

# POLITECNICO DI TORINO

*Master of Science in Automotive Engineering*

*Master's degree thesis*

Model-based calibration for diesel  
engines: comparison between different methods  
and application to experimental data



Academic supervisor

*Prof. Stefano D'Ambrosio*

*Prof. Roberto Finesso*

*Dott. Alessandro Mancarella*

Candidate

*Giuseppe Longo*

Academic year 2020-2021

# Abstract

The subject of this master thesis is to investigate on model-based calibration, in particular on the modelling stage.

The first phase has been dedicated to the DoE theory, the first step of an engine calibration work: very essential to reduce the tests number and increase the model quality at the same time. Then an alternative method was found for outlier detection.

These theoretical considerations have been applied on experimental diesel engine data: gaussian and polynomial models are compared on different tests, with different engines and different strategies in order to evaluate their precision in fitting and predictivity.

Model-based calibration (MBC) toolbox is the employed software for the activity.

## Ringraziamenti

Questo bellissimo capitolo della mia vita è giunto al termine dopo cinque anni e mezzo e mi lascio alle spalle tanti ricordi, tante emozioni e tanta esperienza accumulata.

Con non poca nostalgia ringrazio prima di tutti il Politecnico per avermi messo alla prova e per avermi dato la possibilità di crescere ed imparare sotto tanti punti di vista.

Ringrazio il prof. Stefano D'Ambrosio, il prof. Roberto Finesso e l'ing. Alessandro Mancarella per avermi dato la possibilità di intraprendere questo percorso di tesi e avermi aiutato sempre con grande disponibilità e gentilezza in questo periodo difficile dal punto di vista dello svolgimento delle attività.

Un infinito grazie alla mia famiglia che mi ha sostenuto sempre al cento per cento per ogni mia scelta e per essermi stata vicina in ogni momento in questa esperienza. Vi sarò per sempre grato per quello che avete fatto e fate per me.

Ringrazio tutti i miei amici, vecchi e nuovi, che hanno reso speciale e meno noioso questo percorso di studi.

Un grazie particolare a Chiara che da circa 2 anni mi è stata sempre vicina e mi ha aiutato in mille momenti e in mille situazioni, regalandomi sempre tanta felicità.

# Contents

Abstract .....	2
<b>Ringraziamenti .....</b>	<b>3</b>
1. Introduction.....	9
2. Design of experiment .....	11
2.1. Classical design .....	12
2.2. Space filling design .....	13
2.3. Optimal design.....	15
2.4. Active DoE.....	16
3.1. Modelling: polynomial and gaussian processes .....	18
3.2. Outliers detection.....	21
3.3. Statistical tool .....	29
3.3.1. Statistical indicators.....	29
3.3.2. Stepwise regression.....	30
3.3.3. Box-Cox transformation .....	33
3.3.4. Prediction error variance view .....	34
4. Modelling.....	37
4.1. 1800 rpm x 27 Nm-PCCI combustion (3L engine).....	37
4.2. 5 operating points-conventional diesel combustion (3L engine) .....	42
4.3. Multiple injections DoE (2L engine).....	43
4.3.1. 1500rpmx2bar pM.....	45
4.3.2. 2000rpmx5bar ppM.....	46
4.4. Active DoE-AVL CAMEO simulator.....	47
5. Conclusions.....	49
6. Acknowledgements .....	50
Appendix.....	51
Bibliography.....	54

# List of figure

Figure 1.1 LOAD/SPEED chart: different operating points changing the test typology .....	9
Figure 2.1 Comparison between DOE and OFAT .....	11
Figure 2.2 Full factorial plan: how increase the experiments number with the input variables .....	12
Figure 2.3 Composite design examples .....	12
Figure 2.4 LHS design on MBC .....	13
Figure 2.5 Lattice design examples: on the left a poor choice on prime number. On the right a good one ..	14
Figure 2.6 Confidential intervals of the model coefficients (b) .....	15
Figure 2.7 Comparison between traditional DoE and Active DoE on Smoke(FSN) measurements .....	17
Figure 2.8 comparison between conventional DoE and Active DoE on Smoke/NOx plot .....	17
Figure 3.14 HC(ppm) measurement over 150 tests (2700rpmx12bar) .....	32
Figure 4.1 error percentage of bsfc validation data .....	39
Figure 4.2 error percentage of bsNOx validation data .....	40
Figure 4.3 Mean and maximum error on validation data: 1800rpmx27Nm PCCI combustion .....	40
Figure 4.4 Mean and maximum error difference of MFB50, Noise of cylinder 1 and EGR percentage.....	41
Figure 4.5 Validation data: measured FSN versus residuals (1500rpmx2bar pM) .....	46
Figure 4.6 bsNOx surface. On the left the predicted surface by 2nd order model. On the right the one predicted by gaussian model.....	46

# List of tables

Table 4.1 3L engine specifications.....	37
Table 4.2 Statistical indicators about 1800x27 PCCI models.....	38
Table 4.3 $R^2$ adjusted and PRESS $R^2$ related to polynomial model(1800x27 PCCI) .....	38
Table 4.4 optimal calibration(1800rpmx27Nm) .....	42
Table 4.5 statistic indicators about 3250rpmx47Nm point .....	43
Table 4.6 list of injection strategies.....	44
Table 4.7 statistic indicators about 1500rpmx2bar pM point .....	45
Table 4.8 ratio between validation RMSE and mean value about 1500rpmx2bar pM point.....	45
Table 4.9 statistic indicators about 2000rpmx5bar ppM point .....	46
Table 4.10 ratio between validation RMSE and mean value about 2000rpmx5bar ppM point .....	47
Table 4.11 Statistic indicators between conventional and active D.o.E. (2200rpm x 375Nm) .....	47
Table 4.12 minimum bsfc (conventional procedure 2200rpm x 375 Nm) .....	47
Table 4.13 minimum bsfc (active procedure 2200rpm x 375 Nm).....	48
Table 4.14 minimum bsfc with constraint on NOx (active procedure 2200rpm x 375 Nm) .....	48
Table 4.15 Statistic indicators between conventional and active D.o.E. (2600rpm x 425Nm) .....	48
Table A.1 1500rpmx2bar pM models: statistical indicators.....	51
Table A.2 2000rpmx5bar pM models: statistical indicators.....	51
Table A.3 2750rpmx12bar pM models: statistical indicators.....	51
Table A.4 1500rpmx2bar ppM models: statistical indicators.....	51
Table A.5 1500rpmx5bar pMa models: statistical indicators.....	52
Table A.6 1500rpmx5bar ppM models: statistical indicators.....	52
Table A.7 1500rpmx5bar ppMa models: statistical indicators.....	52
Table A.8 2000rpmx2bar ppM models: statistical indicators.....	52
Table A.9 2000rpmx5bar pMa models: statistical indicators.....	53
Table A.10 2000rpmx5bar ppM models: statistical indicators.....	53
Table A.11 2500rpmx8bar pMa models: statistical indicators.....	53
Table A.12 2750rpmx12bar pMa models: statistical indicators.....	53

# Nomenclature

## Acronyms

**bsfc**(or **b<sub>e</sub>**) brake specific fuel consumption  
**CI** confidential interval  
**CLF** cumulative distribution function  
**CN** combustion noise  
**CoV** coefficient of variance  
**DoE** design of experiment  
**dof** degrees of freedom  
**DT** dwell time  
**ED** error difference  
**EGR** exhaust gas recirculated  
**FSN** filter smoke number  
**GP** gaussian process  
**HC** unburned hydrocarbons  
**IQR** interquartile distance  
**LHS** latin hypercube sampling  
**MSE** mean square error  
**MFB50** crank angle where the 50% of injected fuel is burned  
**NO<sub>x</sub>** nitrogen oxides  
**OFAT** one factor a time  
**PCCI** premixed combustion compression ignition  
**PEV** prediction error variance  
**PRESS** predicted residual sum of squares  
**RDE** real driving emissions  
**RMS** root mean square  
**RMSE** root mean square error  
**SCR** selective catalytic reduction  
**SE** square exponential  
**SSE** sum of square error

## Symbols

***D*** Cook's distance  
***e*** residual  
***H*** hat matrix  
***h*** leverage  
***L*** number of levels  
***p*** number of model parameters  
***r*** studentized residual  
***s<sub>415</sub>*** filter smoke number measured by smoke meter

**$X$**  regressor matrix

**$Xr\_CO2$**  exhaust gas recirculated percentage measured by CO<sub>2</sub> gas analyzer

**$y$**  measured output

**$\bar{y}$**  mean output value

**$\hat{y}$**  predicted output

**$\lambda$**  exponential parameter of Box-Cox transformation

**$\nu$**  degrees of freedom



# 1. Introduction

Upcoming legislation of Real Driving Emissions (RDE) brings new challenges for the automotive world. In order to improve efficiency in calibration, the measurement methodologies have evolved, at the same time empirical model prediction capability was improved through new model types - the reason behind this being to improve the accuracy with lower measurement effort.

These improvements allowed fulfilment of the emission legal limits with minimum fuel consumption but also handling the complexity of an increasing number of actuators in the powertrain system. Up to some years ago, emissions were optimized within the certification cycle area, while outside of this area, the focus was fuel consumption minimization.

With upcoming regulations, the calibration needs to be robust against a wider powertrain operating range, under different environmental and driving conditions. Due to the high number of variants, it is not feasible to test and calibrate the emissions in all conditions. Therefore the model-based approach has to be extended to non-standard conditions.

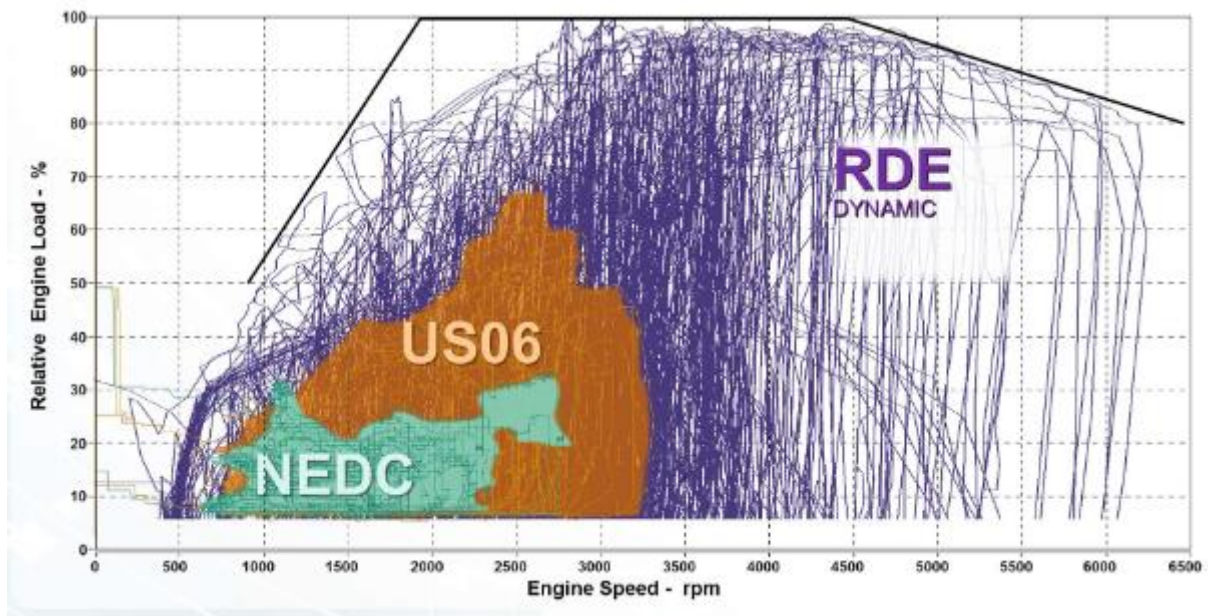


Figure 1.1 LOAD/SPEED chart: different operating points changing the test typology

A model-based approach, which can predict emissions under real driving conditions, is required. But “What is a model?”

The two types of models are:

- The physically based, mathematical models are typically applied in the early phases of a development process in simulation tools. E.g. for designing the gas path of an engine including the after treatment system. Such models are using well understood relationships and can reach a high degree of detail with high reliability in the prediction quality. The input parameters in this example are gas condition parameters and design parameters of the engine to be developed. The challenge lies in the real time capability.
- In case of very complex relationships – such as combustion processes, a system is just observed under defined, different input conditions. The Design of Experiment techniques (DoE) allow correct placement of input settings, in order to identify empirical, time free, mathematical

models, from the observed results. This is achieved with low measurement effort. The calculation speed of such models is remarkable higher and far beyond real time - but they are typically not capable to predict outside the investigated range. [1]

In the engine calibration field, with high number of actuators, the idea is to measure some output (consumption, emissions, noise) and to construct a mathematical model by regression. This is the fastest method. However it is impossible to measure and test all the possible combinations. To overcome this problem the Design of Experiment theory is employed.

The used software for modelling is MBC toolbox, a Matlab tool, in this thesis.

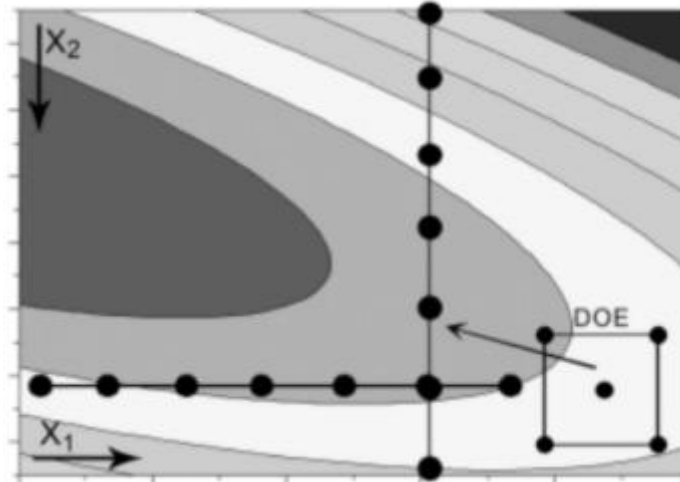
Model-Based Calibration Toolbox provides apps and design tools for modeling and calibrating complex nonlinear systems. It can be used in a wide range of applications, including powertrain systems such as engines, electric machines, pumps, and fans, as well as nonautomotive systems such as jet engines, marine hydrofoils, and drilling equipment.

It consists of two user interfaces and both have been used in the following thesis work:

- **Model Browser:** provides the tools for applying the DoE technique and for statistical modeling. The tool provides the techniques for DoE construction: Optimal, Space filling, Classical designs.  
Importing data, it also allows the comparison between various statistical models and experimental projects and there is a large library of pre-built models that can be used or alternatively it is possible to create new ones. It is possible to study the model fitting of different model process and finally, after creating the model, it can be exported to Matlab, Simulink or CAGE for further analysis.
- **CAGE Browser:** provides tools for analytical calibration. It is a simple graphical interface created with the aim of calibrating the lookup tables of the electronic control unit. Through analytical methods it allows to easily calibrate lookup tables. There are also optimization tools and data testing, to find optimal point (with or without constraints) and trade-off graphs.

## 2. Design of experiment

The term Design of experiment (DoE) refers to a series of experiments to be performed and to be analyzed. The goal is to design in the best way a series of experiments drawing conclusions from his behavior in the shortest possible time. The DoE technique, unlike the intuitive approach that bases the search for the best by changing sequentially every single variable, changes all factors simultaneously allowing to find the optimal solution quickly. The basic concept is a symmetrical distribution of the experiments around a central point. A range is established for each factor, the center point is calculated and the distribution is created symmetrical. The Figure 2.1 compares the traditional OFAT (one-factorial-at-time) and the DoE approach.



*Figure 2.1 Comparison between DOE and OFAT*

In the traditional method, the experimenter does not know which value of  $X_1$  or  $X_2$  to change, he goes to attempts, changing their value one at a time. DoE instead creates a symmetric distribution of experiments around the center point, simultaneously changing all factors, outlining a direction that allows for better results.

The big disadvantages of OFAT is the long time to perform all the experiments especially if the inputs are more than two and the risk to conclude the process finding a local optimum point, not a global one [2].

The steps of DoE methodology are:

1. Recognition of and statement of the problem
2. Selection of the response variable
3. Choice of the factors, levels and ranges (inputs domains)
4. Choice the experimental design
5. Performing the experiment

For what about the point 4, there are three different approaches: Classical, Space-filling, Optimal.

## 2.1. Classical design

This approach is the simplest one. After defining the input variables, their ranges and their levels the experimenter develops the variation list considering all the possible combinations. It is possible to represent this graphically by square (2 inputs), cube (3 inputs) or hypercube (more than 3 inputs). The number of all possible experiments are:

$$N = \prod_{i=1}^D L_i$$

$L_i$  is the number of levels of the input  $i$ ,  $D$  is the number of the input variables.

This specific approach is called Full factorial plan.

It is clear that this solution is very expensive from time point of view: for example in engine calibration situation, like the following figure, the number of measurements is huge.

