Optimization for engine calibration

Hoël Langouët, Ludovic Métivier, Delphine Sinoquet, Quang-Huy Tran

EngOpt 2008 - International Conference on Engineering Optimization
Rio de Janeiro, Brazil, 01 - 05 June 2008.

Optimization for engine calibration

Hoël Langouët, Ludovic Métivier, Delphine Sinoquet* and Quang-Huy Tran

IFP, Rueil-Malmaison, France
hoel.langouet@ifp.fr ludovic.metivier@ifp.fr delphine.sinoquet@ifp.fr q-huy.tran@ifp.fr

1. Abstract
Nowadays, automotive manufacturers are submitted to strong constraints in engine calibration such as: low fuel consumption, emission-control legislation and driver requests for driving comfort and performances. These constraints lead to an increasing complexity of the engines and thus an increasing number of parameters to be tuned, making the empirical engine calibration by a scan of parameter values impossible at engine test-bench. New methodologies in automated engine calibration based on statistics and optimization have emerged in order to limit the number of experimental tests to be run.

The optimization problem of engine calibration consists in the determination of engine tuning parameters that minimize the cumulated fuel consumption and pollutant emissions on a driving cycle generally associated with legislation norms. This cycle is decomposed in a set of stationary operating points of the engine characterized by its speed and its torque (the transient behaviors of the engine are not taken into account in the stabilized calibration). Then, the optimal tuning parameters of the engine should be defined for each operating points, the functions defining these parameters on the whole engine operating domain are called the engine maps. These two-dimensional optimal engine maps are then integrated in the engine control unit in the vehicle.

We illustrate the difficulties associated with this application and propose adapted optimization methodologies: LoLiMoT models for engine map parameterization in order to handle intrinsic constraints on the map regularity, multi-objective optimization method based on CMA-ES approach. Finally, application on real dataset obtained at IFP automated test-bench for a diesel engine are presented.

2. Keywords: Engine calibration, LoLiMoT, Multi-objective optimization, Evolutionary algorithm

3. Introduction
Engine calibration consists in fulfilling the engine tuning maps that are used in engine controls of the vehicle, i.e. in defining the optimal tuning of parameters used by engine control strategies. Due to the highly increased number of these parameters (especially for diesel engines but spark ignition engines are following the same trend) and the reduction of the development schedule available for the calibration process, manual tuning of engine parameters is now replaced by mathematically assisted calibration process. Such a process is based on the design of experiments with associated modeling methods, in order to reduce the number of tests used to build engine response models depending on engine control parameters, and optimization techniques to determine the optimal settings within the model definition domain. In order to perform the tests in a more productive way, these mathematical techniques are generally associated with test automation, requiring well controlled measurement methods and reliable test equipments.

This paper describes the optimization methods developed for this application and illustrates their effectiveness on a real case of a common rail diesel Engine. The first section introduces the classical steps of the calibration process and discusses the associated difficulties. In the second section, we propose the Multi-Objective Covariance-Adaptation Evolutionary Strategy method for solving the optimization problem associated with a given engine operating point defined by the engine speed and the engine load. In the third part, an integrated approach is proposed in order to directly optimize the engine maps on the whole driving cycle (associated with legislation norms) instead of the individual optimization of each engine operating point.

4. Engine calibration
4.1. Sketch of the engine calibration process
The emission calibration workflow is classically divided into four steps:

1. a preliminary phase consisting in choosing a sample of operating points (referred to as OP in the

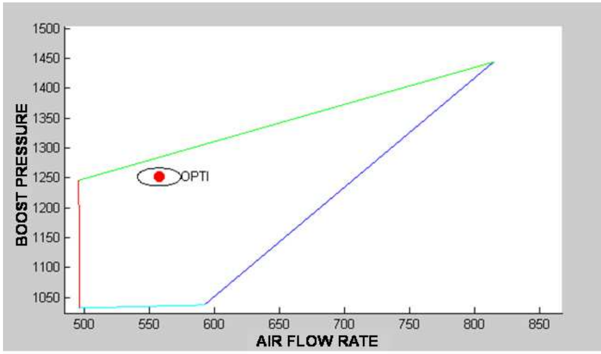

Figure 2: Projection of the domain of parameter variations in the (boost pressure, air mass) plane. The limits of the domain are generally modeled by linear constraints on the engine control parameters. The ellipse indicates the parameter dispersions (here projected on the 2D space) modeled by a multidimensional Gaussian probability law. The ellipsoïd should remain in the engine physical domain in order to avoid that any control parameter takes some value outside this domain after the implementation in the control unit.

following) to be studied and emissions targets associated with each OP, these targets being called *allocations* hereafter.

