

# Parameter Identifiability and Parameter Estimation of a Diesel Engine Combustion Model

Lilianne Denis-Vidal\*, Zohra Cherfi, Vincent Talon, El Hassane Brahmi

University of Technologie of Compiègne France, LMAC , ROBERVAL Compiègne, France  
Email: \*[lilianne.denis-vidal@math.univ-lille1.fr](mailto:lilianne.denis-vidal@math.univ-lille1.fr), [zohra.cherfi@utc.fr](mailto:zohra.cherfi@utc.fr), [vincent.talon@renault.fr](mailto:vincent.talon@renault.fr)

Received December 2013

---

## Abstract

In this paper an original method based on the link between a piecewise identifiability analysis and a piecewise numerical estimation is presented for estimating parameters of a phenomenological diesel engine combustion model. This model is used for design, validation and pre-tuning of engine control laws. A cascade algebro-differential elimination method is used for studying identifiability. This investigation is done by using input-output-parameter relationship. Then these relations are transformed by using iterated integration. They are combined with an original numerical derivative estimation based on distribution theory which gives explicit point-wise derivative estimation formulas for each given order. Then new approximate relations, linking block of parameters and outputs (without derivative) are obtained. These relations are linear relatively to the blocks of parameters and yield a first estimation of parameters which is used as initial guess for a local optimization method (least square method and a local search genetic algorithm).

## Keywords

Calibration, Genetic Algorithms, Global Optimization, Identifiability, Identification Algorithms, Internal Combustion Engines

---

## 1. Introduction

The use of model to predict system behavior in order to make technical choices or to understand its functioning has become very important during the last decade. This is particularly true in automotive industry, especially because of the increase of emissions constraints and the will to optimize fuel consumption. A phenomenological model has been used. Such models have the main advantage to be quite simple and directly linked with observable phenomenon. In this paper a Wiebe double phase combustion model is proposed and investigated. On the other hand their use is effective only if they were beforehand calibrated.

The main advantage of this mathematical model is its execution speed. Thus, it helps to carry out real-time simulations as for HIL (Hardware In the Loop). But its lack of physics makes impossible any form of explora-

---

\*Corresponding author.

tion out of its training zone. Some of its parameters must be estimated [1].

Before performing parameters estimation, it is necessary to investigate their identifiability. Identifiability falls into the category of inverse problem. It is a priori question: do two distinct parameters give distinct output?

In case of negative answer, one will not be able to estimate it except by some additional a priori knowledge: the model has to be reconsidered. There are three main approaches for studying identifiability of controlled and uncontrolled non linear dynamical systems described by differential systems. The first one is based on equivalence theory [2,3], the second one considers the Taylor series expansion [4], and the last one comes from elimination theory [5]. This last theory gives a possible procedure for estimating parameters without any information about them. Few studies use it efficiently [6]. The differential elimination method necessitates model with smooth function which is not the case of our model and more some output derivatives of high order may be estimated. A piecewise differential elimination is proposed in order to take into account the non smooth functions which appear in our model. Then it is shown how the resulting calculus can be used to build a parameter estimation procedure for this type of model.

The paper is organized as follows. In the second section the phenomenological model is presented as well as the resulting mathematical system. The third section is devoted to identifiability: some definitions are recalled, piecewise identifiability tests and results are given. The fourth section treats the transformation of the relations obtained from identifiability previous study. Section 5 is devoted to numerical procedure: the numerical estimation of derivatives and the resulting estimation procedure and numerical results are given. Finally, conclusions are drawn in the last section.

## 2. The Diesel Combustion Model

### 2.1. The Phenomenological Model

The Wiebe model with double phases facilitates the modeling of traditional diesel combustions. It is based on first thermodynamics principle. The internal energy variation follows the equation:

$$dU / d\theta = dQ + dW + \sum h_i dm_i$$

with

U: internal energy [J],  $\theta$ : crankshaft degree, dQ: energy exchange with outside [J/s], dW: work with outside [J/s]

$h_i$ : specific enthalpy [J/kg],  $dm_i$ : mass flow with outside [kg/s]

In the case of the combustion chamber we can identify:

- The energy exchanges with outside are equal to losses on the walls and the contribution of combustion energy.
- The work is the one of the piston on gases.
- The exchanges of matters are mass flows through the exhaust and intake valves.