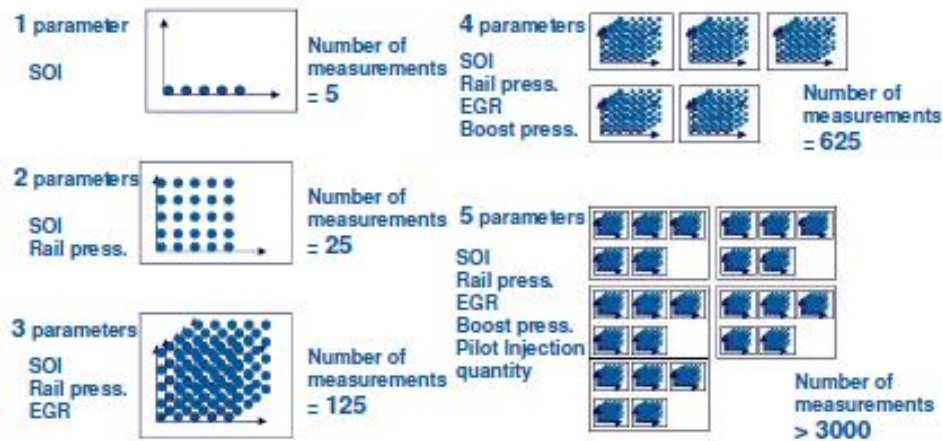


Figure 2.2 Full factorial plan: how increase the experiments number with the input variables

To overcome this issues it is possible to adopt a fractional factorial design, choosing only a fraction of the full factorial design.

Another design type for a classical approach is the composite design. It combines the investigations done by a fractional design, the corners and replicated central-points, with the use of axial experiments. These last ones are called star points. In general their locations on the cube representation in function of the distance from central point and the other points.

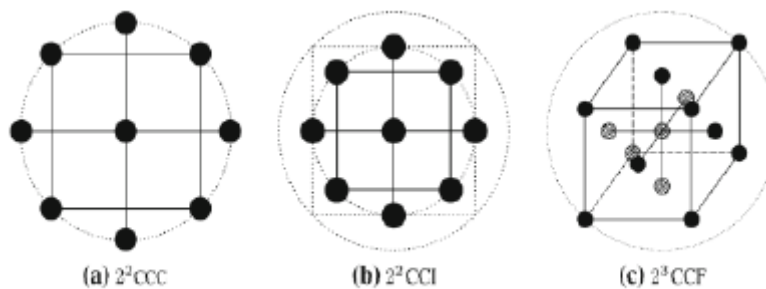


Figure 2.3 Composite design examples

These approaches in engine calibration are practically never used because are the most expensive approach from time (and so cost) point of view, and because there are procedures (like optimal design) that give better model results with less tests.

## 2.2. Space filling design

Space-filling design should be used when there is little or no information about the underlying effects of factors on responses. For example, they are most useful when you are faced with a new type of engine, with little knowledge of the operating envelope. These designs do not assume a particular model form. The aim is to spread the points as evenly as possible around the operating space. These designs literally fill out the n-dimensional space with points that are in some way regularly spaced.

There are different types of space filling design.

*Latin Hypercube Sampling (LHS)* are sets of design points that, for an N point design, project onto N different levels in each factor. Here the points are generated randomly. You choose a particular Latin Hypercube by trying several such sets of randomly generated points and choosing the one that best satisfies user-specified criteria.

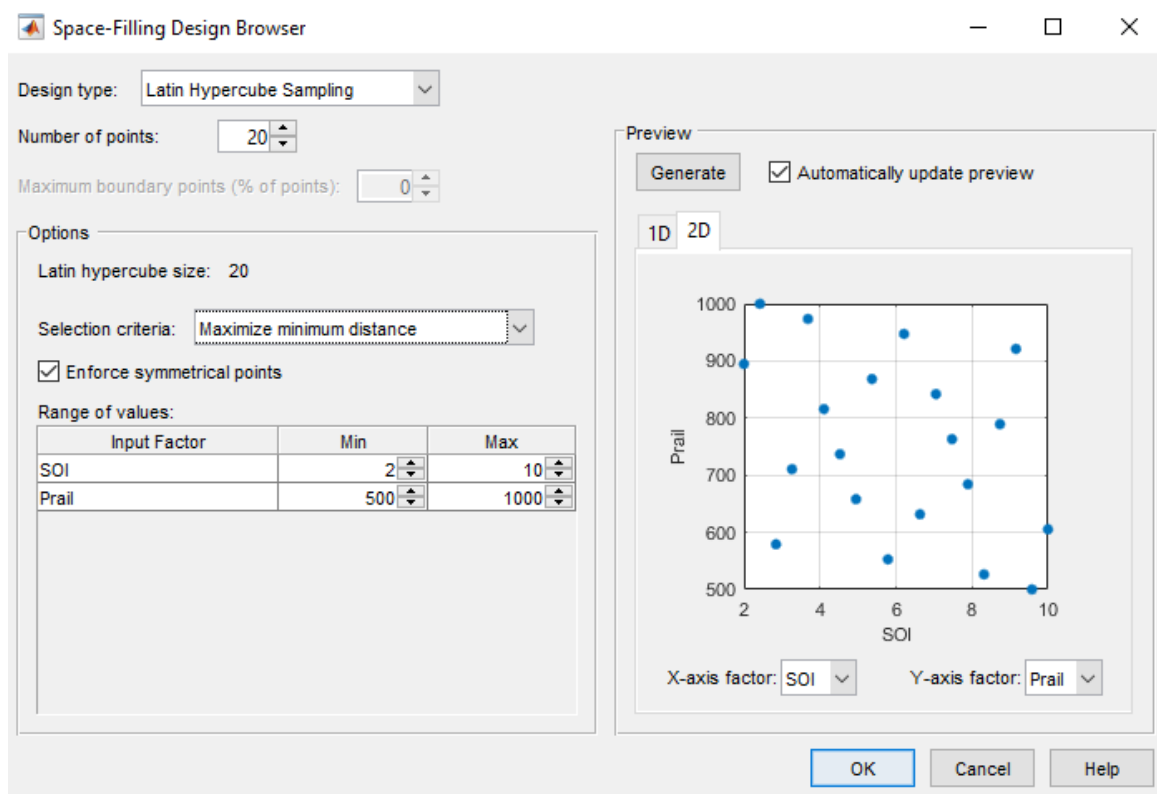


Figure 2.4 LHS design on MBC

On MBC there are different selection criteria for the points distribution:

The Selection criteria drop-down menu has these options:

- Maximize minimum distance (between points).
- Minimize maximum distance (between points)
- Minimize discrepancy — Minimizes the deviation from the average point density

- Minimize RMS variation from CDF — This option minimizes the Root Mean Square (RMS) variation of the Cumulative Distribution Function (CDF) from the ideal CDF.
- Minimize maximum variation from CDF — Minimizes the maximum variation of the CDF from the ideal CDF.

*Lattice designs* project onto N different levels per factor for N points. The points are not randomly generated but are produced by an algorithm that uses a prime number per factor. If good prime numbers are chosen, the lattice spreads points evenly throughout the design volume. A poor choice of prime numbers results in highly visible lines or planes in the design projections. If all the design points are clustered into one or two planes, it is likely that cannot estimate all the effects in a more complex model.

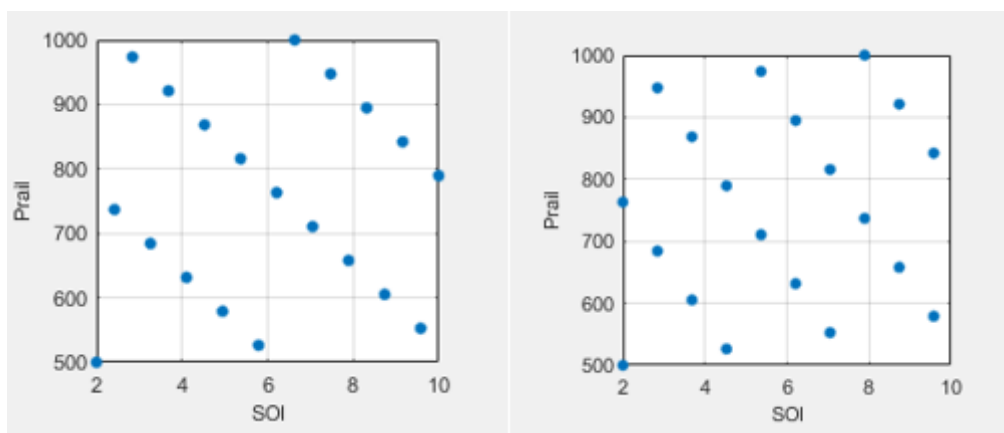


Figure 2.5 Lattice design examples: on the left a poor choice on prime number. On the right a good one

When design points are projected onto any axes, there are a large number of factor levels. For a small number of trials (relative to the number of factors) LHS designs are preferred to Lattice designs. This is because of the way Lattice designs are generated. Lattice designs use prime numbers to generate each successive sampling for each factor in a different place. No two factors can have the same generator, because in such cases the lattice points all fall on the main diagonal of that particular pairwise projection, creating the visible lines or planes described above. When the number of points is small relative to the number of factors, the choice of generators is restricted and this can lead to Lattice designs with poor projection properties in some pairwise dimensions, in which the points lie on diagonals or double or triple diagonals. This means that Latin Hypercube designs are a better choice for these cases.

*Stratified Latin Hypercubes* separate the normal hypercube into N different levels on user-specified factors. This can be useful for situations where the preferred number of levels for certain factors might be known; more detail might be required to model the behavior of some factors than others. They can also be useful when certain factors can only be run at given levels. All these methods require a symmetric design space [3].

*S-optimal design* is also a space filling technique. The space is filled by maximizing the minimum distance between the points which leads a much more equally distributed coverage of the design space. Points on the domain border are integrated specifically, so that a better border coverage is achieved. For a very low number of design points in respect to the variation space dimension, only borders will be covered. In this case the LHD would be a better choice. S-optimal design fully supports asymmetric design space: this is the best advantage [4].

### 2.3. Optimal design

The optimal design approach is the most refined method for DoE construction. The starting point is the classical equation between inputs and output for linear model:

$$Y = X * \beta + \varepsilon$$

The  $\beta$  is the parameters vector. If it will be estimated by least square method, it is:

$$\beta = (X' * X)^{-1} * X' * y$$

The experiment tests are expressed in the matrix  $X$ . So the optimal design goal is to find the best matrix  $X$  according the chosen criteria. From the previous considerations, it is clear that the optimal design is suitable for linear model, because the dimension of matrix  $X$  is function of polynomial order and structure.

The *D-optimal design* employs a criterion on the selection of design points that results in the minimization of the volume of the joint confidence region of the regression coefficients. The DoE is chosen to maximize  $\det(X' * X)$ . The matrix  $(X' * X)$  is called information matrix. When fitting the model to experimental data, the experimental error and the error of the model are transmitted to the coefficients. Geometrically, the coefficients and their errors are represented as (hyper)ellipsoids whose axes describes these errors (figure 2.6). So, the smaller the axes, the more precise are the coefficients and consequently, the more accurate are the predictions. The volume of this ellipsoid is inversely proportional to the square root of the determinant of the information matrix. So this criteria tries to maximize this determinant, which is the same as minimizing the volume of the ellipsoid.

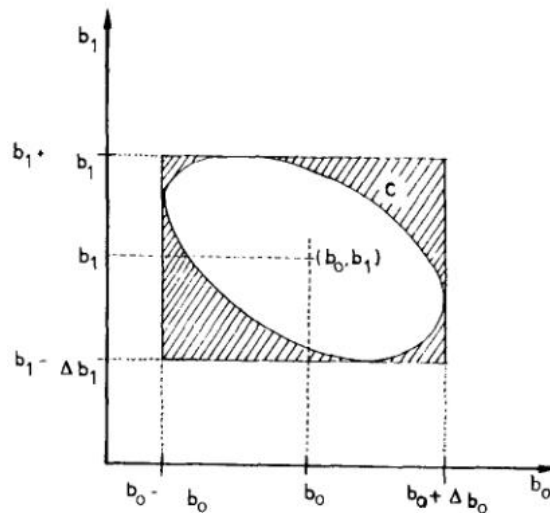


Figure 2.6 Confidential intervals of the model coefficients ( $b$ )



The *A-optimal design* criteria minimizes the trace(sum of the diagonal elements) of the matrix  $(X' * X)^{-1}$ , called dispersion matrix. The practical meaning of this operation is minimizing the average variance of the estimated coefficient and reduces the asphericity of the confidence ellipsoid.

The variance function is a measure of the uncertainty in the predicted response. It is described as follows:

$$var(\hat{y}_i) = x_i' * (X' * X)^{-1} * x_i * \sigma^2 = PEV(x_i) * \sigma^2$$

where PEV is the prediction error variance(read the chapter xx for more details). The ideal situation would be when the variance function is as small as possible, leading to minimum deviations between predicted and true values. When PEV is equal to 1, the variance in the prediction will be the same as the variance of the experimental method [5]. To increase the precision PEV must be as low as possible in all the inputs domain. Therefore the prediction error variance is function of the regressor matrix X, that contains the DoE.

The *G-optimal design* minimizes the maximum value of prediction error variance in the design region.

The *V-optimal design* minimizes the mean value of prediction error variance in the design region.

The optimal designs are the smartest approaches to design a variation list: it is adaptable to all types of inputs domain and potentially the only limit to the tests number is the coefficients number (for example, at least 10 tests are necessary to compute a polynomial model with 10 coefficients). With this approach the calibration engineer can have better modelling results from its DoE than one designed by a classical approach also with a smaller tests number.

## 2.4. Active DoE

Active DoE is a method which iteratively adapts the test design during the testrun.

On the basis of already measured data, models are calculated online. These models are used to adapt the test design during the testrun in such a way that measurements are executed within the required ranges (Active DoE is also called COR DoE = Customized Output Range DoE).

This DoE type is available on AVL Cameo (not available on MBCmodel).

The testrun is continued until the model quality matches a termination criterion. This has the advantage that it is not necessary to specify the number of required measurement points in advance and that only as much measurements as required are executed.

With Active DoE the test design can be adapted advantageously:

- Selected measurement quantities can be distributed equally within a range - increasing, therefore, the model quality in those ranges which are of interest to users.
- Selected measurement quantities can be minimized or maximized - providing, therefore, the possibility to run an online optimization.
- Selected measurement quantities must not fall below or exceed a predefined value - avoiding, therefore, limit violations through the selection of variation parameters.

The below figure shows the comparison between conventional DoE (to the left) and Active DoE (to



the right):

- With the conventional approach, the test design is created without information on the system behavior. Consequently, it might be possible that too little points are measured in ranges showing nonlinearity to sufficiently model the behavior and, therefore, a further test must be started. As a result, with the conventional approach measurements may be executed in ranges which later on do not end in additional benefit during the evaluation process.
- With Active DoE, the test run is adapted automatically, thereby well covering relevant ranges and nonlinearities without requiring re-measurements [6].

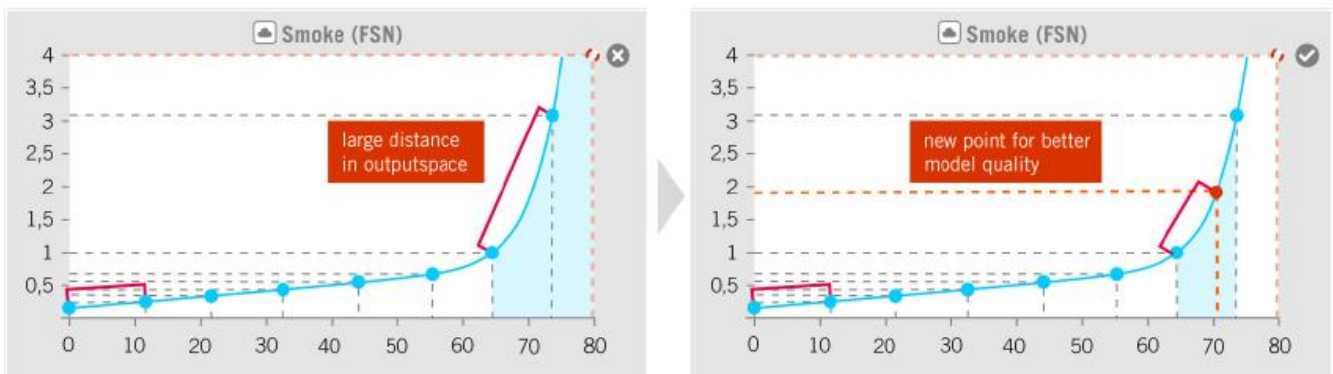


Figure 2.7 Comparison between traditional DoE and Active DoE on Smoke(FSN) measurements

The Active DoE can be applied for local and global models.

For example in global situation(varying load and speed) analyzing the Smoke/NO<sub>x</sub> plot can be useful for a diesel engine. With an active DoE, focussing on pareto front is possible avoiding to waste time on uninteresting points.