2. the optimization of engine responses on each OP according to these targets,

3. the fulfilling of the maps with a smoothing step between these optimal settings.

The preliminary step deduces from a simple model of the vehicle the trajectory of the driving cycle in the engine speed-load operating domain (see Figure 1): the transient effects on the accelerations of the cycle are neglected and the cycle is thus considered as a sum of stabilized points. A limited number of specific Operating Points are chosen to represent the cycle in the engine working range. Figure 1 gives an example of NEDC simulation and a selection of 17 OP. The cumulated levels of pollutant emissions along the cycle are computed as weighted sums of the pollutant emissions for each chosen OP.

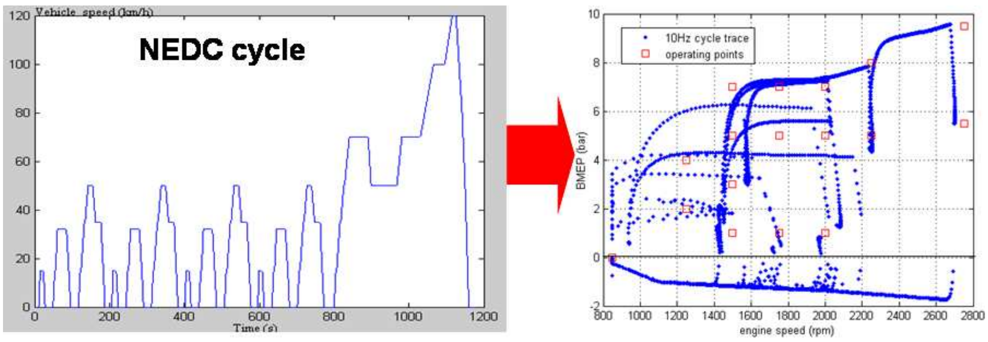

Figure 1: Simulation of the cycle from European legislation (NEDC: New European Driving Cycle) and selection of operating points

Phase 2 consists of five steps:

i) defining the domain of variations of the engine control parameters: this is an essential step of the process as it defines the validity domain of the models. The complexity of the models to be used for engine response depends on the size of this domain: for tiny domains low order polynomials (second order) are usually sufficient to accurately model engine responses. However, choosing too small domains leads to difficulty in coherently fulfilling sub-optimal engine maps.

ii) building the test matrix: various types of experimental designs can be used to build a test matrix: D-Optimal, space filling... The choice of the type of design as well as the number of tests to be done are directly correlated with the assumed complexity of the model and thus with the size of the considered domain. D-Optimal test designs are often used with hyper-cubic tiny domains.

iii) running this test matrix on the test bench: as the tests are predefined, the experiment can be performed in an automated way, which drastically improves the productivity of the global process. In this case, special attention must be paid to the validation of the experimental data.

iv) modeling the engine responses: the type of mathematical models depends on the complexity of the engine responses and on the size of the domain of parameter variations. In practice, the engineer often prefers to limit the parameter variations in order to avoid critical engine tunings that may lead to difficulty to realize the experiment at the test bench. However, as already mentioned, reducing the size of the parameter variations may lead to difficulties to build coherent optimal engine maps from the obtained optimal settings in those tiny domains. Often, low order polynomial functions are used coupled with D-optimal design of experiment. For some response as HC or CO emissions, this type of models may be too limited, then non-parametric response surfaces may be chosen as RBF or kriging. We will not detail this step of the calibration process in this paper, the reader may find more details on this topic in [3].

v) optimizing the engine control parameters to meet the allocations. The problem may be formulated as a classical mathematical problem of optimization under constraints or as a multi-objective optimization (searching for compromises between antagonist objectives) as it is detailed in the following section. For this classical approach, the optimization is performed one OP after the other, considering the allocations of each OP as constraints. We refer in the following to this optimization problem as the OP optimization problem.

When the optimal settings are found, the last step consists in integrating them in the reference engine maps (if available) or building maps from these settings. For the drivability of the target vehicle and because sharp evolution of air loop parameters are not easily feasible during transient, it is necessary to provide smooth engine maps. Thus, the settings are often moved away from their optimal values in order to build smooth engine maps, especially for air loop parameters. This smoothing process can be performed with respect to local constraints (such as maximum gradients), as well as to keep some predefined shapes. The difficulty of this step is thus to stay as close as possible to the local optima while preserving a smooth map shape, in order to keep all the benefits of the optimization work and satisfy the targets. This delicate step is usually time consuming and is likely to deteriorate the work performed during the optimization phase by providing sub-optimal engine tunings. This is the motivation for our proposal of an integrated method of map optimization described in section 6.