The Wiebe phases represent the combustion of pre mixture and diffusion. These two combustions succeed one another during one engine cycle. Below we describe shortly process of diesel combustion.

After the intake and gas compression, fuel is introduced into the combustion chamber. This fuel is believed in vapor form in the case of Wiebe. The compression overheats the fuel. This heating will cause its ignition. This starts the two flames of combustion of pre mixture and diffusion.

The first phase represents the flame of pre mixture: at this stage, air and fuel are perfectly mixed. When the time of ignition is reached, a mass combustion occurs. At this stage there is a multitude of points of ignition in the combustion chamber. We say that the combustion is controlled by chemistry.

The second phase represents the diffusion flame: in this case chemistry becomes infinitely fast because of the rise of temperature due to the pre mixture flame. The combustion is then controlled by the speed of fusion of the air and the fuel. That is why we call it diffusion flame.

The Wiebe model thus makes possible to model these two flames. We need to associate to this model of heat release, additionally work of the piston on gases and the exchanges with the walls:

- Work is in traditional form—P.dV, where P is the pressure in the chamber and dV is the volume variation.
- Exchanges with the walls are in form  $h_w.S.(T - T_w)$ , where H, S, T and  $T_w$  are respectively, the coefficient of

exchange, the surface of exchange, the temperature of gases in the combustion chamber and the temperature of the walls.

With internal energy (U) and mass (M) in the chamber, we can calculate pressure and temperature of gases.

This temperature is calculated with the first law of Joules  $U = C_n M T$

with  $C_n$  = heat capacity for constant volume [J/kg/K],  $T$  = temperature [K].

The pressure is calculated with the law of perfect gases:  $P = (M r T) / V$

where,  $P$  = pressure [Pa],  $r$  = perfect gases constant [J/kg/K] and  $V$  = volume [m<sup>3</sup>].

## 2.2. Mathematical Model

This model can be described in state space form as follows

$$S(p, \nu): dX / d\theta = F(X, \theta, p) + C\nu(\theta), \text{ and } X(\theta_0) = X_0, Y = h(X, \theta)$$

Using the notations of the previous subsection, this system is explained as follows.

$X$  denotes the state vector in the model described in the previous section,  $X = (M, U)'$ ,  $Y$  denotes the output and  $Y = X$  in our problem,  $\theta \in [0, \Theta]$ ,  $p = (p_1, \dots, p_l)$  denotes the vector of parameters, each parameter is a positive number. In the case of the phenomenological model the vector of parameters is given by:

$$P = (\alpha_{pp}, f_{pre}, \Delta_p, m_p, \Delta_d, m_d) \text{ and } \nu \text{ is a square wave input,}$$

$\nu = Q_{inj} \max(H(-ADV) - H(\theta - ADV - \theta_{inj}))$  with  $Q_{injmax} > 0$  and  $\theta_{inj} > 0$ ,  $\Omega_{ad}$  denotes the set of such inputs.  $H$  is the well known Heaviside function:

$$H(x) = 1 \text{ if } x > 0 \text{ and } H(x) = 0 \text{ if } x < 0$$

Putting:

- $C = (1, C_{carb}, T_{carb-Lvap})'$ ,  $F = (F1, F2)'$ , with:  $F1(X, \theta, p) = 0$  and  $F2(X, \theta, p) = k1(X) - Q_{paroi}(\alpha_{pp}, X) + (f_{pre})f1(p, \theta)H(\theta - ADV_p) + (1 - f_{pre})f2(p, \theta)H(\theta - ADV_d)$

where

- $k1(X) = (\gamma - 1)(-U/V)(dV/d\theta)$  is a known function,  $\gamma$  is the constant of adiabatic compression
- $K = PCI_{Minj}$ ,  $PCI$  and  $Minj$  are known constants,  $f1(p, \theta)$  and  $f2(p, \theta)$  are the derivatives relatively to  $\theta$  of

$$\text{Exp}[-(\alpha_p / \Delta p^{1+mp}) (\theta - ADV_p)^{1+mp}] \text{ et } \text{Exp}[-(\alpha_d / \Delta d^{1+md}) (\theta - ADV_d)^{1+md}]$$

- $Q_{paroi}(\alpha_{pp}, X) = \alpha_{pp} S_p k2(X)$  where  $S_p$  is a known constant. and  $k2(X)$  is a known function and  $\alpha_{pp}$  is a parameter to be identified.