This technique ensures respect conventional ones:

- Maximum accuracy if the optimization results
- Focus on area of interest
- 30% less data needed for better results
- Added protection from wear and drift [7].

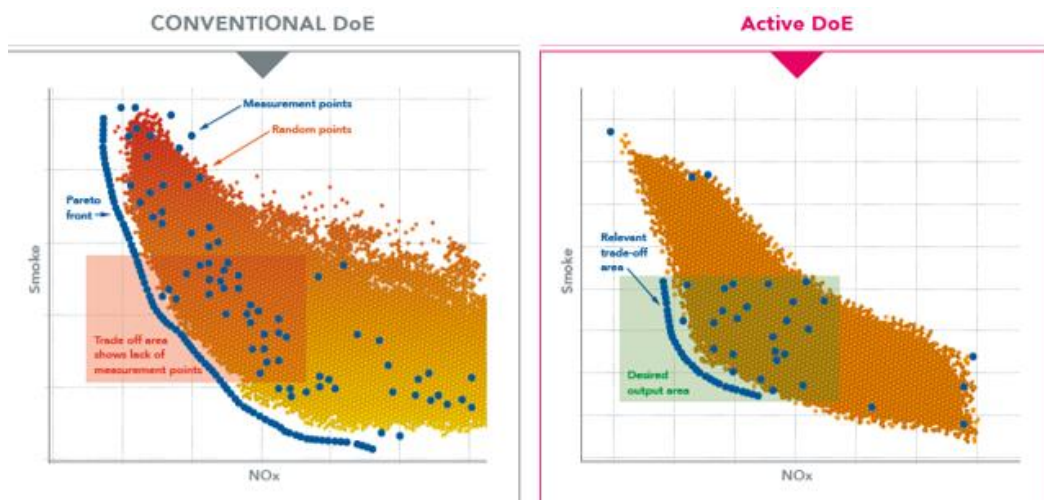


Figure 2.8 comparison between conventional DoE and Active DoE on Smoke/NO<sub>x</sub> plot

### 3. Model-based calibration

In engine calibration the aim of modeling is to assign to an output function (such as NOx emissions, fuel consumption) a function or a system of functions that depends on several inputs (such as injection timing, injection pressure).

Since the internal combustion is a complex process, which is influenced by thermodynamics, fluid dynamics and chemistry, a precise modeling and simulation of this process is not practicable today. Hence, nowadays only very simplified physical models of the combustion are used for simulation in engine calibration. Further in engine calibration, models are used in domains where the combustion does not need to be considered.

In base engine calibration many effects of the combustion cannot be neglected. A typical example is the modeling of different emissions, e.g. soot, with respect to many different adjustment parameters, e.g. input pressure. Obviously, in this example the fluid flow and the chemical reactions in the cylinder have a major influence on the formation of soot and cannot be neglected.

However, as said above, a detailed physical modeling and numerical simulation is not practicable today. Therefore, in base engine calibration measurements are taken from the test bench and with this data a black box modeling is done.

A set of measurements  $\mathbf{D} = \{(\mathbf{x}_n, \mathbf{y}_n) | \mathbf{n} \in \{1 \dots N\}\}$  where N is the number of measurements, contains the values of inputs  $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_N)$ , and the measured values (outputs)  $\mathbf{Y} = (\mathbf{y}_1, \dots, \mathbf{y}_N)$ .

The calculation of a black box model is called modelling: this process starts from the measurements and the system that links the inputs and the output is estimated according the type of regression (gaussian process for gaussian models, least square method for polynomial models,..).

In engine model-based calibration it is possible to distinguish:

- Local model: the load and the engine speed are not inputs, so the model is valid in a single point of LOAD-SPEED plot
- Global model: load and engine speed are model inputs. In this way it is possible to predict the model output in offline mode in every working condition.

#### 3.1. Modelling: polynomial and gaussian processes

The models used in data regression are divided in linear models and nonlinear ones.

The linear models are linear combinations of fixed functions of input variables, and can be written as:

$$y(x, \theta) = \sum_{j=1}^M \theta_j \phi_j(x) = \theta^T \Phi(x)$$

This does not mean that this class of models is linear in the input  $x$ , but rather these models are linear in its model parameters  $\theta$ . Therefore, this class of models shares simple analytical properties and yet can be nonlinear with respect to the input variables.

One advantage of linear modelling is that these can directly obtain a closed-form solution for the minimization of the SSE (sum of square errors), like the Least Square method: generally speaking, given the measured system inputs and outputs the polynomial coefficients are computed minimizing SSE.

If the basis functions  $\phi_j$  are chosen to be:

$$\phi_j(x) = x_j$$

in this case it is a polynomial model.

An example can be:

$$bsfc = 3.45 + 0.21 * SOI + 2 * EGR + 3 * SOI^2 + 2.65 * EGR^2 - 23 * SOI * EGR$$

Instead the nonlinear regression models have in common that are not linear in the model parameter  $\theta$ .

One type of this one can be designed by the Gaussian process. Polynomial regression has some significant drawbacks, since the basis functions  $\phi_j$  have to be chosen in advance, before the model training. However, typically the basis functions suitable for the training data are unknown. Therefore, one idea can be to work with an infinite number of basis functions, which can be achieved with GP regression.

The Gaussian process viewpoint is a non-parametric approach, and this type of modeling is somewhat different than the other modeling techniques.

In order to gain a better understanding of GP regression, the derivation of the formulas is started with the dual representation, in which initially the well known parametric viewpoint is considered (for more detailed information see [8]).

The formula of the model output is:

$$y(x) = k(x)^T * (K + \lambda * I_N)^{-1} * y$$

$k(x)$  is the kernel function,  $K$  is the gram matrix that is function of  $k(x)$  and  $\lambda$  is the regularization parameter.

The solution of the fitting problem can be expressed completely by the kernel function, without explicit calculation of the basis functions. This is the major GP advantage.

A common choice for the kernel function is the squared exponential kernel:

$$k_{SE}(x, x') = \theta_\sigma^2 * \exp \left( - \sum_{j=1}^D \frac{(x_j - x'_j)^2}{2 * \theta_{l,j}^2} \right)$$

with the signal variance  $\theta_\sigma^2$  and the length-scale parameters in each input dimension  $\theta_{l,j}^2$ .

Since the Gaussian process regression is a form of non-parametric modeling, they are called hyperparameters in the area of machine learning.

The length-scale hyperparameters have an interesting property. They can be estimate the values of all hyperparameters out of the training data. In doing so, it is possible that different inputs obtain different values for the length-scale parameters.  $\theta_{l,j}$  becomes high, the function becomes relatively insensitive to the corresponding input variable  $x_j$ .

Hence, with the squared exponential kernel it becomes possible to detect input variables that have little or much effect on the model [9].

Therefore, we are able to interpret the model also from a physical viewpoint. Inputs which have a high or low value for  $\theta_{l,j}$ , have a low or high nonlinear behavior. This determination of the importance of a certain input is called automatic relevance determination, and it is well known in machine learning. Instead of going into more detail on this technique, it is referred to the literature [8].

Polynomial regression has several advantages compared to Gaussian processes. Polynomials have a simple form, are well known and easy to understand. Further, as polynomial regression is a special form of linear modeling, this modeling is computationally cheap and easy to implement. In order to avoid overfitting, statistical tests can be used (stepwise). These tests remove parameters which are not significant and therefore not needed in the model. In this way, a big set of admissible basis functions can be chosen for regression, which increases the potential flexibility of the modeling, without the fear of obtaining overfitting.

However there are some drawbacks of polynomial regression in theory and practice.

One disadvantage of polynomial regression is a bad extrapolation of the data. Polynomials, which are not constant over the whole input space, tend very fast to high (positive or negative) values outside the region of the measurement data. In comparison to that, using the SE kernel, Gaussian processes tend to the mean of the data, if every measurement is far away from the prediction.

Another drawback is, that polynomials of high order tend to waviness and 'end-effects'. This can be illustrated by Runge's phenomenon, which describes the problem of oscillation at the edges of an interval.

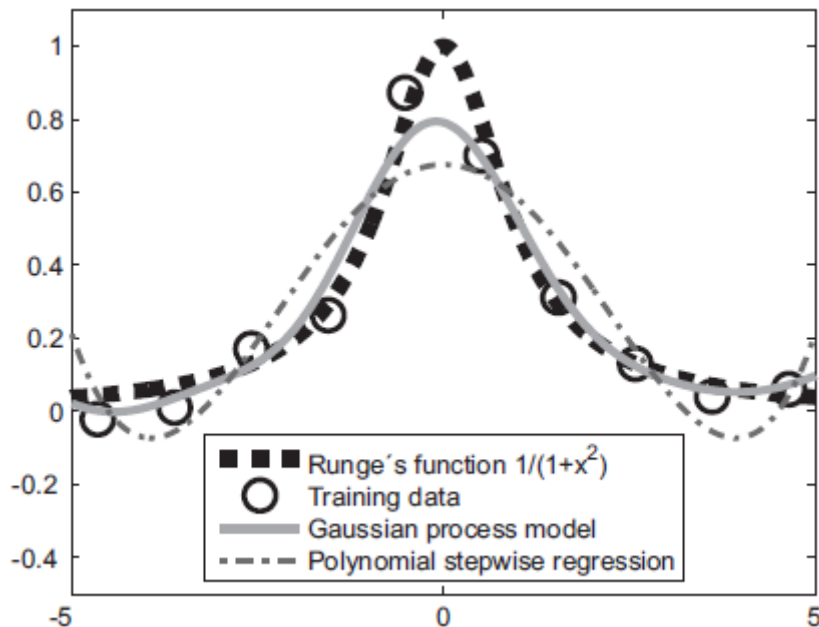


Figure 3.1 Runge's function

The big advantage of GP regression is the use of kernel functions, which can express an infinite dimensional vector of basis functions

However the gaussian models often suffer of overfitting: the model matches very well the training data but with "unnatural way" that makes the model poor in prediction. The following figure shows an overfitting example: the coefficient of determination  $R^2$  tends to 1 but the model is too conditioned by training data and will have insufficient predictivity quality.

Seeing the response surface view is a method to detect overfitting. Another possible method can be to compare the RMSE and the PRESS RMSE (see the chapter "statistic indicators" for more details): the overfitting is probable if there is one or more order of magnitude of difference between these values.

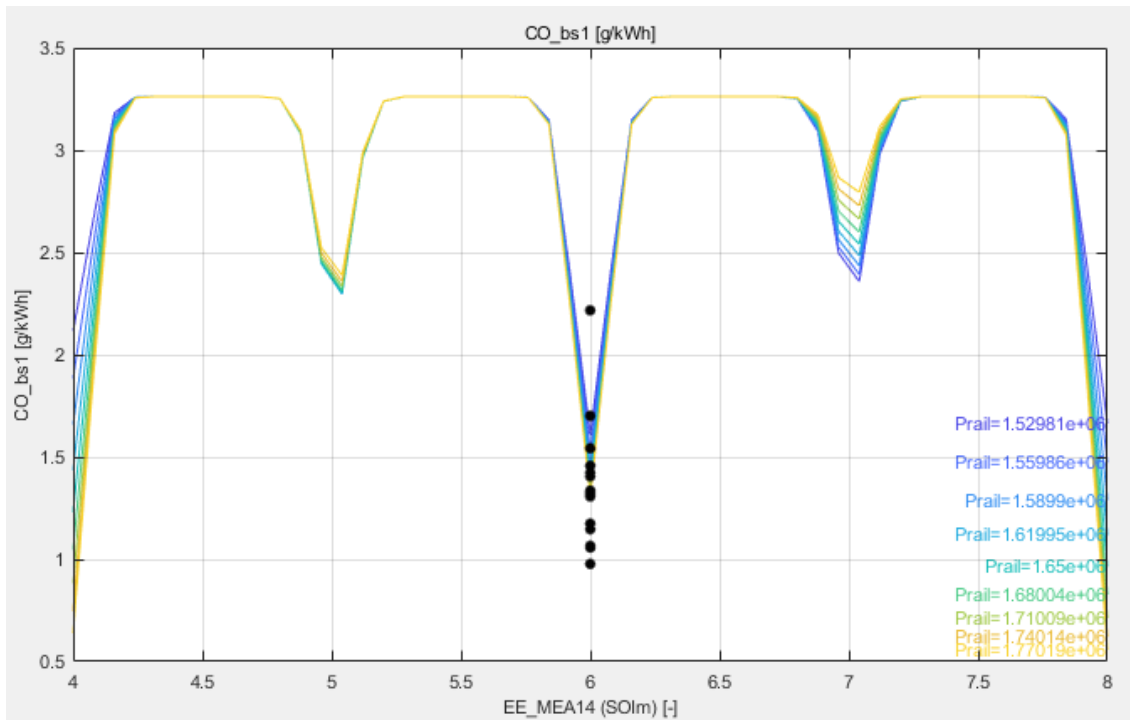


Figure 3.2 Multiline plot of brake specific CO in function of SOI for different values of rail pressure: overfitting

### 3.2. Outliers detection

Outliers are observations that are very different from the bulk of the data. For example, consider the data in Figure 3.3 : observation A seems to be an outlier because it falls far from the line implied by the rest of the data.

The outlier may be a “bad value” that has resulted from a data recording or some other error. On the other hand, the data point may not be a bad value but it is only badly predicted by the model and may be a highly useful piece of evidence concerning the process under investigation [10].

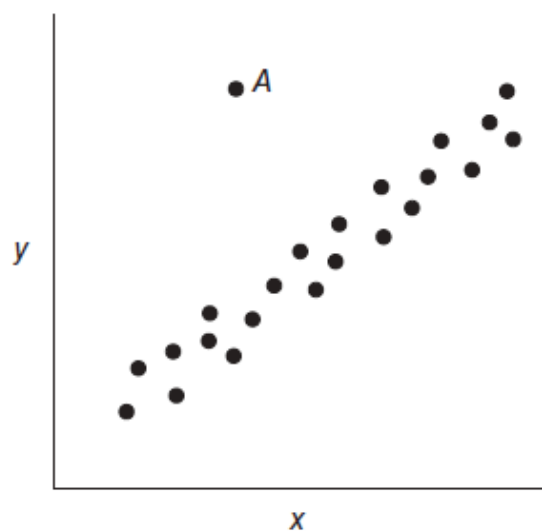


Figure 3.3 A possible outlier example

Detecting outliers is a significant problem that has been studied in various research and application areas. Researchers continue to design robust schemes to provide solutions to detect outliers efficiently.

There are different methods to identify outliers. The simplest approach is graphical: the experimenter checks the INPUT/OUTPUT graph (like the previous figure) in order to find data that are abnormal like A.

However this method is only figurative and the risk is to cancel a test that can be precious for the model.

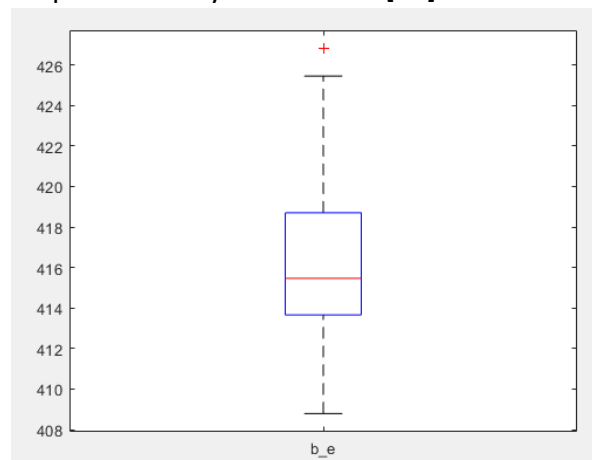
For model-based calibration purpose, a more reliable and impartial method is necessary. So the statistics and the data analysis are two important subjects to use.

A simple approach can be the data quartiles representation by the box plot. It is a pre-modeling approach: it describes simultaneously the most important characteristics like the median value, the dispersion and the identification of the data far from the median value.

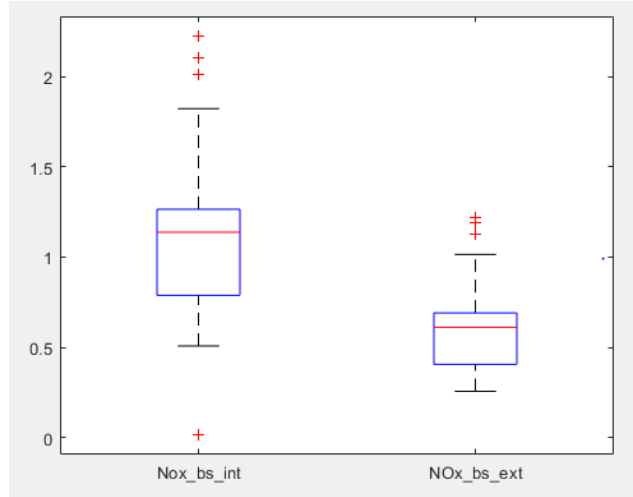
Knowing the median value (value that splits the data in two equal parts,  $q_2$ ), the first quartile (it divides the data between the 25% and 75% of them,  $q_1$ ) and the third quartile (it divides the data between the 75% and 100% of them,  $q_3$ ) the box can be plotted with a line inwardly. The distance between  $q_3$  and  $q_1$  is called interquartile distance (IQR) and is the box long side.  $q_2$  is the red line.