4.2. The OP optimization problem formulation

The OP optimization problem consists in minimizing some engine responses under constraints on others engine responses whereas the engine control parameters are kept inside the domain usually modeled by linear constraints coupling the parameters (see Figure 2). It leads to the following multi-objective constrained optimization problem:

$$\begin{cases} \min\limits_{p \in \mathbb{R}^{N_p}} (f_1(p), f_2(p), \ldots, f_m(p)) \\ \\ \text{subject to} \\ l \leq Ap \leq u \\ f_1(p) \leq s_1 \\ f_2(p) \leq s_2 \\ \ldots \\ f_d(p) \leq s_d \end{cases} \tag{1}$$

In turbo-diesel engine, a classical bi-objective optimization problem is the minimization of the particulate emissions and the NOx emissions under constraints on fuel consumption (CO_2 emissions), on CO and

HC emissions and engine noise level. Generally, this multi-objective optimization problem (1) is replaced by a constrained optimization problem where only one of the objective is effectively minimized, the other being constrained to remain below some thresholds:

$$\begin{cases} \min\limits_{p \in \mathbb{R}^{N_p}} (f_i(p)) \\ \\ \text{subject to} \\ l \leq Ap \leq u \\ f_j(p) \leq s_j \quad j \in [1, d] \end{cases} \tag{2}$$

This constrained single-objective optimization problem (2) is then solved by a classical SQP method run several times with different initial points to search for the global optimum.

5. Solving the OP multi-objective optimization problem

In this section, we present a method to solve the constrained multi-objective problem (1). It is based on the algorithm proposed by [5] namely the Multi-objective Covariance Matrix Adaptation Evolutionary Strategy method (MO-CMA-ES).
The MO-CMA-ES method belongs to the class of evolutionary strategy methods which allow to solve global optimization problem. These methods are popular for their ability to find a global minimum but their slow convergence rate is often criticized. In [4], Hansen has proposed the CMA-ES method for optimization of real-space functions, based on an original mutation operator that increases the convergence rate thanks to the adaptation of the mutation distribution at each generation.

5.1. CMA-ES method for single-objective optimization
During the CMA-ES evolutionary algorithm, a new population is created from the current population by generating realizations of a Gaussian probability law

$$x_k^{(g+1)} \sim \mathcal{N}\left(m^{(g)}, \sigma_k^{(g)2} \cdot C_k^{(g)}\right) \tag{3}$$

where $m^{(g)}$ is the barycenter of the best μ individuals with respect to the objective function to be minimized,

$$\begin{aligned} & m^{(g)} = \sum_{i=1}^{\mu} w_i x_{i:\lambda}^{(g)}, \\ & \text{with} \sum_{i=1}^{\mu} w_i = 1 \text{ and } w_\mu \leq w_{\mu-1} \leq \ldots \leq w_1 \leq 1, \end{aligned} \tag{4}$$

$x_{i:\lambda}^g$ being the i^{th} best individual of the current population $x^{(g)}$ with respect to the function. The next population is thus directed towards the best points.
An ideal choice for the covariance matrix would be the inverse of the Hessian matrix of the objective function, the idea proposed by Hansen is then to approximate this matrix as for a Quasi-Newton method but without any computation of the derivatives. The update formula for the covariance matrix is then

$$C^{(g+1)} = (1 - c_{cov})C^{(g)} + c_{cov} \sum_{i=1}^{\mu} w_i \left(\frac{x_{i:\lambda}^{(g+1)} - m^{(g)}}{\sigma^{(g)}}\right)\left(\frac{x_{i:\lambda}^{(g+1)} - m^{(g)}}{\sigma^{(g)}}\right)^T \tag{5}$$

which amounts to adding to the current covariance matrix a μ-rank term that takes into account the repartition of the best μ individuals of the new population. An other update formula is proposed by Hansen which does not take into account only the last iteration as in (5) but all the previous iterations from the beginning.
In the same way, an update formula of the standard deviation is built: the main idea is here to avoid that two consecutive iterations lead to a displacement of the barycenter of the population $m^{(g)}$ in the

same direction, replacing these two iterations by one with a larger standard deviation would have been a better choice. Likewise, a small standard deviation is preferred to a large one which leads to two iterations with opposite displacements of the barycenter of the population. The reader is referred to [4] for more details. Figure 3 illustrates on a 2D single-objective optimization example the adaptation of the covariance matrix and of the standard deviation leading to an efficient evolution of the population towards the global minimum.