In this model, we assume that: H1:  $0 < ADV < ADV_p < ADV_d < \Theta$  and H2:  $k2(X)$  not equal to 0.

## 3. Identifiability

### 3.1. Definitions

Now some useful definitions are recalled. In the following the vector of parameter is denoted  $p = (p_1, \dots, p_l)$

#### Definition1

The parameter  $p_i$  is said to be globally identifiable if, for any  $(p_h, p_t)$  and  $p_i$  not equal to  $p_i$  there exists an input such that the systems  $S(p_h, \nu)$  and  $S(p_t, \nu)$  yield different outputs.

#### Definition2

The model is said to be globally identifiable if all parameters are globally identifiable.

### 3.2. Identifiability Analysis

In this section, an identifiability study is proposed by using piecewise relations linking successive known functions and parameters. The main identifiability results are given in the following.

**Theorem1** The diesel combustion model  $S(p, \nu)$  is globally identifiable.

This result is a consequence of the following proposition.

**Proposition1** Input-output-parameter relations are given successively by:

- 1) If  $\theta \in [0, ADV[$

$$-\frac{dU}{d\theta}(\theta, p) = k_1(U(\theta, p)) - \alpha_{pp} S_p k_2(M(\theta, p), U(\theta, p)) \quad (\text{ES1})$$

2) If  $\theta \in ]ADV_p, ADV_d[$

$$\left[\left(\frac{\partial y_1}{\partial \theta}\right)^2 - y_1 \frac{\partial^2 y_1}{\partial \theta^2}\right](\theta - ADV_p) = -m_p \frac{\partial y_1}{\partial \theta} y_1 - f_{pre} \frac{\partial^2 y_1}{\partial \theta^2} [\theta - ADV_p] + m_p f_{pre} \frac{\partial y_1}{\partial \theta} \quad (\text{ES2})$$

$$-\frac{\partial y_1}{\partial \theta} = \frac{-\alpha_p}{\Delta \theta_p^{m_p+1}} z_2 (m_p + 1) (y_1 + f_{pre}) \quad (\text{ES3})$$

3) If  $\theta \in ]ADV_d, \min(\Theta, ADV + \theta_{inj})[$

$$-\left[\left(\frac{\partial y_2}{\partial \theta}\right)^2 + (1 - y_2 - f_{pre}) \frac{\partial^2 y_2}{\partial \theta^2}\right](\theta - ADV_d) = -m_d \frac{\partial y_2}{\partial \theta} (y_2 - 1 - f_{pre}) \quad (\text{ES4})$$

$$-\frac{\partial y_2}{\partial \theta} = \frac{-\alpha_d}{\Delta \theta_d^{m_d+1}} z_4 (m_d + 1) (y_2 + f_{pre} - 1) \quad (\text{ES5})$$

where:

$$-y_1(M, U, p) = \frac{dU}{d\theta} - K_1(U) - \alpha_{pp} S_p k_2(M, U) - Q_{inj \max} (C_{pcarb} T_{carb} - L_{vap})$$

$$-y_2(M, U, p) = y_1(M, U, p) + K \omega f_{pre} f_1(p, \theta)$$

$$-z_2(\theta, p) = (\theta - ADV_p)^{m_p} \text{ and } z_4 = (\theta - ADV_d)^{m_d}$$

For the proof of this proposition a piecewise differential elimination is performed and successively input-output parameters relations are obtained and then global identifiability of parameters and model.

## 4. Iterated Integration, Input-Output-Parameter Relations and Derivatives Estimation

The main objective of this paper is to give an efficient procedure for a numerical estimation of parameters. This procedure begins with input-output-parameter relations. Often these relations require the knowledge of derivatives of observed signal. The observations are noisy and estimation of derivatives becomes an ill posed problem. In order to overcome this problem, this paper proposes a new computation algorithm for determining the parameters with noisy observed signal. This algorithm is based on two steps described in next sections.