Now the whiskers can be defined: the low whisker connects  $q_1$  to the bigger value within  $1,5 \cdot \text{IQR}$  from  $q_1$ ; the high whisker connects  $q_3$  to the lower value within  $1,5 \cdot \text{IQR}$  from  $q_3$ . The value that are outside the box and the whiskers are outliers.

In the figure these data are represented by red crosses [11].



*Figure 3.3 Box plot of brake specific fuel consumption*



*Figure 3.4 Box plots of brake specific NOx at intake and exhaust manifolds*

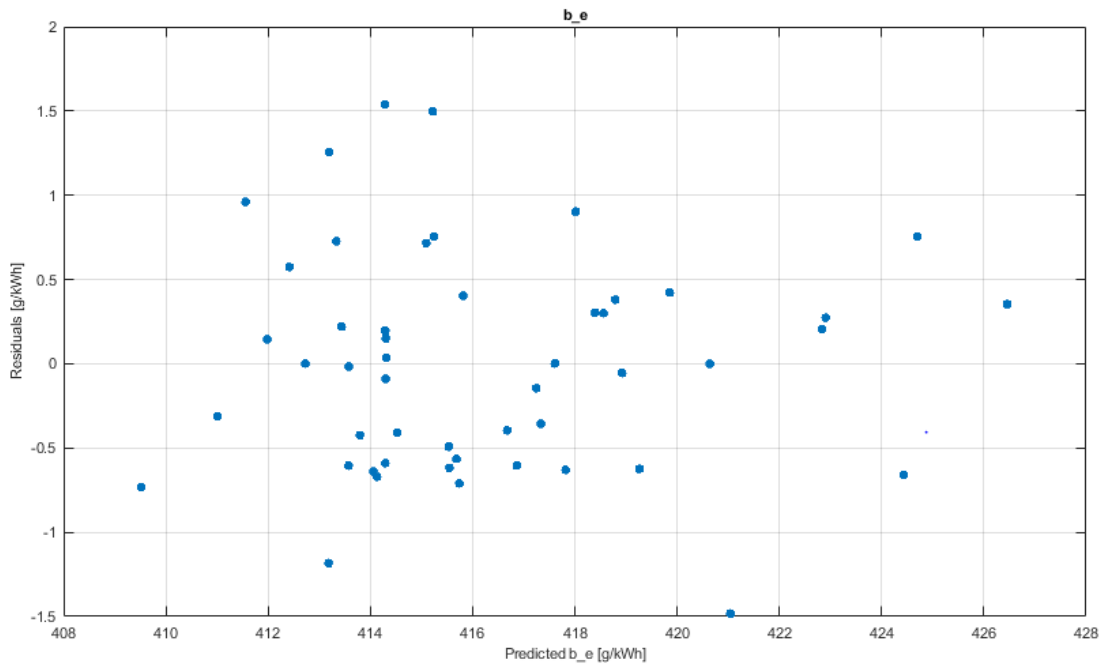
The main drawback is that it is only graphic and does not take into account the modelling results. Furthermore, this method eliminates data taking into account the median value of the considered output and the values surrounding it. In engine calibration, for example in a single engine point (fixed load and speed), the emission values may vary widely, by varying only some input parameters such as SOI or the opening of the EGR valve. This method could therefore discard a priori data of this kind, even if they are not “bad” data and which can give very useful information for the model creation.

To overcome this limitation and not discard data that may be useful, it is necessary to use another method that also takes into account the modelling results and it is necessary to analyze other statistical indicators taking into account the inputs that generate the measured output.

A widely used approach is the detection from residual plot. The residual formula is:

$$e_i = y_i - \hat{y}_i$$

The residual  $i$  is the difference between the output data  $i$  and the model predicted output  $i$ . So each observation has its residual.



*Figure 3.5 Residual plots of brake specific fuel consumption*

In the figure 3.5 the predicted brake specific fuel consumption is plotted versus the residuals. This is the most common choice. Another useful view can be the plot Observation number vs residuals: for example if the residuals are constant in a value or grows up with the observation number may mean that after a certain time same test bed measurement devices have not worked correctly.

The drawback of this approach is the lack of objectivity: the experience of the operator decides what is the threshold for each output (maximum value for the residual), that is different for each of them. For example in a data set the brake specific fuel consumption mean value can be 450g/kWh. The operator will use a threshold residual value of  $\pm 10\text{g/kWh}$ . If he takes into account the brake specific NOx (assuming 40g/kWh as mean value) he cannot use  $\pm 10\text{g/kWh}$  because it is very relevant respect the measured values.

It is necessary another indicator that is more objective and usable for each output regardless its range and its mean value.

The approach used by MBC analyzes the studentized external residuals.

The studentized residual is a residual that is normalized according the error variance. This last quantity is:

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - p}$$

Where the numerator is the SSE (sum of square errors), n is the number of tests and p is the number of model parameters.

Moreover the measured outputs and the model ones can be connected by the hat matrix H:

$$\hat{y} = H * y$$

Hat matrix because gives the hat at each output.

$$\hat{y}_1 = h_{11} * y_1 + h_{12} * y_2 + \dots + h_{1n} * y_n$$

$$\hat{y}_2 = h_{21} * y_1 + h_{22} * y_2 + \dots + h_{2n} * y_n$$

...

$$\hat{y}_n = h_{n1} * y_1 + h_{n2} * y_2 + \dots + h_{nn} * y_n$$

The  $h_{ij}$  terms are only functions of the inputs X and are very simple to compute (see [10] for more details). Moreover  $h_{ji} = h_{ij}$  and the diagonal terms are:

$$0 < h_{ii} \leq 1$$

The  $h_{ii}$  quantifies the distance of the point  $(x_{i1}, x_{i2}, \dots, x_{in})$  from the mean of all points belonging to data set. Accordingly high value means that this data is far from the data center. In the figure 3.6 this concept is explained by a DoE with two inputs: X1 and X2. The highlighted point has inputs values that are much higher than average. So it has high value of h.



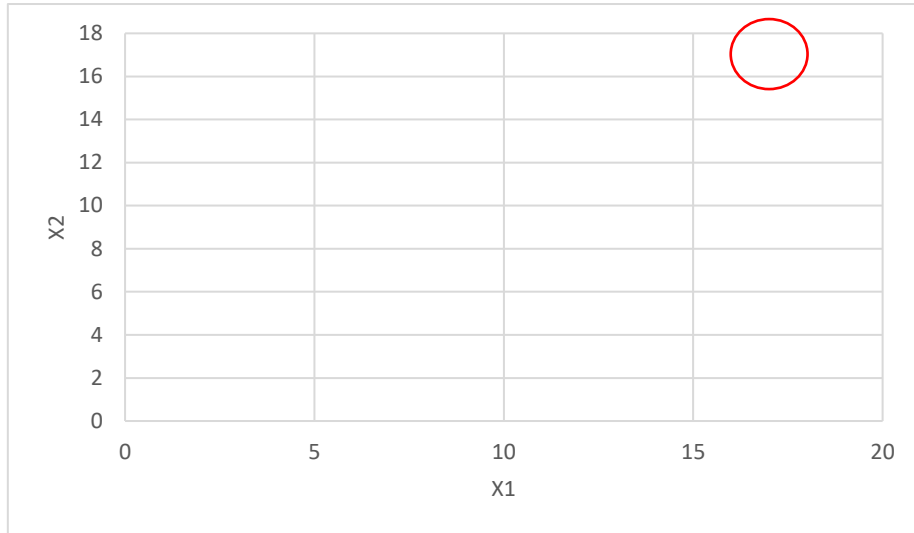


Figura 3.6 DOE with 2 inputs(X1 and X2)

The studentized residual  $i$  related to the test  $i$  is:

$$r_i = \frac{e_i}{\sqrt{\widehat{\sigma}^2 * (1 - h_{ii})}}$$

The quantity at the denominator is the correct standard error of data  $i$ .

The studentized residual can be internal or external.

In the internal case, all the data are used for the calculation of error variance.

But if the  $i^{\text{th}}$  residual is suspected of being improbably large, it is prudent to exclude the  $i^{\text{th}}$  observation from the process of estimating the variance.

In this case the error variance is different for each test and is:

$$\widehat{\sigma}_{(i)}^2 = \frac{\sum_{j=1, j \neq i}^n (y_j - \hat{y}_j)^2}{n - p - 1}$$

This formula is used for external studentized residual computation.

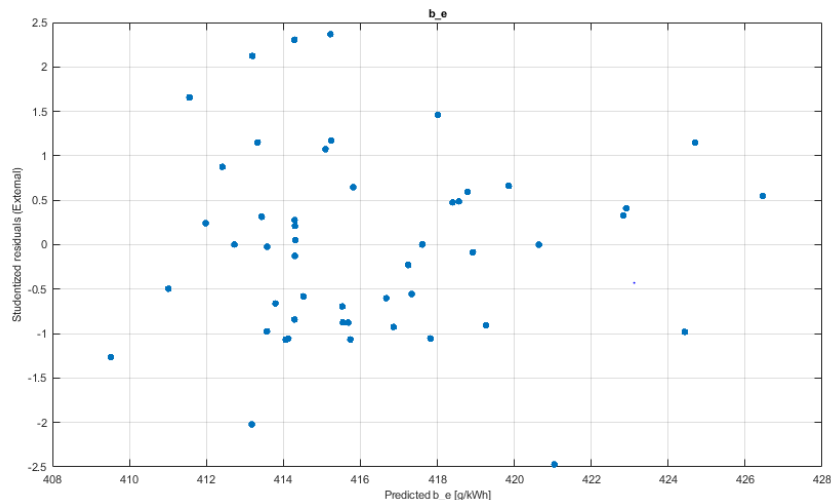


Figure 3.7 Studentized(external) residuals plot of brake specific fuel consumption

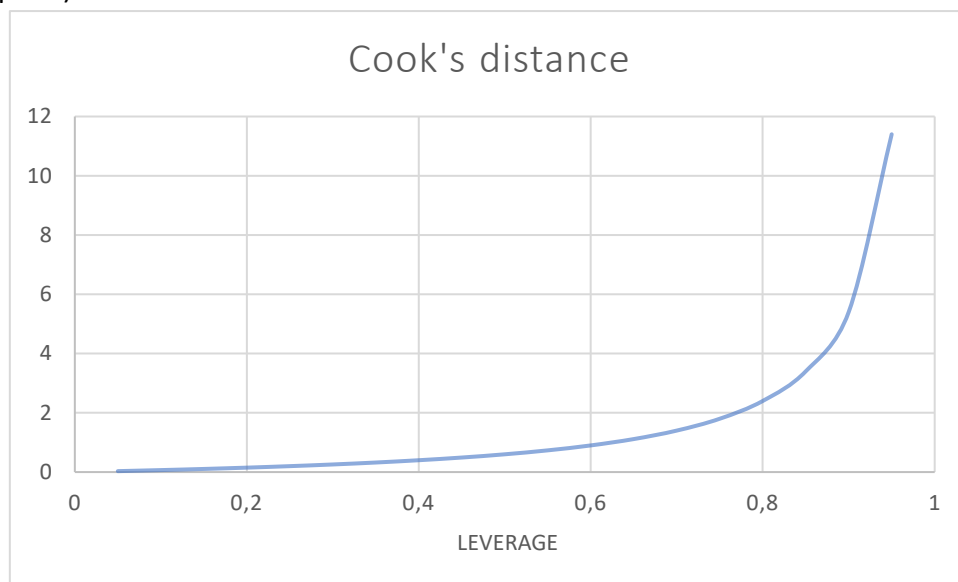
The cut-off value for MBC is 3. This approach is absolutely more efficient than residual approach because a more objective indicator is adopted: it is normalized by the correct standard error for each data, so it is a pure number. By this it is possible to adopt a cut-off value that can be used for each output (bsfc, bsNOx, MFB50, Noise,...).

However this strategy for outlier detection deletes the data considering only the modelling capability. For example it is possible that a data, that has  $|r_i| > 3$  (so it is outlier for MBC algorithm) can be a data that is not well predicted by the model and including it can be very important: its elimination can decrease the model predictivity.

To avoid the elimination of important test, one way is to measure the observation influence. For this aim an excellent diagnostic tool is the Cook's distance:

$$D_i = \frac{r_i^2}{p} * \frac{h_{ii}}{1 - h_{ii}}$$

Clearly if the  $i^{\text{th}}$  data is influent, its elimination will involve a great variation for the model. The first factor contains the square of studentized residual that suggests how the model is suitable to  $i^{\text{th}}$  observation; the second term indicates how the  $i^{\text{th}}$  observation inputs is far from the data rest. In the figure 3.8 shows as the Cook's distance grows up with the leverage  $h_{ii}$  (fixed studentized residual=2, p=15)



*Figure 3.8 Cook's distance versus leverage*

According the literature, the cut-off value for this indicator is 1: if an observation has  $D_i > 1$ , it is defined influential. This means that is precious for the model creation and it is recommended not to take it off [11].

Another way to detect an influential observation is to observe the Cook's distance of each observation: if all the distances are lower than 1 but there are observation where  $D_i$  is enough higher than the others, it is convenient to take into account them as influential observations.

In the figure 3.9 the Cook's distances are plotted for the brake specific NOx model. The highest value is about 0.4: it is less than 1. Theoretically if it had a  $r_i$  greater than 3, it should be eliminated. However it has the most influential data and so it is counterproductive for the model its elimination. On the contrary the point circled in red (because its studentized residual is higher than 3) has a low

value of Cook's distance and this value is practically on the average. In conclusion this data can be an outlier and its elimination wouldn't get worse the model quality.

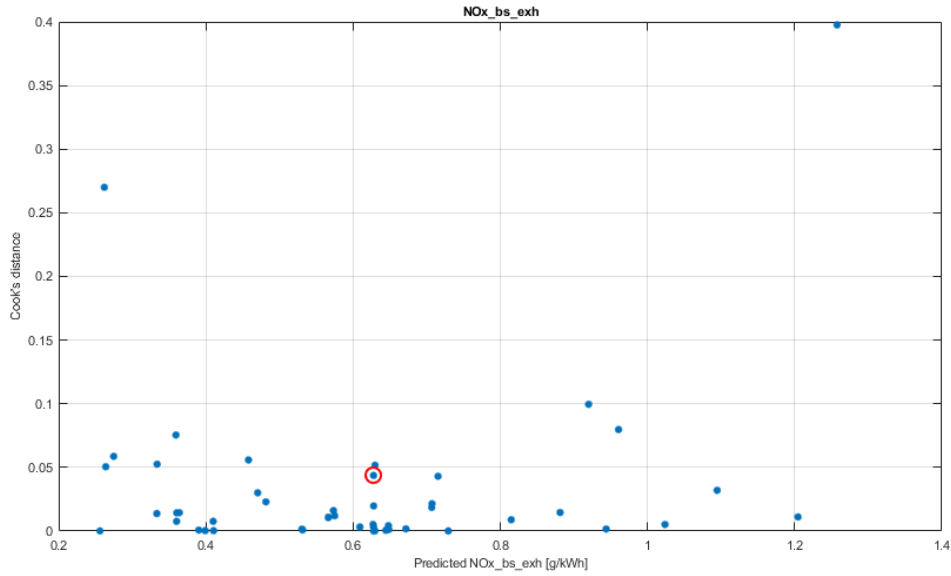


Figure 3.9 Cook's distance of brake specific NOx model

In the case of figure 3.10 the outlier is detected for a data that has the highest  $D_i$ . So its exclusion can be risky for the model quality because it is an important observation. The suggestion is to maintain this observation.