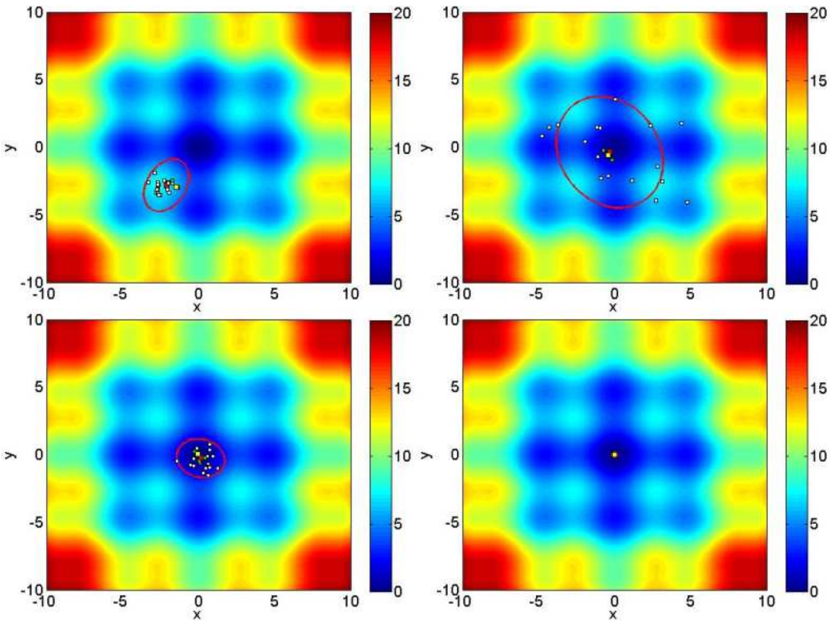

Figure 3: An application of CMA-ES method on the analytical function $f(x,y) = \frac{x^2+y^2}{10} - \frac{1}{2}\left(\cos(\frac{2\pi}{5}x) + \cos(\frac{2\pi}{5}y)\right) + 3$. From top left to bottom right, the population at iterations 1, 5, 8 and 20 are represented. The ellipse represents the 95%-isovalue of the Gaussian probability density of the mutation operator with the adaptive covariance matrix.

5.2. MO-CMA-ES method for multi-objective optimization

Adapting the CMA-ES method to multi-objective optimization requires the definition of new criteria to define what is a *good* and a *bad* individual for selection step of the evolutionary algorithm (see [5]). Two criteria are implemented in MO-CMA-ES:

1. A classical criterion for multi-objective optimization is the Pareto dominance:

   Definition 1. *An individual x is said to dominate another individual x', which is symbolized by $x \prec x'$, if and only if*

$$\begin{cases} \forall m \in \{1,...,M\} : f_m(x) \leq f_m(x') \\ \exists m \in \{1,...,M\} : f_m(x) < f_m(x'). \end{cases} \tag{6}$$

   Definition 2. *The set of Pareto optimal points of the set X is thus defined by $\{x \mid \nexists x' \in X : x' \prec x\}$.*

   The Pareto front is composed of the values taken by the objectives for all the Pareto optimal solutions. Examples of Pareto fronts are given in Figure 5. The individuals of a given population A are sorted by their non-dominance level.
   The non-dominated solutions of A are given by: $ndom(A) = \{a \in A \mid \nexists a' \in A : a' \prec a\}$. We consider that these non-dominated solutions have a non-dominance level of 1. Recursively, we define the points of non-dominance level equal to 2 as the non-dominated points of the population A without the non-dominated solutions. We thus obtain the recursive definition:

$$ndom_i(A) = ndom(dom_{i-1}(A)) \text{ with } dom_i(A) = dom_{i-1}(A) \backslash ndom_i(A) \text{ and } dom_0(A) = A. \tag{7}$$

2. The second criterion is a diversity criterion of the solutions to obtain an uniform repartition of the solutions along the Pareto front. A measure of the dominance hyper-volume at a non-dominated solution is defined by considering the surface of the rectangles defined in Figure 4 (one rectangle for each non-dominated point).