### 4.1. Iterated Integration and Input-Output-Parameter Relations

The first step consists in integrating as far as possible the input-output-parameters relations in order to remove derivatives and by using integration by parts and iterated integrals using the well known **Cauchy's formula**:

$$\int_a^x dt \dots \int_a^t f(u) du = \int_a^x f(t) \frac{(x-t)^{i-1}}{(i-1)!} dt,$$

This leads to integro-differential relations which are used to build our estimation procedure. However, in these relations, some first order derivatives remain. Then, derivatives have to be estimated. This approach has been developed in [6].

### 4.2. A Method for Estimating Derivatives

#### Description of the method

For estimating the derivatives an original method is given in [6]. It is based on distribution theory and the

main tools are the test functions which are infinitely differentiable functions with compact support.

In this paper the test functions are deduced from the following function with support in  $[-1;1]$ ,

$$\psi(x) = \begin{cases} \frac{1}{2} e^{\frac{1}{2} - |x|} & |x| < 1 \\ 0 & |x| \geq 1 \end{cases}$$

A translation and a dilatation allow obtaining an infinitely differentiable function  $\psi_{I_\varepsilon}$  with a support included in a given interval  $I_\varepsilon$ . They play the role of sliding windows. The method is based on the combination of a truncated Taylor expansion of the signal and test functions. The use of the truncated Taylor expansion for the derivatives estimation is classical (finite difference, etc...).

The difference is the way to use it. Let us consider  $y$  an analytic function in the interval  $[a,b] \subset \mathbb{R}$ . Such that the length of  $I_\varepsilon$  is  $2\varepsilon$  the truncated Taylor expansion  $y_N$  of  $y$  is an approximation of  $y$ . It is given by:

$$y_N(\theta) = \sum_0^N y^{(n)}(\theta_0) \frac{(\theta - \theta_0)^n}{n!} \quad \text{and then} \quad y_N^{(k)}(\theta) = \sum_{n=k}^N y^{(n)}(\theta_0) \frac{(\theta - \theta_0)^{n-k}}{(n-k)!} (E_k) \quad \text{for } k = 0 \dots N.$$

Multiplying this equality by  $\psi_{I_\varepsilon}$  and integrating on  $I_\varepsilon$ , because the support of  $\psi_{I_\varepsilon}$  is included in  $I_\varepsilon$  and by using integration by part, the following linear system triangular is obtained

$$\sum_{n=k}^N y^{(n)}(\theta_0) \alpha_{n-k} = (-1)^k \int_{I_\varepsilon} y_N(t) \psi_{I_\varepsilon}^{(k)}(t) dt \quad \text{for } k = 0 \dots N. \quad (\text{Ss})$$

The unknown are  $y^{(k)}(\theta_0)$  for  $k = 0 \dots N$ , and the determinant of the system is  $\alpha_0^N$  which is not equal to zero. Then  $y^{(k)}(\theta_0)$  for  $k = 0 \dots N$  can be obtained as solution of this system. Replacing  $y_N$  by  $y$  in (Ss) a numerical estimation  $y_{\varepsilon,N}^k(\theta_0)$  of  $y^{(k)}(\theta_0)$  is obtained.

#### Error due to the noise on the signal

In [7], the authors bring out two kinds of perturbations. The first ones, like the constant perturbations whose amplitude is unknown are said structured. They can be annihilated by linear differential operators and it is, for example, the case of a random noise whose mean is constant and unknown. The second ones are said unstructured and correspond to a high frequency perturbation which can be reduced by low pass filters like iterative integrals. Thus, any noise can be written as the sum of such noises. The method based on the distribution theory allows annihilating a part of the structured noises. On the other hand the integration reduces the effect of unstructured noise which corresponds to high frequency perturbations.

## 5. Methods for Estimating Wiebe Model Parameters and Numerical Results

In numerical study, the cylinder pressure  $P$  has been used instead of the internal energy  $U$  because the observed data in our possession are those of the cylinder pressure  $P$ .

### 5.1. Description of the Methods

Two methods have been set up. Each method consists in 2 steps. The first one is the same for both approaches.

#### Estimation and integro-differential equations

This first step is based on successive global optimizations which are made on successive intervals.