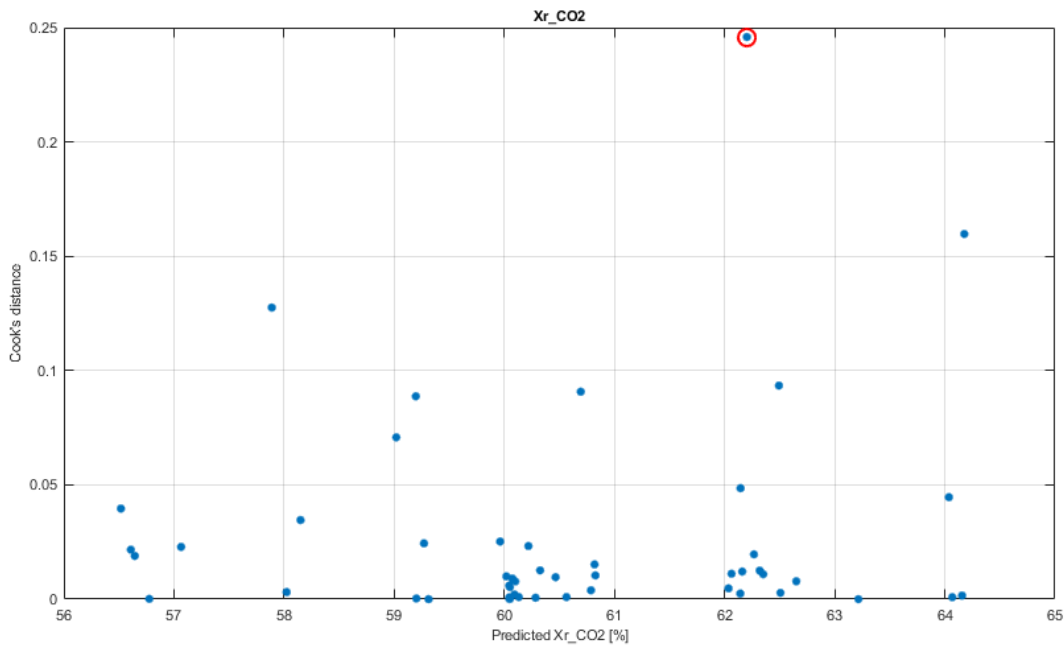


Figure 3.10 Cook's distance of EGR model

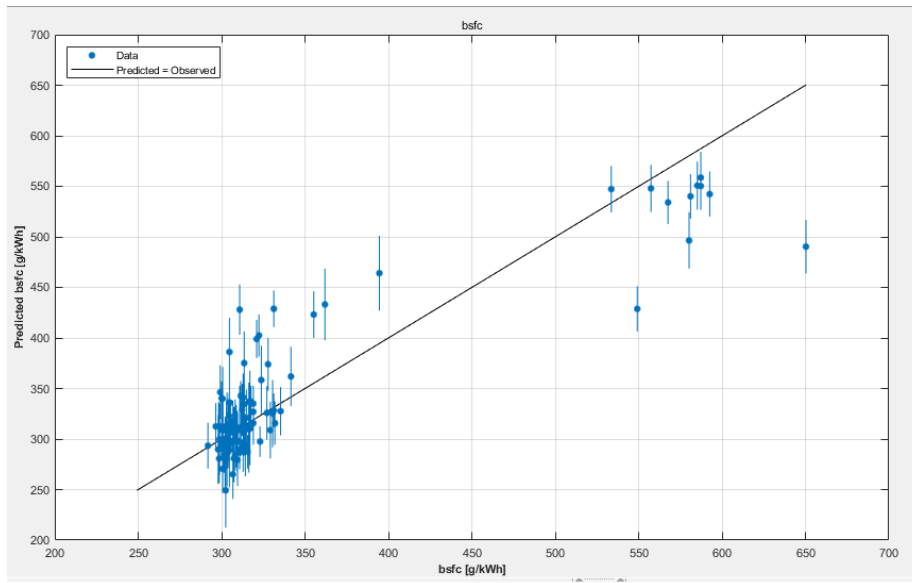
At the end the following algorithm can be implemented.

The  $i^{\text{th}}$  observation may be eliminated if respects all the following equations:

- $|r_i| > 3$  (high discrepancy between model and observation)
- $D_i < 1$  (low importance of the observation in absolute value)
- $D_i < K * D_{MAX}$  (low importance of the observation in relative value)

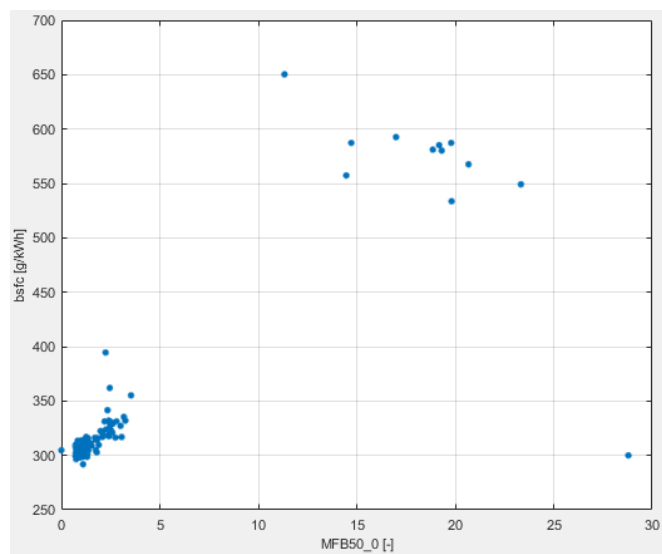
The parameter K is calibrated by some practical test but a good value of it should be 0.5.

However it is important that, first to create the model and find the outliers, the user scans the raw data in order to find measurement errors or wrong test series. The following example (figure 3.11) represents a possible issue if the previous recommendation is bypassed: the bsfc in a operating point is modeled and two clouds of data are present in the graph Predicted/Observed. For an engineer is obvious that a value of bsfc higher than 500g/kWh is very uncommon. Probably there were some problems during these tests running. However the program (or the algorithm) cannot define them as outliers because there are quite a large number and the model construction is highly influenced and fitted by them.



*Figure 3.11 Predicted/Observed graph of bsfc*

The user for example can check some data that highlight the strangeness like the CoV between the cylinder of same outputs. A typical output for this type of analysis is the MFB50 (the crank angle after the top dead center where the combustion of the 50% of fuel is completed) like the following figure.



*Figure 3.12 MFB50\_CoV versus bsfc*

To have a stable combustion process, the maximum value of CoV is 4%-5%. In this case there are value much higher. Moreover they are points where the bsfc is very high: definitely these are points very unstable. So they are detrimental for the model and they must be eliminated before fitting the model.

In the figure there is also a point with a bsfc equal to 300g/kWh and a very high MFB50\_CoV (about 29%). It is possible that can be a error in CoV calculation or this is a bad test. However the good choice is to maintain it and to refit the model. Without the eliminated tests, the model is “purified” by a big number of instable tests. So with the outlier algorithm, the user can now try to find outlier and decide if it is one.

### 3.3. Statistical tool

This chapter is dedicated to statistical tools used during the modelling to refine and study its quality.

#### 3.3.1. Statistical indicators

To evaluate the model quality, it is necessary to check some statistical parameters.

The first term to take into account is the coefficient of determination:

$$R^2 = \frac{SS_{reg}}{SS_{tot}} = 1 - \frac{SSE}{SS_{tot}}$$

Where:

$$SS_{reg} = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$$

is the sum of squares of deviations between predicted (calculated by means of regression) values and mean value .

$$SS_{tot} = \sum_{i=1}^n (y_i - \bar{y})^2$$

is the sum of squares of deviations between measured values and mean value.

$$SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

is the sum of squares of error .

From the previous definition:

$$SS_{tot} = SS_{reg} + SSE$$

The coefficient of determination varies from 0 and 1 and shows how precisely the model fits to the measured values. More the predicted values are close to measured ones, SSE is smaller and  $R^2$  tends to one

For linear models with p as the parameters number, a variation of  $R^2$  is the adjusted  $R^2$ :

$$R_{adj}^2 = 1 - \frac{SSE * (n - 1)}{SS_{tot} * (n - p)}$$

For the same model this indicator has a lower value than  $R^2$  taking into account the degrees of freedom of the model equation. This is an important element to identify overfitting: for example a 4<sup>th</sup> order polynomial model can have a value of  $R^2$  close to 1, but a value of  $R^2_{adj}$  lower than 0.7. This means that the data fitting is excellent but it is due to a high order model and probably for different inputs respect the training data the predictions are poor (overfitting).

The PRESS (predicted residual sum of squares)  $R^2$  is:

$$PRESS R^2 = 1 - \frac{PRESS}{SS_{tot}}$$

$$PRESS = \sum_{i=1}^n (y_i - \hat{y}_{(i)})^2$$

The term  $\hat{y}_{(i)}$  (predicted value for the  $i^{th}$  measurement) is the predicted model value  $i$  without taking into account the value  $i$  in the model construction.

This indicator can sense the model predictivity: in general  $PRESS R^2 \geq 0.7$  is the requirement for a good model.

Two additional estimators help to analyze more extensively the goodness of fit for the model.

The root mean square error (RMSE) is a measure of accuracy to quantify the differences between values predicted by a model and the values actually observed: low RMSE value reflects greater accuracy.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}$$

With previous considerations, it is possible to define the PRESS RMSE:

$$PRESS RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_{(i)})^2}{n}}$$

The benefit of these terms is its unit of measure that is the same of  $y$  so it is directly comparable.

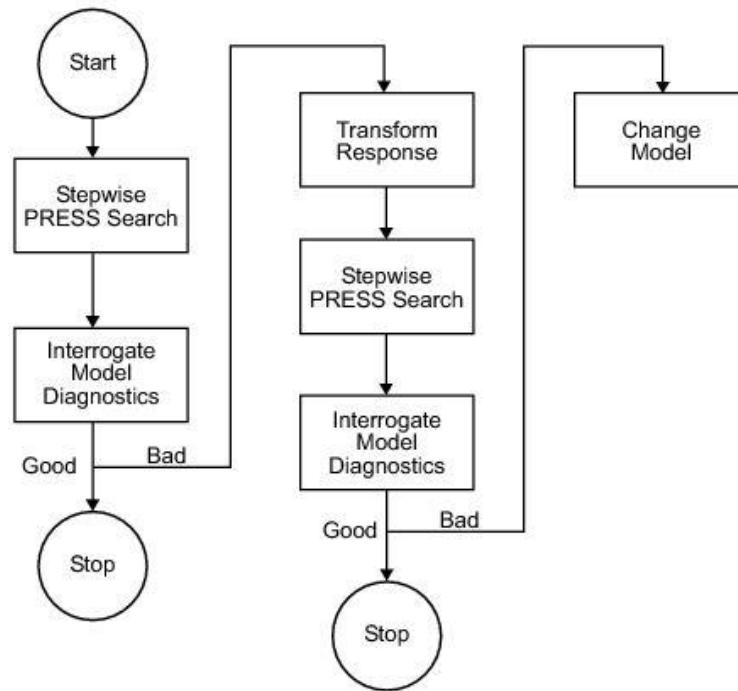
### 3.3.2. Stepwise regression

Stepwise function helps to search for a good model fit. The goal of the stepwise search is to minimize PRESS. Minimizing Predicted Error Sum of Squares (PRESS) is a good method for working toward a regression model that provides good predictive capability over the experimental factor space.

The PRESS statistic gives a good indication of the predictive power of model: it is useful to compare PRESS RMSE with RMSE as this may indicate problems with overfitting. RMSE is minimized when the model gets very close to each data point; 'chasing' the data will therefore improve RMSE. However chasing the data can sometimes lead to strong oscillations in the model between the data points; this behavior can give good values of RMSE but is not representative of the data and will not give

reliable prediction values where you do not already have data. The PRESS RMSE statistic guards against this by testing how well the current model would predict each of the points in the data set (in turn) if they were not included in the regression. To get a small PRESS RMSE usually indicates that the model is not overly sensitive to any single data point [12].

Once the user sets up the model, he should use the stepwise function and examine the diagnostic statistics to search for a good model fit. The process is pointed out by the following flow chart:



*Figura 3.13 Flow chart of linear model construction [12]*

1. Begin by conducting a stepwise search(automatically or using the stepwise window in MBC).The goal of the stepwise search is to minimize PRESS.
2. Remove outliers at this stage. The criteria is subjective, however in the previous chapter an algorithm is recommended.
3. After removing outliers(if there are), review the stepwise window and minimize again the PRESS. If the minimization of PRESS has been implemented in the first phase of model set up, this step is done automatically by the program
4. A transform of the response feature might prove beneficial. A useful set of transformations is provided by the Box and Cox family. It is recommended a transformation if the statistic indicators are still deficient. After repeat the stepwise PRESS search. This procedure has been not recommended at first attempt and the reason are explained in the next chapter. Then remove the outliers if there are.
5. After this if the statistic indicators of the model are still insufficient, the current model is probably inadequate. The user should try to use another model type.

There is also another final conclusion: the data cannot be fitted by any model, for example in situation where the measurement are very compromised by the uncertainty of the instrument. An example can be the measurement of soot in a PCCI(premixed combustion compression ignition)

combustion at low load. In the figure the FSN of 50 different tests is showed. The value are too low and too influenced by the uncertainty: no models can describe this output.

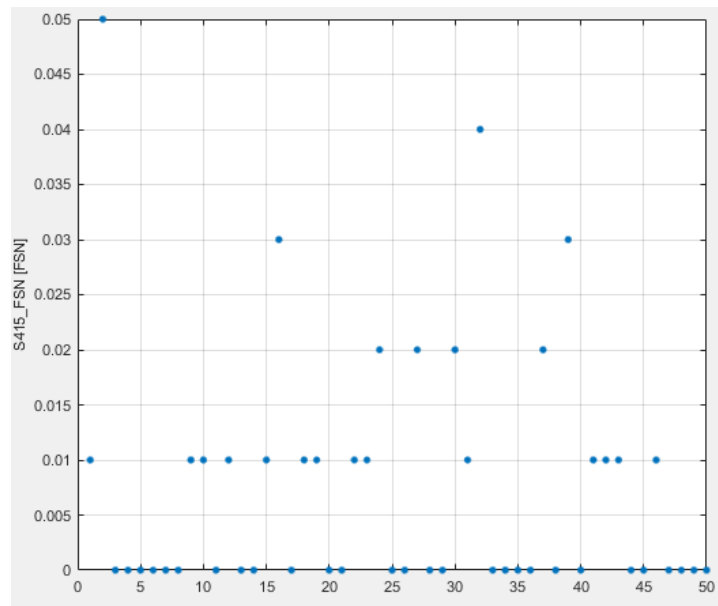


Figure 3.14 Filter smoke number measurements over 50 tests (PCCI 1800rpmx27Nm)

In this field another example can be the measurement of unburned hydrocarbon in a normal diesel combustion at medium/high load. It is possible to have measurement about 5-10 ppm that are comparable to measurement uncertainty .

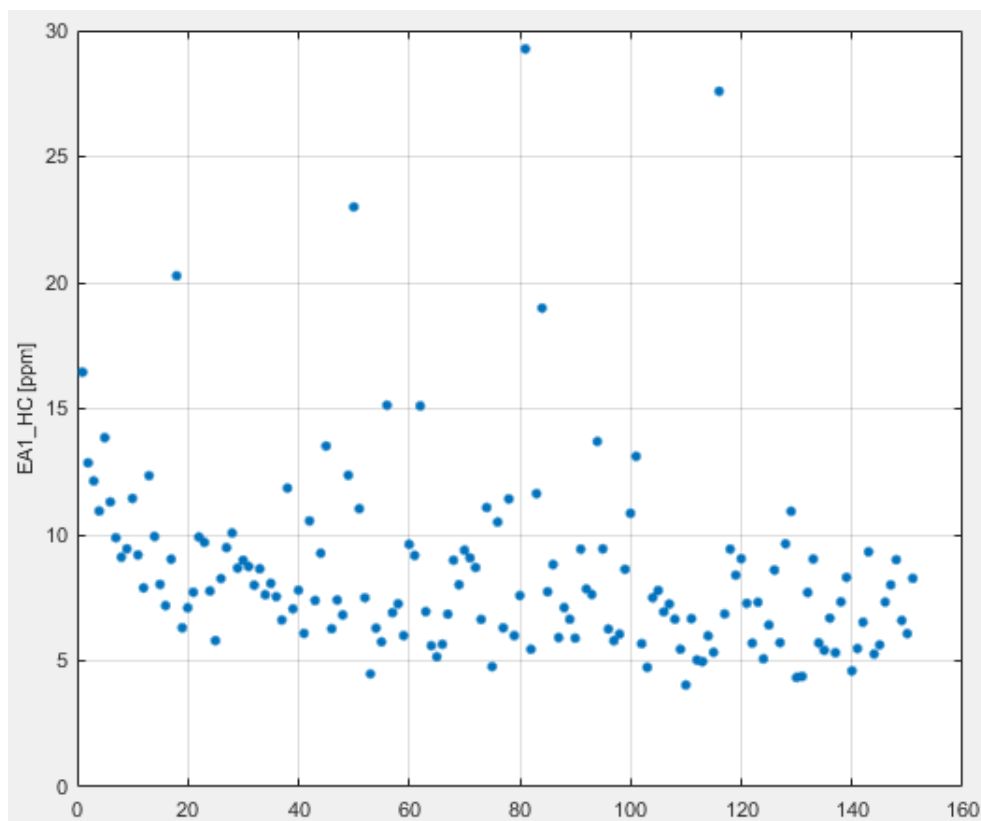


Figure 2.94 HC(ppm) measurement over 150 tests (2700rpmx12bar)



### 3.3.3. Box-Cox transformation

Generally, transformations are used for three purposes: stabilizing response variance, making the distribution of the response variable closer to the normal distribution, and improving the fit of the model to the data. This last objective could include model simplification, say by eliminating interaction terms. Sometimes a transformation will be reasonably effective in simultaneously accomplishing more than one of these objectives. Power family of transformations  $y^* = y^\lambda$  can be very useful, where  $\lambda$  is the parameter of the transformation to be determined (e.g.,  $\lambda = 0.5$  means use the square root of the original response). Box and Cox (1964) have shown how the transformation parameter  $\lambda$  may be estimated simultaneously with the other model parameters (overall mean and treatment effects). The theory underlying their method uses the method of maximum likelihood. The actual computational procedure consists of performing, for various values of  $\lambda$ , a standard analysis of variance on:

$$y^{(\lambda)} = \begin{cases} \frac{y^\lambda - 1}{\lambda * \dot{y}^{\lambda-1}} & \lambda \neq 0 \\ \dot{y} * \ln(y) & \lambda = 0 \end{cases}$$

Where  $\dot{y} = \ln^{-1}[\frac{\sum \ln(y)}{n}]$  is the geometric mean of the observations. The maximum likelihood estimate of  $\lambda$  is the value for which the error sum of squares SSE is a minimum. This value is usually found by plotting a graph of SSE versus  $\lambda$  and then reading the value that minimizes SSE from the graph.