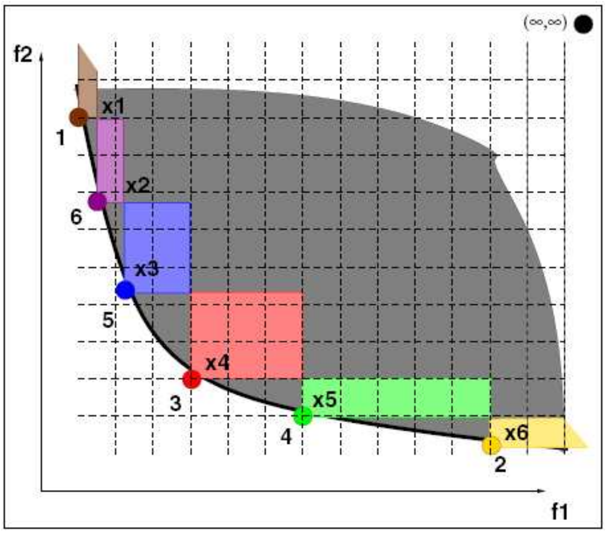

Figure 4: Contributing dominance hyper-volumes for the evaluation of the diversity criterion

Combining these two criteria allows to sort the individuals of the current population with respect to the multiple objectives and then to apply the CMA-ES update formula of the covariance matrix adaptation for the mutation. For more details on the MO-CMA-ES method, the reader is referred to [5].

5.3. Application of MO-CMA-ES method to a real case of engine calibration

In this section we present results obtained from a real dataset associated with a turbo-diesel engine. We consider the optimization problem (1) at a given engine operating point (1750 rd/mn - 1 bar). Models of the engine responses have been built from experimental data, the tests being designed thanks to a classical D-optimal criterion: the models are cubic polynomials.

The optimization is performed on six engine control parameters: the main injection timing, the pilot fuel injection quantity, the pilot injection timing, the fuel injection pressure, the mass air flow and the boost pressure.

The optimization objectives are the particulate emissions and the NOx emissions, CO2, HC and CO emissions and the engine noise being constrained to remain smaller than a given threshold. Linear constraints defining physical limits of the domain of variations of the parameters are introduced. These linear constraints may take into account the parameter dispersions, these new constraints are denoted by robustness constraints hereafter.

Thus, the multi-objective optimization algorithm has been modified to deal also with linear and non linear inequality constraints: they have been introduced via a l_1 penalization term added to the objectives (if the i^{th} constraint is not satisfied, the misfit with the bound is added to the objective functions, the penalization weight varying during the iterations).

The results are presented in Figure 5 : two optimizations are compared, one without taking into account the parameter dispersions and one with the robustness constraint which consists in taking into account the possible dispersion of the engine control parameters when applied on the vehicle. This dispersion is modeled by an uncorrelated Gaussian probability law, thus by one standard deviation by parameter. As shown in Figure 2, a vicinity of the engine setting is defined and this vicinity should remain inside the physical limits. The impact of this constraint is noticeable in Figure 5 by an important shift of the Pareto front and a modification of its shape. Engine experts were very interested in the different compromises between the objectives that have been obtained by the multi-objective optimization method.

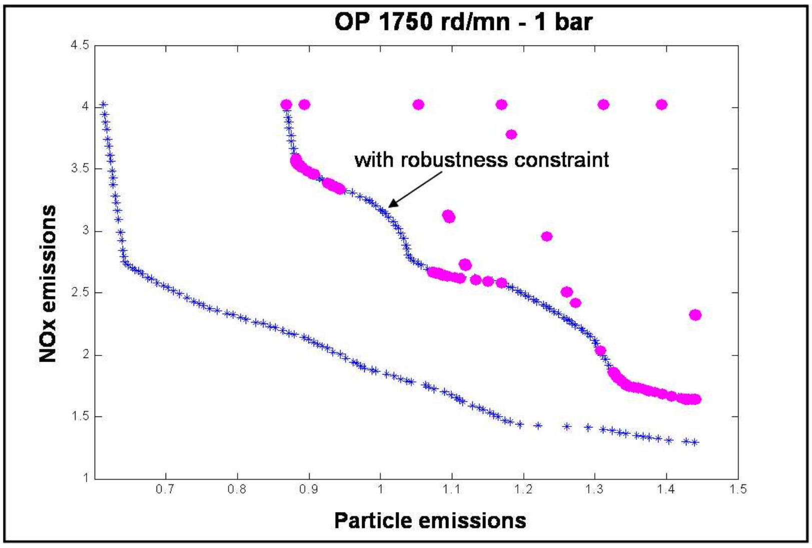

Figure 5: Pareto fronts (in dark blue) obtained with the MO-CMA-ES method for the OP 1750 rd/mn - 1bar. The front at the right side of the Figure is associated with an optimization taking into account the parameter dispersion (robustness constraints that ensure the effective engine control parameter values to remain inside the engine physical domain). The other front results from an optimization without these constraints. The pink points are solutions obtained by single-objective optimizations (weighted sum of the two objectives) associated with different initial points and different penalization weights.