Firstly we focus on the parameter  $P = (\alpha_{pp}, f_{pre}, \Delta_p, mp, \Delta_d, md)$ . The **exact** observation, corresponding to the **exact** value of  $p$ , is more or less perturbed by noise. So, integro differential equations are not satisfied by the noisy output. Moreover, the observation is not carried out at any time and the data are a set of measures:

$$Pk,mes = P(tk,pex) + \varepsilon k, \quad \text{for } k = 1 \dots q$$

$\varepsilon k$ , for each  $k$ , is assumed to be a normal random noise with zero mean. Each integro-differential equation corresponds to an integro-differential polynomial. The outputs are noisy, and are carried at each discretization time  $tk$  for each polynomial and the derivatives are estimated, therefore a least square minimization problem is performed. This step gives us a first estimation. But the crankshaft degree interval:  $[ADVp, ADVd]$  is small and the number of discretization points is also small. Consequently the estimations of some integrals are not accurate enough and then, some parameters are not well estimated: the pressure reconstruction is not satisfactory. This fact leads us to implement local optimization methods to improve the accuracy. The first estimation obtained

with this first step can be used as initial guess for a local optimization. In the following it is denoted  $p_0$ .

Local optimization

Having studied sensitive analysis, we noticed that the thermodynamic quantity  $dQ_{cyl}/d\theta$  is sensitive to some parameters while the pressure is less. This makes necessary the joint use of these two quantities.

More, the cylinder pressure  $P$  and the heat release  $dQ_{cyl}/d\theta$  are two thermodynamic quantities of different nature. In order to perform the local optimization, the system  $S(p, \nu)$  is used and not input output integro-differential equations. The first step of the methods doesn't allow the use of the heat release: this quantity don't appear in the input-output relations. These two thermodynamic quantities can be used for local optimization by considering an augmented system with the classical link between cylinder pressure and heat release.

Now, we will describe precisely the step 2 corresponding to each of the two methods.

First method, step 2

The objective function considered in this second step for the first method is the function  $J$ :

$J(p) = w_1 F_1(\theta, p) + w_2 F_2(\theta, p)$  where  $w_i$  are weights,  $F_1$  and  $F_2$  are respectively the least square residual corresponding respectively to the pressure and to the heat release.

The Levenberg Marquardt has been used in order to solve the minimization problem: find  $p$  which minimize  $J$  with initial guess  $p_0$  given by the first step.

Second method, step 2

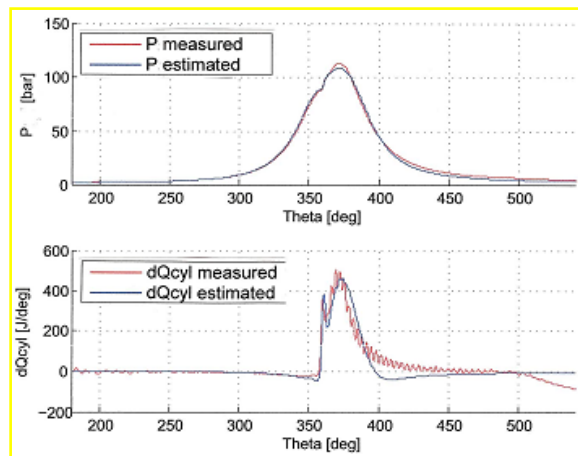
The previous objective function is a weighted sum. Because of a lack of information on the physical phenomenon, a relevant choice for the weights is difficult. In order to overcome this problem a multi-objective optimization is considered in the second method. NSGA-II algorithm [8] has been used for solving the minimization problem find  $p$  which minimize  $J_2$ , with  $s$  and  $p$  taken in a domain around  $p_0$ .

**5.2. Numerical Results**

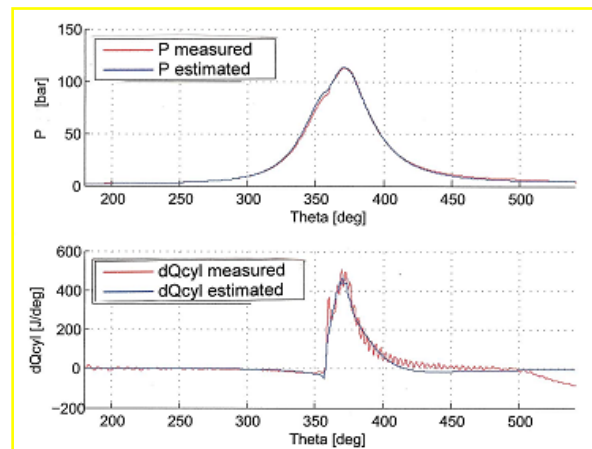
In this section, numerical results by using these two methods, denoted GM + L.M for the first one and GM + NSGA. II for the second one, are presented and compared.