A problem arises in  $y$  when  $\lambda$  tends to 0, and the fraction tends to infinite. The component  $\frac{y^\lambda - 1}{\lambda}$  alleviates this problem because as  $\lambda$  tends to zero,  $\frac{y^\lambda - 1}{\lambda}$  goes to a limit of  $\ln(y)$ . The divisor component  $\dot{y}^{\lambda-1}$  in the equation rescales the responses so that the error sums of squares are directly comparable. In using the Box-Cox method, it is recommended that the experimenter use simple choices for  $\lambda$  because the practical difference between 0.5 and 0.58 is likely to be small, but the square root transformation ( $\lambda=0.5$ ) is much easier to interpret. Obviously, values of  $\lambda$  close to unity would suggest that no transformation is necessary. SSE versus  $\lambda$  has in general a parabolic behaviour, like the figure 3.15:

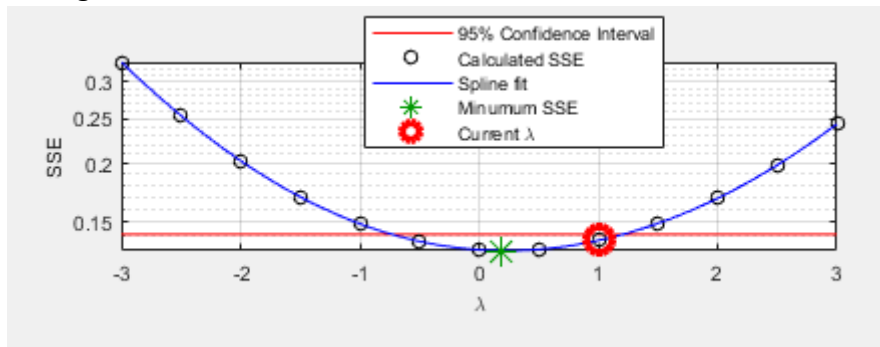


Figure 3.15 Box-Cox transformation: SSE versus  $\lambda$  on MBC

MBC computes the SSE varying  $\lambda$  from -3 to 3 with a step of 0.5. In this case the best choice is  $\lambda$  equal to 0.

An approximate 95% confidence interval for  $\lambda$  can be found by computing:

$$SSE^* = SSE(\lambda_{\text{minimum } SSE}) * (1 + \frac{t_{\alpha, \frac{2}{v}}^2}{v})$$

Where  $v$  is the number of degrees of freedom (number of tests minus number of model parameters) and plotting a line parallel to the  $\lambda$  axis at height  $SSE^*$  on the graph (red line in the previous figure). The value of  $\lambda$  that has SSE lower than  $SSE^*$  are better than the others for the model at 95% of confidence. If this confidence interval includes the value  $\lambda = 1$ , this implies that the data do not support the need for transformation.

In the previous example under the CI there are: -0.5, 0, 0.5 and 1. The best choice is 0 for SSE minimization, however it is preferable to maintain  $\lambda = 1$ , because it is under the CI of  $SSE(\lambda = 0)$  and it is preferable not to transform if possible.

Indeed the Box-Cox transformation is suggested only in a second moment, after that the model (after stepwise process for a linear model) is still not satisfactory. There are two main reasons for this:

- The normal output function ( $\lambda=1$ ) possesses a natural engineering interpretation. It is unlikely that the behaviour of a transformed version of a response feature is as intuitively easy to understand.
- Outliers can strongly influence the type of transformation selected. Applying a transformation to allow the model to fit bad data well does not seem like a prudent strategy [13].

### 3.3.4. Prediction error variance view

Prediction Error Variance (PEV) is a very useful way to investigate the predictive capability of model. It gives a measure of the precision of a model's predictions. There are the PEV for designs and for models. It is useful to remember that:

$$PEV(\text{model}) = PEV(\text{design}) * MSE$$

So the accuracy of the model's predictions is dependent on the design PEV and the mean square errors in the data. A low PEV (close to zero) means that good predictions are obtained at that point.

If the design  $PEV < 1$ , then the errors are reduced by the model fitting process. If design  $PEV > 1$ , then any errors in the data measurements are multiplied. Overall the predictive power of the model will be more accurate if PEV is closer to zero.

Starting with the design matrix, for example, for a quadratic in SOI (start of injection) and P\_rail (rail pressure), it is a  $N \times 6$  matrix where  $N$  is the number of tests:

$$X = \begin{bmatrix} 1 & SOI_1 & P\_rail_1 & P\_rail_1 * SOI_1 & SOI_1^2 & P\_rail_1^2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & SOI_n & P\_rail_n & P\_rail_n * SOI_n & SOI_n^2 & P\_rail_n^2 \end{bmatrix}$$

The actual model can be written as:

$$Y = \beta * X + \varepsilon$$

where  $\varepsilon$  is the measurement error with variance:

$$var(\varepsilon) = MSE$$

The predicted coefficient of the model are:

$$\hat{\beta} = (X^T * X)^{-1} * X^T * y$$

Which have variance:

$$var(\hat{\beta}) = (X^T * X)^{-1} * MSE$$

Let  $x^*$  be the vector that contains the input of a new point for the evaluation:

$$x^* = \{1 \quad SOI^* \quad P\_rail^* \quad SOI^* * P\_rail^* \quad SOI^{*2} \quad P\_rail^{*2}\}$$

Then the model prediction for this point is:

$$\hat{y} = \hat{\beta} * x^*$$

At this point the PEV is:

$$\begin{aligned} PEV(x) = var(\hat{y}) &= (x^* * (X^T * X)^{-1} * X^T) * (X * (X^T * X)^{-1} * x^{*T}) * MSE = \\ &= x^* * (X^T * X)^{-1} * x^{*T} * MSE \end{aligned}$$

The PEV(x) for a design (without MSE) has this effect of the new observations - if it is greater than 1 it will magnify the error, and the closer it is to 0 the more it will reduce the error [13].

In MBC environment it is possible to see the PEV plot: an example in the figure 3.16 This is the PEV of lambda in 3D where in the x and y axis there are SOIm ( start of main injection) and qPil1(fuel quantity in the pilot injection 1).

This is a good trend because the value are really small. To quantify this there are the optimality criteria:

- G is the maximum value of PEV in the space considering all the possible combination of the inputs in the domain region. It is 0.719 and there is also the location of this point
- V is the average value of PEV in the space considering all the possible combination of the inputs in the domain region.
- A e D are not value related to PEV but only to DOE design(see previous chapter)

In general the PEV plot has a “cup” shape in 3D for each combination of inputs. In particular it has the minimum in the domain middle: this is intuitive because it represent the barycenter of the tests so the expected precision of prediction is the highest.

On the other hand, the point with highest PEV are the ones at the domain extremes because these points are adjacent only on one side to the other tests and are very far from the tests barycenter.

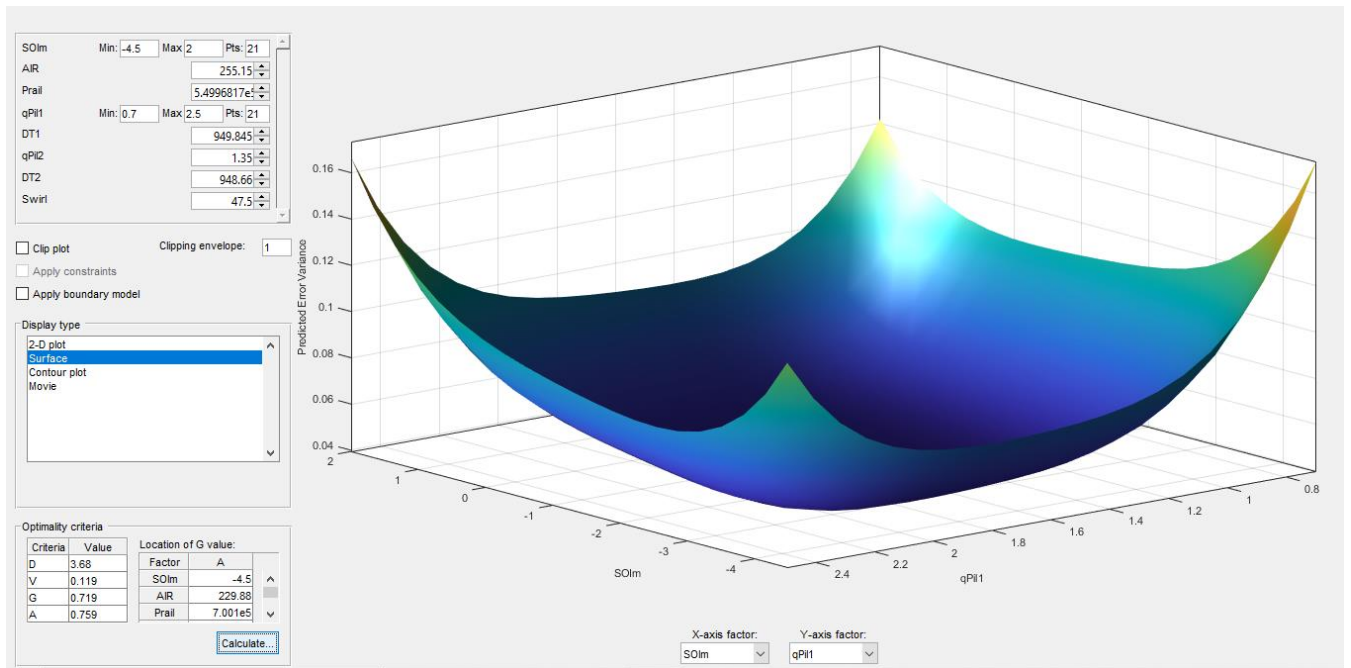


Figure 3.16 Prediction error variance view on MBC

## 4. Modelling

This chapter is dedicated to practical modelling on real data. Different DoE are created for different engines in different working conditions.

### 4.1. 1800 rpm x 27 Nm-PCCI combustion (3L engine)

The first models are created considering a PCCI combustion test. The employed engine is an heavy duty one with the following specifics:

Engine type	3.0L Euro VI 16V
Displacement	2998 cm <sup>3</sup>
Bore / stroke	95.8 mm / 104 mm
Connecting rod	158 mm
Compression ratio	17.5
Valves per cylinder	4
Turbocharger	Single-stage variable geometry turbocharger
Fuel injection system	Common Rail, solenoid injectors

*Table 4.1 3L engine specifications*

Premixed charge compression ignition (PCCI), which can be achieved by using large amounts of exhaust gas recirculation (EGR), and advanced or retarded fuel injection timing, can be considered a promising alternative combustion strategy. Under these working conditions, the combustion process and the mechanism of pollutant formation differ significantly from those of conventional diesel combustion. The combined effect of advancing or retarding fuel injections and using high EGR rates (which reduce the oxygen concentration of the intake charge) leads to a slower pre-ignition chemistry, and to a higher ignition delay. This in turn allows a better pre-combustion mixing than conventional diesel combustion, hence the formation of rich mixture pockets within the cylinder is avoided, which is the main cause of soot generation. Moreover, high EGR rates diminish flame combustion temperatures, and thus lower NO<sub>x</sub> emissions. On the other hand, due to heavily reduced oxygen content, low combustion temperatures and early injections, the formation of incomplete combustion products, such as carbon monoxide (CO) and unburned hydrocarbons (HC), can be significant, and can require a higher conversion efficiency of the diesel oxidation catalyst. Penalties in fuel consumption have also been observed. In addition, due to sharp rises in the in-cylinder pressure, high noise levels are generally related to PCCI combustion [14].

However this combustion can be adopted only at low load, where it is possible to work with high EGR and low oxygen presence.

The models have been created in the operating point 1800rpmx27Nm. The following input parameters have been identified as the ones that mainly affect PCCI combustion: the rail pressure, the start of injection, the position of variable geometry turbine and the position of the backpressure flap valve used to regulate the EGR rate, since the EGR poppet valve was fully open in all the PCCI working conditions. The DoE is created by V-optimal design approach with 50 tests.

There are also 33 tests in the same operation point with a different variation list: they have been used for the model validation.

For every output, 2 alternative models are presented: a second order polynomial model with stepwise and a gaussian one.

	Mean value	R <sup>2</sup>		RMSE		PRESS RMSE		Validation RMSE	
	50 tests	2nd order	gaussian	2nd order	gaussian	2nd order	gaussian	2nd order	gaussian
b_e [g/kWh]	416.4	0.969	0.976	0.739	0.568	0.85	0.899	1.424	1.538
Nox_bs_exh [g/kWh]	0.6	0.995	0.996	0.02	0.016	0.023	0.022	0.024	0.02
CO_bs_exh [g/kWh]	50.67	0.996	0.994	0.535	0.415	0.585	0.59	2.983	2.982
HC_bs_exh [g/kWh]	7.18	0.985	0.993	0.172	0.108	0.193	0.19	0.893	0.976
Noise_1[dB]	85.02	0.985	0.989	0.171	0.133	0.19	0.194	0.488	0.473
MFB50_1 [deg aTDC]	3.2	0.995	0.995	0.133	0.116	0.151	0.165	0.284	0.272
PMAX_2[bar]	51.05	0.994	0.994	0.182	0.152	0.208	0.222	0.69	0.669
Xr_CO2[%]	60.6	0.996	0.996	0.149	0.118	0.171	0.16	0.212	0.206

*Table 4.2 Statistical indicators about 1800x27 PCCI models*

	R <sup>2</sup> adjusted	PRESS R <sup>2</sup>
b_e [g/kWh]	0.961	0.947
Nox_bs_exh [g/kWh]	0.996	0.99
CO_bs_exh [g/kWh]	0.991	0.989
HC_bs_exh [g/kWh]	0.982	0.977
Noise_1[dB]	0.982	0.977
MFB50_1[deg aTDC]	0.994	0.992
PMAX_2[bar]	0.992	0.989
Xr_CO2[%]	0.995	0.993

*Table 4.3 R<sup>2</sup> adjusted and PRESS R<sup>2</sup> related to polynomial model(1800x27 PCCI)*

From the previous tables, the first consideration is the negligible differences comparing the statistical indicators between the polynomial models and gaussian ones.

Secondly from the tables it is possible to see that the statistic indicators that measure how the models fit the measurements are extremely good. From a first view also PRESS RMSE and Validation RMSE (that measure the predictive power) look satisfactory. However they express a mean error value: it may therefore be convenient to analyze the tests individually.

The following plots show the error(in percentage) of the models in the 33 validation tests. The test number is on the x-axis while the error in the y-axis. The error formula is:

$$E[\%] = \frac{y_{i(model)} - y_{i(validation)}}{y_{i(validation)}}$$

In the figure 4.2 the brake specific fuel consumption error is represented: the values are always under the 1%(in absolute value). This means that the bsfc model predicts very well the validation data. This figure also highlights that the difference between the two model typologies is not existent practically.