The points indicated in red, green and blue are associated with successive runs of a local single-objective optimization with different initial points and with a weighted sum of the two objectives (different weight values have been tested). We notice that these optimizations lead to a very few points on the Pareto front: the other points are some local minima. Moreover, as expected, the points obtained on the Pareto front are located on the convex parts only.

The MO-CMA-ES method has performed on this example 3 times more evaluations of the objective functions than the local optimization with different initial points and different weights. Even if this performance is already very encouraging, some work on the stopping criteria may reduce the number of evaluations of MO-CMA-ES.

6. Solving the map optimization problem

In the classical calibration process described in section 4, the engine map building in the smoothing step is delicate and time consuming. In this section, we propose an alternative method that consists in directly optimizing the cumulated engine responses over the cycle via deformations of the engine maps themselves instead of optimizing individually the selected OP and building afterward the engine maps by the smoothing step. The map optimization problem is formulated as

$$\left\{\begin{array}{l} \displaystyle\min_{m^p \in \mathbb{R}^{N_p}} \int_0^T F_i\left(r(t), c(t), m^{p_1}(r(t),c(t)), m^{p_2}(r(t),c(t)), \ldots, , m^{p_{N_p}}(r(t),c(t))\right) dt \\ \\ \text{subject to} \\ l(r,c) \leq A m^p(r,c) \leq u(r,c) \\ \displaystyle\int_0^T F_j\left(r(t), c(t), m^{p_1}(r(t),c(t)), m^{p_2}(r(t),c(t)), \ldots, , m^{p_{N_p}}(r(t),c(t))\right) dt \leq S_j \quad j \in [1,d] \end{array}\right. \tag{8}$$

where $(r(t), c(t))$ indicate the trajectory of the cycle within the engine speed-load domain, F_i is the model of the engine response i depending on the engine control parameters but also on the speed and load, m^{p_i} are the 2D engine maps of the control parameters.

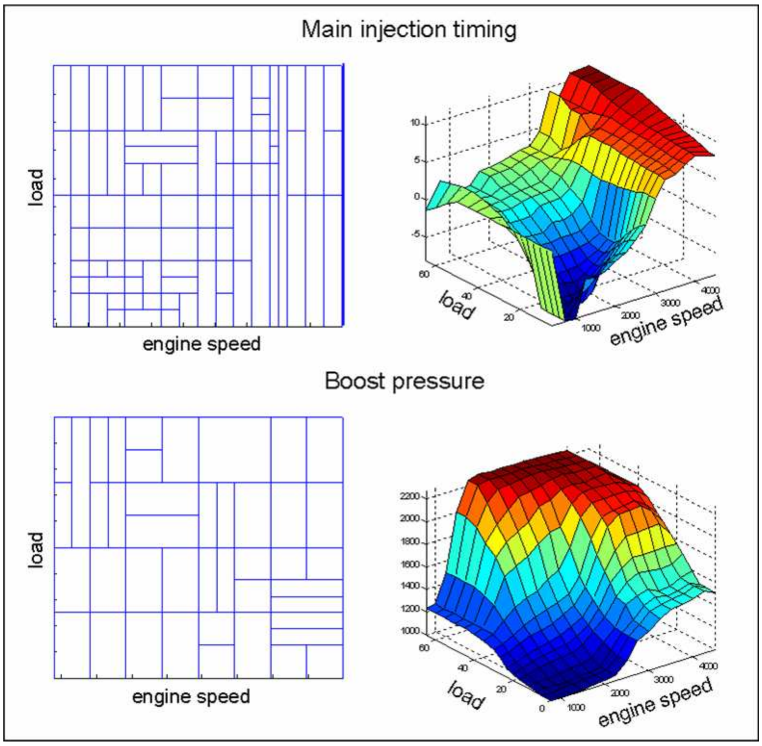

Figure 6: Examples of LoLiMoT parameterization of engine map. Left: patching of the domain. A local linear model is defined on each patch. Right: the resulting surfaces modeled by LoLiMoT for two different engine control parameters from given discrete reference maps.

