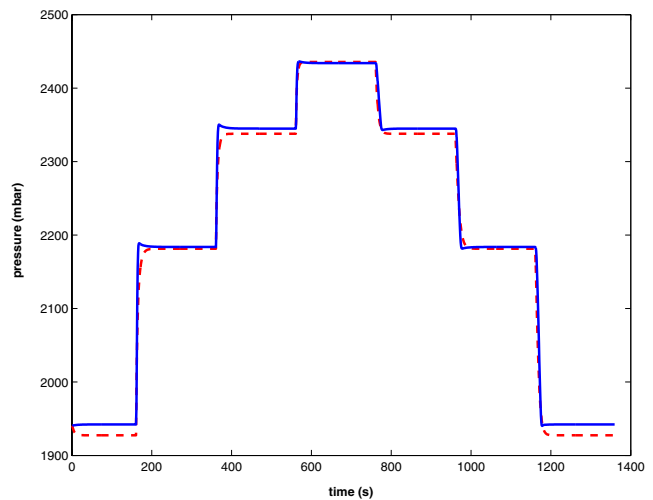


b)

Fig. 5. Model validation for $(N,W)=(3000 \text{ rpm}, 64 \text{ mm}^3/\text{cp})$: model prediction (dashed line), system output (solid line); a) 1-step-ahead predictor output (standardized data); b) long-term predictor output.



a)



b)

Fig. 6. Model and system step responses for $(N,W)=(3000 \text{ rpm}, 64 \text{ mm}^3/\text{cp})$: a) small amplitude; b) high amplitude; model output (dashed line), system output (solid line).

- [11] J. Sjöberg, Q. Zhang, L. Ljung, A. Benveniste, P. G. B. Delyon, H. Hjalmarrsson, and A. Juditsky, "Nonlinear Black-box Modeling in System Identification: a Unified Overview," *Automatica*, vol. 31, no. 12, pp. 1691–1724, 1995.
- [12] S. A. Billings, S. Chen, and M. J. Korenberg, "Identification of MIMO nonlinear systems using a forward-regression orthogonal estimator," *International Journal of Control*, vol. 49, no. 6, pp. 2157–2189, 1989.
- [13] S. A. Billings and S. Chen, "Extended model set, global data and threshold model identification for severely nonlinear systems," *International Journal of Control*, vol. 50, no. 6, pp. 1897–1923, 1989.
- [14] H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 716–723, 1974.
- [15] L. A. Aguirre and S. A. Billings, "Improved structure selection for nonlinear models based on term clustering," *International Journal of Control*, vol. 62, no. 3, pp. 569–587, 1995.
- [16] L. Piroddi and W. Spinelli, "An identification algorithm for polynomial NARX model based on simulation error minimization," *International Journal of Control*, vol. 76, no. 17, pp. 1767–1781, 2003.
- [17] S. A. Billings and W. S. F. Voon, "Correlation based model validity tests for non-linear models," *International Journal of Control*, vol. 44, no. 1, pp. 235–244, 1986.
- [18] S. A. Billings and Q. M. Zhu, "Nonlinear model validation using correlation tests," *International Journal of Control*, vol. 60, no. 6, pp. 1107–1120, 1994.
- [19] R. Haber, "Nonlinearity tests for dynamic processes," in *IFAC Identification and System Parameter Estimation*, York, UK, 1985, pp. 409–413.