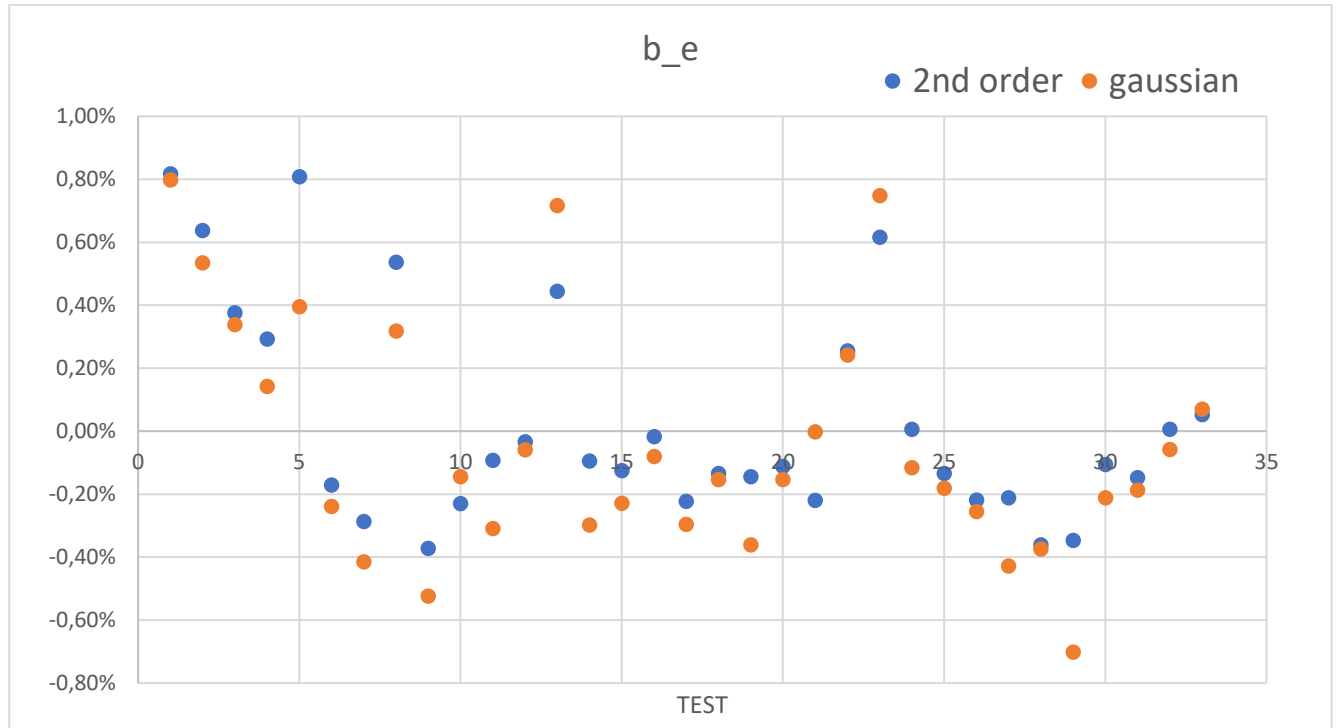


Figure 4.1 error percentage of bsfc validation data

The figure 4.3 describes the bsNOx errors. The values are higher than previous ones on average but there are in a safety margin: the worst prediction is the test 10, where the error is -10%. However this value is quite low and it can be concluded that the NOx models are reliable.

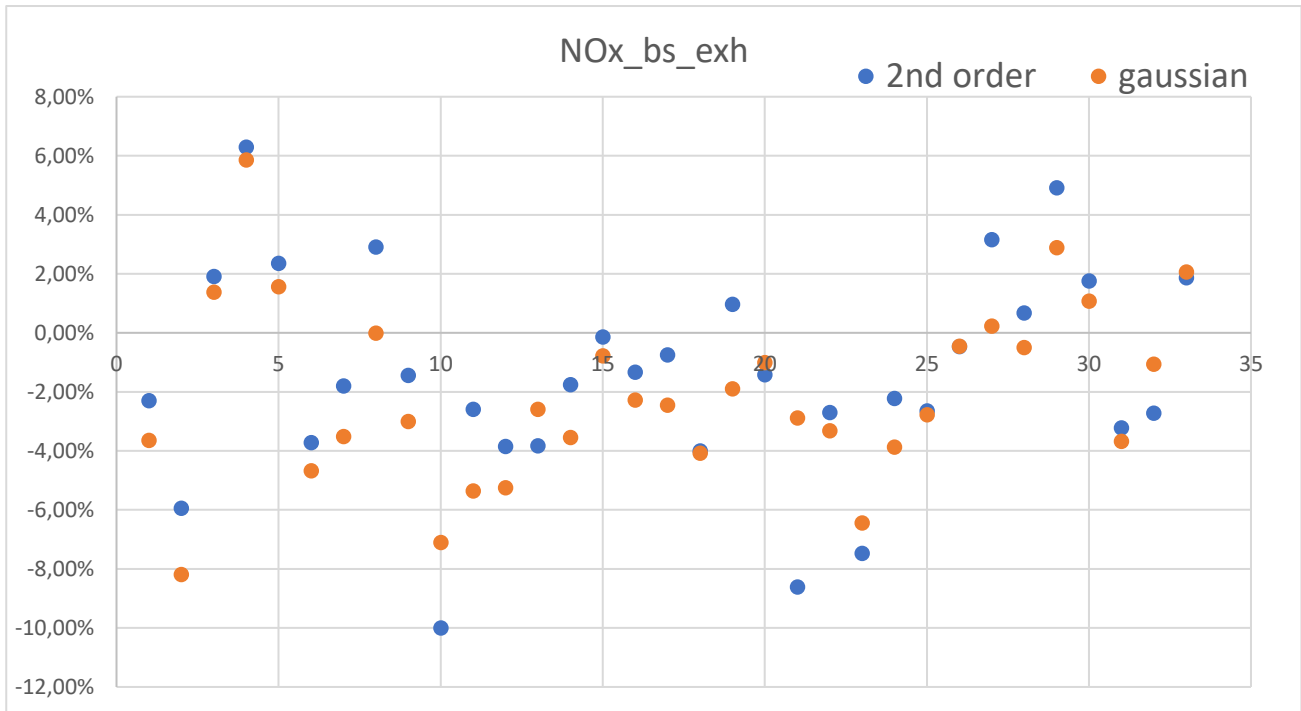


Figure 4.2 error percentage of bsNOx validation data

The following two pictures summarize the validation study showing the mean error and the maximum one over the  $n^*$  (33 in this case) validation tests. They are computed as:

$$\bar{E} = \frac{\sum_{i=1}^{n^*} |E_i|}{n^*}$$

$$E_{MAX} = \max(|E_i|)$$

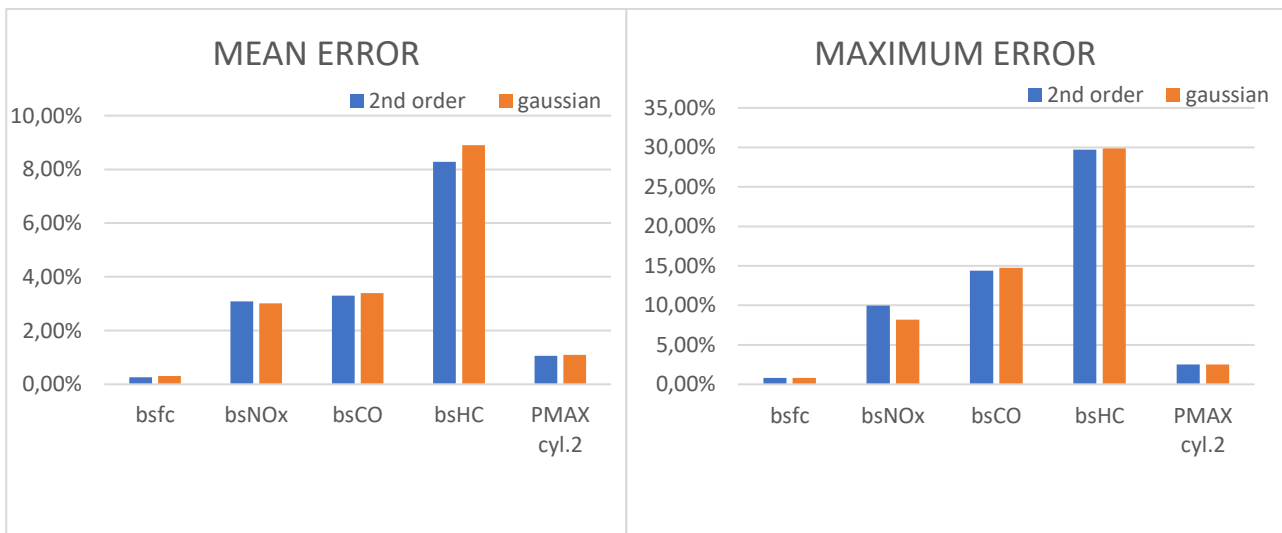


Figure 4.3 Mean and maximum error on validation data: 1800rpmx27Nm PCCI combustion

For the other outputs(Noise1, MFB50\_1, Xr\_CO2) it is preferable to study the mean error difference and the maximum error difference:



$$\text{mean}(ED) = \frac{\sum_{i=0}^{n^*} |y_i - \hat{y}_i|}{n^*}$$

$$\text{max}(ED) = \max(|y_i - \hat{y}_i|)$$

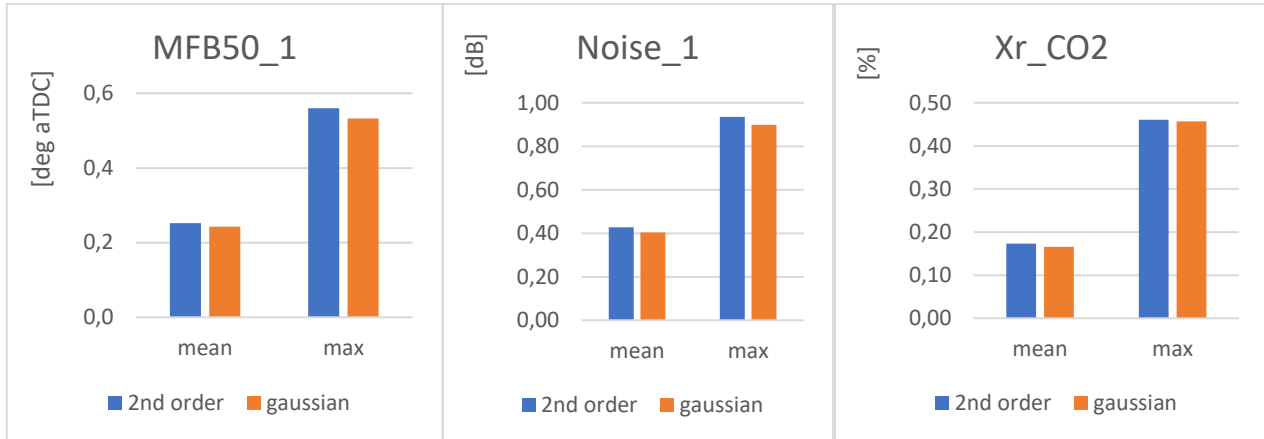


Figure 4.4 Mean and maximum error difference of MFB50, Noise of cylinder 1 and EGR percentage

The first impression is the irrelevant difference between the two model typologies also in validation. Seeing the first image, the mean errors of HC models are about 8% with a maximum value of 30%. These are alarming values and so these models cannot be defined reliable.

Instead the other models have a mean error value under the 4%. The CO models measure the maximum error value(excluding HC models). However it is about 15% and can be consider a passable prediction.

From right to the left in the figure 4.4:

- MFB50\_1(related to 1<sup>st</sup> cylinder) model have a mean error of 0.2 degrees and a maximum error of less than 0.6 degrees. These are negligible numbers for this type of measurement and so the model is satisfactory.
- Noise\_1(related to 1<sup>st</sup> cylinder) model have a mean error of 0.4 dB and a maximum error of 0.9 dB. The measured value during the validation tests goes from 82.8 dB to 87.2 dB. So the errors are negligible respect the measured values.
- Xr\_CO2(EGR percentage measured by gas sensors) model have a mean error of 0.19% and a maximum error of 0.45%. The measured value during the validation tests goes from 59,5% to 64%. Also in this case the errors are negligible respect the measured value.

In conclusion the created models have a good predictivity power except for the HC model.

A possible reason can be the extremely high amount of exhaust gas recirculated: it is possible that in some validation tests this high EGR value may have created an high unbalancing of exhaust gases between the cylinders. This could have led to an extremely high HC production and since this phenomenon cannot be expected by this model there is a large discrepancy between the predicted value and the measured value.

Further investigation has been done on these models: the optimal point for each models has been searched using cage (calibration generation, a MBC tool). The following table points out the results:

Optimization on:		constraints	model	Value	SOI[°]	FLAP[%]	VGT[%]	Prail[bar]
1	b_e[g/kWh]		2nd order	409.22	19.56	73.24	45.02	500.04
			gaussian	409.28	19.63	73.44	45.02	500.04
2	Nox_bs_exh [g/kWh]		2nd order	0.26	17.97	80.57	25.00	588.90
			gaussian	0.26	17.97	80.57	25.00	568.14
3	HC_bs_exh [g/kWh]		2nd order	5.19	20.43	73.24	37.64	499.97
			gaussian	5.12	20.70	73.24	38.28	499.96
4	CO_bs_exh [g/kWh]		2nd order	37.75	17.97	73.24	40.08	499.97
			gaussian	37.78	17.97	73.24	40.49	499.97
5	b_e [g/kWh]	Nox_bs_exh<0.6g/kWh	2nd order	411.85	20.45	77.86	45.01	500.01
			gaussian	412.01	19.88	77.83	45.01	500.01
6	b_e[g/kWh]	HC_bs_exh<7g/kWh	2nd order	411.85	20.42	77.85	45.01	500.01
		Nox_bs_exh<0.6g/kWh	gaussian	412.01	19.88	77.83	45.01	500.01
7	b_e[g/kWh]	HC_bs_exh<7 g/kWh	2nd order	411.85	20.42	77.85	45.01	500.01
		CO_bs_exh<50 g/kWh	gaussian	412.01	19.88	77.83	45.01	500.01

Table 4.4 optimal calibration(1800rpmx27Nm)

Also in calibration stage the models give practically the same results. The only careless difference is on the 2<sup>nd</sup> optimization: the models suggest a rail pressure that differs of about 20 bar. However this difference considering the range(from 500 bar to 700bar) is absolutely negligible and moreover the NOx emission is almost constant changing the rail pressure in this condition(PCCI combustion). For this DoE it was not possible built the Soot model. This because the FSN measurements are too low and too influenced by measurement uncertainty. These measurements are displayed in the figure XXX

#### 4.2. 5 operating points-conventional diesel combustion (3L engine)

The previous described engine is adopted for further tests: 5 operating points are selected and a V-optimal DoE of 25 tests has been created for each of them. The engine is used in normal combustion mode and the only 2 test variables are: rail pressure and start of injection.

The operation points are listed in the following table:

	Engine speed[rpm]	Torque[Nm]
1 <sup>st</sup> point	3250	47
2 <sup>nd</sup> point	2500	161
3 <sup>rd</sup> point	1800	161
4 <sup>th</sup> point	1200	260
5 <sup>th</sup> point	1200	23

Table 4.5 The tested operating points for DoE with 2 variables(3L engine)

Consumption, emissions and noise models are created for each operation point. The results are similar for each operating point from model quality point of view. For simplicity only the 1<sup>st</sup> point results are showed in the following table:

	mean value (25 tests)	R <sup>2</sup>		RMSE		PRESS RMSE	
		2nd order	gaussian	2nd order	gaussian	2nd order	gaussian
b <sub>e</sub> [g/kWh]	352.00	0.984	0.992	0.408	0.259	0.481	0.435
Nox <sub>bs</sub> [g/kWh]	8.32	0.999	1	0.109	0.033	0.13	0.078
Soot <sub>bs</sub> [g/kWh]	0.21	0.981	0.999	0.017	0.004	0.019	0.009
CO <sub>bs</sub> [g/kWh]	4.00	0.97	0.998	0.092	0.022	0.103	0.051
HC <sub>bs</sub> [g/kWh]	1.10	0.888	0.996	0.054	0.01	0.061	0.033
Noise <sub>1</sub> [dB]	88.29	0.93	0.999	0.481	0.05	0.578	0.135

*Table 4.5 statistic indicators about 3250rpmx47Nm point*

The results are very satisfying: the 2<sup>nd</sup> order model and the gaussian one produce similar results also in this case(with only 2 inputs). The fitting is excellent for always models. From the surface response plots there are not overfitting issues.

For these DoE there are not validation tests.

#### 4.3. Multiple injections DoE (2L engine)

The following models are created considering V-optimal DoE related to a 2L engine. Its characteristics are listed in the table 4.6

Engine type	2.0L Euro 5
Displacement	1956 cm <sup>3</sup>
Bore × stroke	83.0 mm × 90.4 mm
Compression ratio	16.3
Valves per cylinder	4
Turbocharger	Twin-stage with valve actuators and WG
Fuel injection system	CR 2000 bar piezoelectric indirect acting injectors
Specific power and torque	71 kW/l–205 N m/l
EGR system type	Short-route cooled EGR

*Table 4.6 2L engine specifications*

The following DoE are more complex than the previous ones. The following parameters were considered as the most relevant input variables for the procedure: rail pressure, swirl actuator position, dwell times (DT) between consecutive injections, main injection timing , the injection quantities in each shot, the boost pressure(not for each DoE) and the inducted air per stroke and per cylinder.