The objectives to be minimized (or constrained) are the engine responses cumulated on the considered driving cycle: the cumulated responses are here (see (9)) the weighted sums of the local models defined at chosen representative OP but if global models of the engine responses are available (models including engine speed and load dependency), a finer optimization may be performed with a fine sample of the integrals in the original optimization problem (8). Additional smoothing constraints such as global smoothing constraints (to preserve the regularity of the original maps) or more local constraints (for example limits on the gradients of the maps) can also be introduced.

$$\min_{m^p \in \mathbb{R}^{N_p} \times \mathbb{R}^{N_{OP}}} \sum_{l=1}^{N_{OP}} \omega_l F_i^{(l)} \left(m^{p_1}(r_l, c_l), m^{p_2}(r_l, c_l), \ldots, , m^{p_{N_p}}(r_l, c_l) \right) \tag{9}$$

6.1. Modeling the engine maps

The formulations (8) and (9) require an adapted parameterization of the engine maps that must be flexible enough to model the very different shapes of engine map surfaces (Figure 6) and should not require too many parameters to limit the number of unknowns in the optimization process. LoLiMoT models ([1], [2]) seem to be a good compromise between flexibility, accuracy and complexity: some very simple local models (linear or bilinear) are combined by a weighted sum

$$m(r, c) = \sum_{i=1}^{M} \hat{m}_i(r, c) \Phi_i(r, c) \text{ with } \hat{m}_i(r, c) = \omega_{0i} + \omega_{ri} r + \omega_{ci} c \tag{10}$$

where the weights $\Phi_i(r, c)$, normalized Gaussian functions, control the degree of smoothness of the global surface:

$$\Phi_i(r, c) = \frac{\mu_i(r, c)}{\sum_{j=1}^{M} \mu_j(r, c)} \text{ with } \mu_i(r, c) = \exp\left(-\frac{1}{2\alpha} \frac{(r - {r_i}^0)^2}{{\sigma_i^r}^2} + \frac{(c - {c_i}^0)^2}{{\sigma_i^c}^2} \right) \tag{11}$$

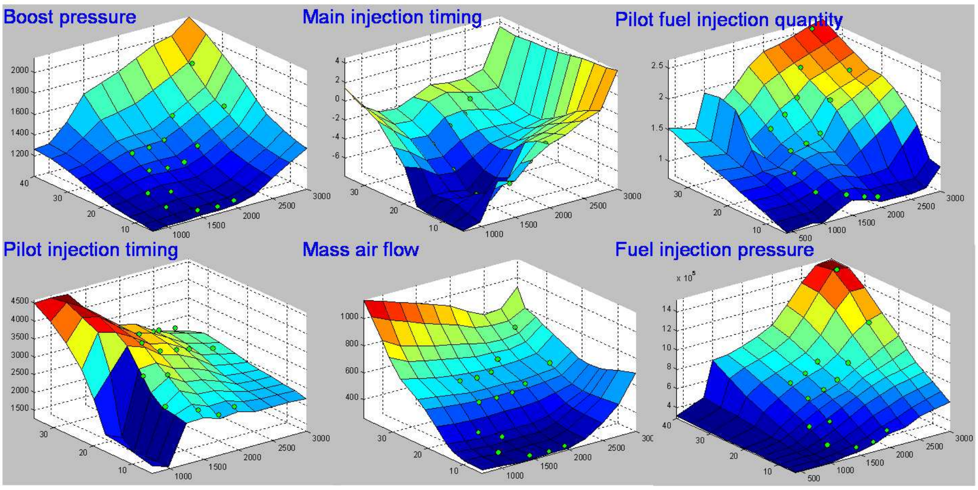

Figure 7: Optimal engine maps obtained by the map optimization on the NEDC driving cycle. 378 LoLiMoT parameters are necessary to model the 6 maps with the expected regularity. The green points indicate the engine operating points for which models of engine responses are available.

This representation allows an adaptive refinement of the surface: the patching associated with the definition domains of the local models may be refined during the optimization process. A finer patching allows a finer optimization (the number of degrees of freedom being increased) but may lead to a cumbersome optimization of a large number of parameters. The parameterization, namely the patch definition for LoLiMoT description of the maps should reflect the degree of smoothness the user expects for the maps: for some parameters like boost pressure, the map should remain smooth, for others like main injection timing, the smoothing constraint is not as strong.

Thus, from some reference engine maps or some a priori information, a LoLiMoT parameterization of each map of engine tunings is defined (Figure 6). The unknowns of the optimization are the LoLiMoT parameters (coefficients of local linear models $\omega_{.i}$).

6.2. Application of map optimization on a real case

The map optimization has been applied on the diesel engine application presented in section 5.3. The optimization problem is defined as a constrained single-objective optimization problem where the cumulated particulate emissions are minimized and the cumulated CO2, CO, HC and NOx emissions are constrained to remain under thresholds (given in table 1). Linear constraints are added to take into account the limits of the physical domain. The obtained results are listed in table 1) and the optimized engine maps are displayed in Figure 7. These engine maps have been run at the test bench and particulate emissions have been measured, those measures are displayed in Figure 8.