The first method gives accurate results: relative error is small, however the second one give best results for the accuracy. On the other hand, the first one is the best for the computation time. The genetic algorithm used by the second one computes  $3 \times N \times M$  times the value objective function, where  $M$  is the population size and  $N$  the number of iterations. The following table presents the relative error for the reconstruction of experimental curves. In **Figure 1** and **Figure 2** experimental and reconstructed pressure and heat release are given for each methods.

method	(e.r) pour P	(e.r) pour dQcyl
GM+L.M	0.0021	0.13
GM+NSGA.II	0.0011	0.069



**Figure 1.** Pressure and heat release reconstruction: GM+LM.



**Figure 2.** Pressure and heat release reconstruction: GM + NSGA.II.

The best reconstruction particularly for the heat release is given by the second method, and more, this method is less sensitive to the first estimation than the first one.

## 6. Conclusion

In this paper an approach for analyze identifiability of the Wiebe model combustion with double phases has been proposed. The input-output-parameter relations which come from the identifiability study have been originally exploited in order to perform the parameters estimation step. These relations contain some derivatives of the noisy observations. Estimation of higher derivatives from noisy data is a numerical ill-posed problem. In order to overcome this problem, successive integrations as far as possible have been done and combined, for the remaining derivatives, with an efficient new method for estimating derivatives. This method is based on distribution approach and for the first time, it is applied to an industrial problem. This approach allows decreasing the order of derivatives and then to estimate remaining derivatives of low order efficiently with a new method which play the role of noise filter. Then, this estimation is used as initial guess for a local optimization algorithm and a multi-objective optimization algorithm in order to improve parameter estimation results. This methodology has shown its efficiency and the resulting solutions lead to a good reconstruction of the actual observations.

## References

- [1] Talon, V. (2004) Modélisation 0-1D des moteurs à allumage commandé. Ph.D Thesis, Orléan University, Orléan.
- [2] Denis-Vidal, L. and Joly-Blanchard, G. (2004) Equivalence and Identifiability Analysis of Uncontrolled Nonlinear Dynamical Systems. *Automatica*, **40**, 287-292. <http://dx.doi.org/10.1016/j.automatica.2003.09.013>
- [3] Vajda, S., Godfrey, K.R. and Rabitz, H. (1987) Similarity Transformation Approach to Identifiability Analysis of Non-linear Compartmental-Models. *Mathematical Biosciences*, **93**, 217-248. [http://dx.doi.org/10.1016/0025-5564\(89\)90024-2](http://dx.doi.org/10.1016/0025-5564(89)90024-2)
- [4] Pohjanpalo, H. (1978) System Identifiability Based on the Power Series Expansion of the Solution. *Mathematical Biosciences*, **41**, 21-33.
- [5] Ollivier, M. (1990) Le problème de l'identifiabilité structurelle globale: Approche théorique, méthodes effectives et bornes de complexité. Ph.D Thesis, Ecole Polytechnique, Paris.
- [6] Verdière, N., Denis-Vidal, L. and Joly-Blanchard, G. (2012) A New Method for Estimating Derivatives Based on a Distribution Approach. *Numerical Algorithms*, **61**, 163-186. <http://dx.doi.org/10.1007/s11075-012-9535-4>
- [7] Fliess, M., Mboup, M., Mounier, H. and Sira-Ramirez, H. (2005) Questioning Some Paradigms of Signal Processing via Concrete Examples. *Proc. Summer School: Fast Estimation Method in Automatic Control and Signal Processing*, Paris.
- [8] Deb, K. and Agrawal, R.W. (1995) Simulated Binary Crossover for Continuous Search Space. *Complex Systems*, **9**, 115-148.