Table 1: Results of map optimization on a diesel engine application.

	Initial maps	Optimized maps	Upper bounds
Particulate (g/km)	0.128	0.090	0.090
NOx (g/km)	0.219	0.240	0.241
HC (g/km)	0.137	0.150	0.150
CO (g/km)	0.815	0.847	0.896
CO2 (g/km)	168.89	164.40	168.89

7. Conclusions

In this paper, the optimization problems encountered in engine calibration are addressed. The constrained multi-objective optimization has been successfully solved by the MO-CMA-ES method modified

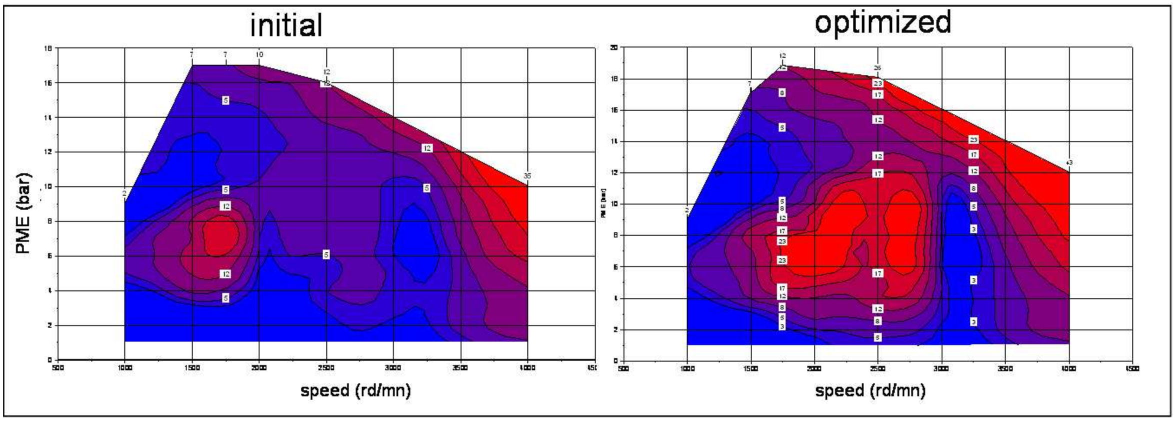

Figure 8: Comparison of measures of particulate emission on the test bench for the initial engine maps and the optimal ones obtained by the proposed map optimization method.

to handle inequality constraints. The performance of this method on the considered real case application is very encouraging. The resulting Pareto front gives to the engine experts a worth information on the possible compromises between antagonist engine responses as particulate emissions and NOx emissions of a diesel engine. In a second part, a new formulation of the engine map optimization is proposed based on a cycle optimization instead of an optimization at given engine operating points. This formulation allows to handle the smoothing constraints on the maps directly in the optimization process, thus to avoid the generally cumbersome step of map smoothing in classical engine calibration process. Results of a constrained single-objective map optimization have shown the efficiency of the method. A next study will be to apply a multi-objective approach as for the OP optimization, the difficulty will be the large size of the parameter space (500 parameters to describe all the engine maps). An other improvement in this approach will be the use of global models of engine responses (defined on the whole engine operating domain instead of the local models at a limited number of OP): this optimization is expected to yield finer results.

8. References

[1] O. Nelles, *Non linear system identification*, Springer-Verlag, 2001.

[2] M. Hafner and R. Isermann, *The use of stationary and dynamic emission models for an improved engine performance in legal test cycles*, International workshop on Modeling, Emissions and Control in automotive engines, Salerno, Italy, 2001.

[3] M. Castagné, Y. Bentolila, F. Chaudoye, A. Hallé, F. Nicolas and D. Sinoquet, *Engine calibration: towards an integrated approach*, IAV conference - DOE in Engine Development - Design of experiments and other modern development methods, Berlin, Germany, 2007.

[4] N. Hansen and A. Ostermeier, *Adapting arbitrary normal mutation distributions in evolution strategies: The covariance matrix adaptation*, In Proceedings of the 1996 IEEE International Conference on Evolutionary Computation, p. 312-317, 1996.

[5] N. Hansen, C. Igel and S. Roth , *Covariance Matrix Adaptation for Multi-objective Optimization*, Evolutionary Computation, 15(1), p.1-28, 2007.

9. Acknowledgments

The authors would like to thank Yohan Entails, Michel Castagné, Fabien Chaudoye, Roland Dauphin, Adrien Hallé and Frédéric Nicolas for their explanations on the engine calibration problem, their advices and last but not least for providing the real dataset for this study.