The next table summarizes the tests:

Injection strategy	Operating point [rpm x bar]
pM(pilot+main)	1500x2, 2000x5, 2750x12
ppM(double pilot+main)	1500x2, 1500x5, 2000x2, 2000x5
pMa(pilot+main+after)	1500x5, 2000x5, 2500x8, 2750x12
ppMa(double pilot+main+after)	1500x5

*Table 4.6 list of injection strategies*

For each DoE a 2<sup>nd</sup> order model and gaussian one have been created for the following outputs:

bsfc[g/kWh]	NOx[g/kWh]	CO[g/kWh]	HC[g/kWh]	soot[g/kWh]	s415[FSN]	Noise2[dBA]	lambda[-]
-------------	------------	-----------	-----------	-------------	-----------	-------------	-----------

The fuel injection strategy can play an important role in simultaneously reducing passenger car diesel engine emissions and combustion noise, without neglecting fuel consumption targets. A multiple injection strategy, adopted to replace a single fuel injection shot with multiple discrete fuel injection events of reduced size, can easily be implemented using Common Rail systems, equipped with the modern injectors.

The implementation of a pilot injection in diesel engines makes the entire amount of the fuel chemical energy be released over a prolonged time interval, thus determining a longer combustion than for the single injection case. Furthermore, the premixed combustion of some of the pilot injected fuel causes a slight increase in the in-cylinder gas pressure and temperature before the main injection has occurred, and therefore leads to a considerable reduction in the ignition delay of the main injection. This reduction in the fuel ignition delay limits the impact of the premixed combustion and generates a less rapid heat release rate during the main injection than in a single injection schedule. As a consequence, the main combustion becomes predominantly mixing-controlled. A pilot injection that is sufficiently close to the main injection has the potential of enhancing combustion efficiency and thus brake specific fuel consumption (bsfc), because the pilot and main combustions are linked smoothly. For the same reason, pilot injections are also effective in decreasing combustion noise (CN), especially at engine idle. Reductions of up to 5–8 dB are generally obtained in the CN, compared to single injection strategies. Since the pilot injection decreases the impact of the overall premixed combustion, it makes the highest flame temperatures diminish and, as a consequence, the NO<sub>x</sub> emissions generally also reduce, compared to the single-injection strategy. However, large pilot injected quantities make the NO<sub>x</sub> produced in the pilot combustion grow, and the increase in the NO<sub>x</sub> amount produced by the pilot combustion can surpass the decrease in the main combustion NO<sub>x</sub> emissions, due to the shortened ignition delay. The smoke emission in pilot–main injections generally tends to increase, compared to single injections. In fact, the pilot injection increases the in-cylinder temperature and decreases the oxygen concentration in the gases before the main injection has occurred. Both of these effects generally make the smoke emission grow [15].

After-injection is efficient in reducing the soot engine-out emissions, which can be up to 40% lower than in the single injection case. In general, after-injections can oxidize part of the unburned fuel and a decrease in CO, HC and PM engine-out emissions is obtained. The benefits increase when mixing

is difficult, i.e. at medium to high loads and under high EGR conditions when the utilization of in-cylinder air is critical. After-injections have also been proposed as a means of reducing turbocharger lag during engine transients, since the pressure and the temperature of the exhaust gas leaving the cylinder can be raised significantly. This makes the turbocharger accelerate more quickly and allows the aggressive increase in the injection quantity to be started sooner, since the fuel quantity growth must follow the increase in the air supply in order to avoid high soot during transients. Finally, after-injections can be applied to raise the diesel oxygen catalyst (DOC) temperature above its light-off temperature after a cold start and post injections can then be supplied to raise the exhaust temperature even further [16].

To simplify the discussion are evaluated only the models that have also validation data. The other model results are in the appendix.

#### 4.3.1. 1500rpmx2bar pM

From fitting point of view, all the models have  $R^2$  higher than 0.8: the only exception is the 2<sup>nd</sup> order model of HC ( $R^2=0.72$ ).

Seeing the table 4.7 and 4.8 the models are validated, almost all with good precision. However the HC, soot and s415 models are poor.

	Model	bsfc[g/kWh]	NOx[g/kWh]	CO[g/kWh]	HC[g/kWh]	soot[g/kWh]	s415[FSN]	Noise2[dBA]	lambda[-]
validation RMSE	2nd order	11.34	0.07	1.47	0.84	0.02	0.14	0.95	0.05
validation RMSE	gaussian	11.66	0.06	1.51	0.84	0.01	0.12	0.84	0.05
range		285.3-335.4	0.37-1.27	4.60-21.90	1.35-8.47	0.003-0.09	0.04-0.79	70.47-82.08	1.71-2.17

*Table 4.7 statistic indicators about 1500rpmx2bar pM point*

	Model	bsfc[g/kWh]	NOx[g/kWh]	CO[g/kWh]	HC[g/kWh]	soot[g/kWh]	s415[FSN]	Noise2[dBA]	lambda[-]
validation RMSE/mean value	2nd order	3.4%	11.2%	12.8%	31.0%	54.9%	45.4%	1.2%	2.5%
validation RMSE/mean value	gaussian	3.5%	9.5%	13.2%	30.9%	45.2%	38.7%	1.1%	2.4%

*Table 4.8 ratio between validation RMSE and mean value about 1500rpmx2bar pM point*

In the image xxx the residual of validated data for s415(2<sup>nd</sup> order model) are plotted. Apparently the errors are very small, but they are important in proportion to the measured values. Indeed the ratio of the validation RMSE and the mean value is very high. This means that there are points with worst error than 45%. The same conclusions can be done for HC and soot models.

Instead the other models are excellent in validation.

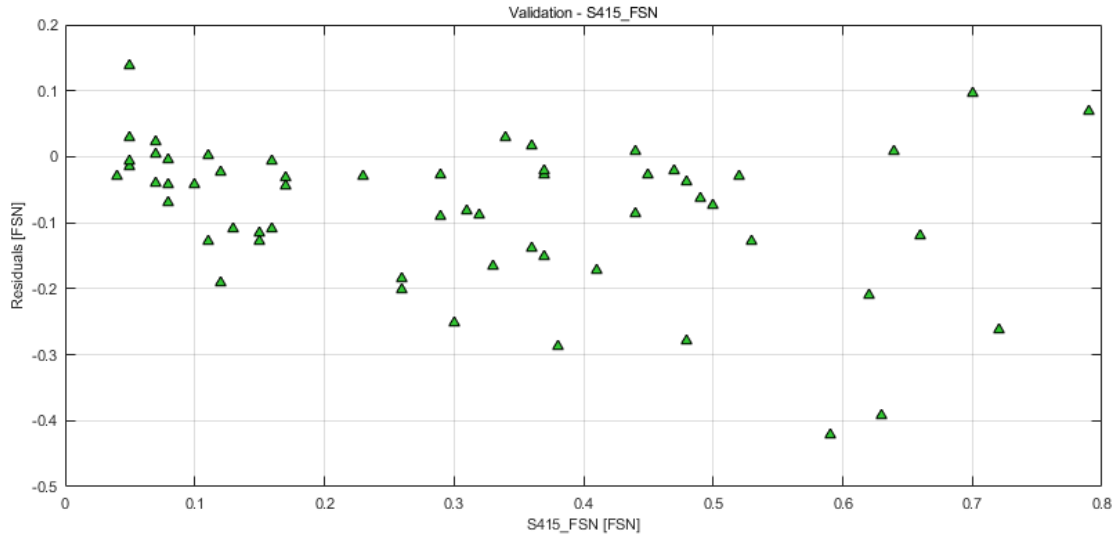


Figure 4.5 Validation data: measured FSN versus residuals (1500rpmx2bar ppm)

#### 4.3.2. 2000rpmx5bar ppm

From fitting point of view, all the models have  $R^2$  higher than 0.8: the only exception is the 2<sup>nd</sup> order model of HC ( $R^2=0.73$ ).

However in some gaussian models, overfitting is present. It is notable by response surface plot. For example the following image points out the NOx emission models ( 2<sup>nd</sup> order and gaussian)

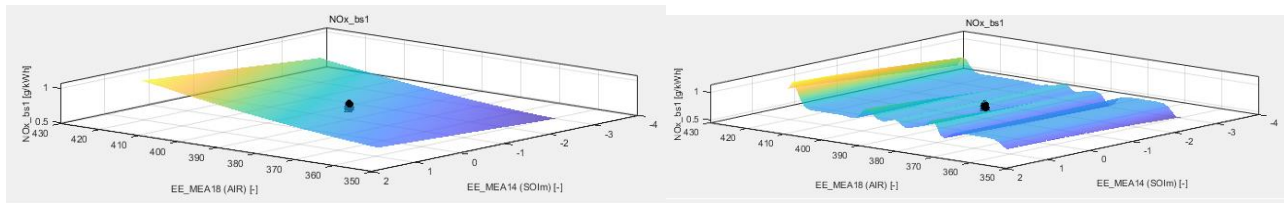


Figure 4.6 bsNOx surface. On the left the predicted surface by 2nd order model. On the right the one predicted by gaussian model

The gaussian surface appears very distorted by the fitting. This is one of the main drawbacks of gaussian process. This problem is present also for bsfc, HC, CO, s415, noise2, lambda and soot.

	Modello	bsfc[g/kWh]	NOx[g/kWh]	CO[g/kWh]	soot[g/kWh]	s415[FSN]	Noise2[dBA]	lambda[-]
validation RMSE	2nd order	1.7	0.05	0.34	0.08	0.14	0.5	0.01
validation RMSE	gaussian	1.5	0.05	0.33	0.06	0.12	0.4	0.01
range		234.6-263.6	0.40-1.19	1.89-6.39	0.08-1.93	0.95-5.21	80.5-86.7	1.36-1.92

Table 4.9 statistic indicators about 2000rpmx5bar ppm point

	Model	bsfc[g/kWh]	NOx[g/kWh]	CO[g/kWh]	soot[g/kWh]	s415[FSN]	Noise2[dBA]	lambda[-]
validation RMSE/mean value	2nd order	0.7%	6.8%	11.3%	18.9%	5.7%	0.6%	0.5%
validation RMSE/mean value	gaussian	0.6%	7.6%	11.0%	14.5%	4.7%	0.4%	0.6%

*Table 4.10 ratio between validation RMSE and mean value about 2000rpmx5bar ppm point*

The predictivity power is remarkable for these models. The only model, on which to do further analysis, is the soot model. However its validation RMSE value is enough low on the mean measured value.

#### 4.4. Active DoE-AVL CAMEO simulator

The Cameo heavy duty engine simulator has been used to test the potentiality of the active DoE. To point out it, two examples are showed.

Two different variation lists have been constructed for two examples with the same point numbers:

- 40 points created by D-optimal design
- 29 points created by D-optimal design plus 11 additional points computed by Active DoE procedure.

The free variables are: SOI, rail pressure, boost pressure, pilot injection quantity, pilot injection timing, air mass for cylinder.

A quadratic polynomial model has been computed for each engine output: brake specific fuel consumption, bsNOx, bsHC, bsCO, combustion noise.

The first example regards the operating point 2200rpm x 375Nm. The strategy of active procedure is to minimize the brake specific fuel consumption. The aim is to add N additional point (in this case 11) in order to find the minimum value.

From the fitting point of view, the difference are negligible between the two model.:

	Conventional DoE	Active DoE
$R^2$	0.991	0.963
$R^2_{adj}$	0.982	0.94
$PRESS R^2$	0.965	0.881

*Table 4.11 Statistic indicators between conventional and active D.o.E. (2200rpm x 375Nm)*

Using CaGe, the minimum point has been found about bsfc (optimal point) for the two models:

bsfc [g/kWh]	NOx [g/kWh]	CO [g/kWh]	HC [g/kWh]	Noise [dB]
199.7	8.7	0.09	0.14	100.9

*Table 4.12 minimum bsfc (conventional procedure 2200rpm x 375 Nm)*

bsfc [g/kWh]	NOx [g/kWh]	CO [g/kWh]	HC [g/kWh]	Noise [dB]
181.1	16.3	1.23	0.18	99.8

*Table 4.13 minimum bsfc (active procedure 2200rpm x 375 Nm)*

The bsfc minimization is about 18 g/kWh. However there is a huge increment about NOx: for a HD engine this means an high increment of urea consumption by SCR.

So a second optimization has been done for the active DoE model, with a constraint on NOx (it must be equal or less than 10g/kWh). The bsfc is anyway less the first minimization and now the increment NOx is moderate:

bsfc [g/kWh]	NOx [g/kWh]	CO [g/kWh]	HC [g/kWh]	Noise [dB]
188.1	10.0	0.67	0.17	100.0

*Table 4.14 minimum bsfc with constraint on NOx (active procedure 2200rpm x 375 Nm)*

The operating points is 2600rpm x 425 Nm for the second example.

The active strategy is to increase the fitting quality for the combustion noise model (called standard strategy). The following table shows that with the same number of measurements (40), the Active DoE method increases the fitting quality.

	Conventional D.o.E.	Active D.o.E.
$R^2$	0.908	0.92
$R^2_{adj}$	0.844	0.893
$PRESS R^2$	0.725	0.86

*Table 4.15 Statistic indicators between conventional and active D.o.E. (2600rpm x 425Nm)*



## 5. Conclusions

This aim of this thesis is to investigate on the model-based calibration for engines. This approach is mandatory in this period where the number of engine and aftertreatment actuators is very high and testing all the engine conditions is necessary at the same time.

The first part focuses on the Design of Experiment choice: different methods are compared from application and mathematical point of view. From comparisons, the optimal design is the best choice to reduce the number of experiments maintaining an high quality modelling. Moreover innovative solutions are researched: Active DoE is a method to perform an online modelling to further increase the efficiency (highest model quality with low measurements).

In the second step, different modelling techniques are compared on experimental data.

A new outlier detection method and a refined fitting procedure are tested with excellent result. Then gaussian and polynomial models have been created and compared on different outputs, inputs, engines and working conditions.

The results highlight that the differences are almost always negligible both for fitting and predictivity properties. However gaussian ones have always  $R^2$  greater than polynomial second order ones but in same cases the GP models suffer from overfitting.

The conclusion is that these two regression processes are comparable and almost always equivalent for these applications.

A possible investigation for future activities can be to study the global modelling techniques and compare polynomial and GP processes on these problems, with engine load and speed as input variables. Another suggestion can be to test an Active DoE on test bed.

## 6. Acknowledgements

AVL is acknowledged for offering the utilization of CAMEO 4 within the frame of the University Partnership Program.

# Appendix

		bsfc[g/kWh]	NOx[g/kWh]	CO[g/kWh]	HC[g/kWh]	soot[g/kWh]	s415[FSN]	Noise2[dBA]	lambda[-]
2nd order	R <sup>2</sup>	0,81	0,86	0,861	0,766	0,825	0,852	0,91	0,975
	RMSE	4,457	0,049	1,47	0,848	0,014	0,107	0,906	0,02
	PRESS RMSE	4,745	0,052	1,584	0,984	0,015	0,115	1,013	0,047
gaussian	R <sup>2</sup>	0,81	0,972	0,97	0,817	0,947	0,957	0,964	0,976
	RMSE	4,184	0,038	0,635	0,624	0,00727	0,055	0,528	0,018
	PRESS RMSE	4,843	0,051	1,359	0,846	0,013	0,1	0,931	0,045

Table A.1 1500rpmx2bar pM models: statistical indicators

		bsfc[g/kWh]	NOx[g/kWh]	CO[g/kWh]	HC[g/kWh]	soot[g/kWh]	s415[FSN]	Noise2[dBA]	lambda[-]
2nd order	R <sup>2</sup>	0,827	0,92	0,841	0,829	0,912	0,979	0,962	0,994
	RMSE	3,347	0,034	0,812	0,052	0,054	0,128	0,358	0,0043
	PRESS RMSE	3,758	0,039	0,868	0,06	0,065	0,152	0,388	0,0048
gaussian	R <sup>2</sup>	0,991	0,916	0,842	0,998	0,972	0,979	0,967	0,993
	RMSE	0,657	0,032	0,71	0,0044	0,045	0,113	0,303	0,0042
	PRESS RMSE	4,935	0,039	0,879	0,093	0,069	0,167	0,383	0,0051

Table A.2 2000rpmx5bar pM models: statistical indicators

		bsfc[g/kWh]	NOx[g/kWh]	CO[g/kWh]	soot[g/kWh]	s415[FSN]	Noise2[dBA]	lambda[-]
2nd order	R <sup>2</sup>	0,878	0,942	0,883	0,837	0,919	0,984	0,862
	RMSE	1,554	0,076	0,229	0,055	0,223	0,128	0,018
	PRESS RMSE	1,614	0,081	0,244	0,059	0,24	0,139	0,019
gaussian	R <sup>2</sup>	0,878	0,94	0,973	0,96	0,985	0,997	0,975
	RMSE	1,492	0,072	0,079	0,026	0,089	0,049	0,0073
	PRESS RMSE	1,642	0,082	0,138	0,043	0,142	0,087	0,011

Table A.3 2750rpmx12bar pM models: statistical indicators

		bsfc[g/kWh]	CO[g/kWh]	HC[g/kWh]	soot[g/kWh]	s415[FSN]	Noise2[dBA]	lambda[-]
2nd order	R <sup>2</sup>	0,816	0,811	0,773	0,805	0,884	0,888	0,951
	RMSE	5,958	1,707	0,509	0,089	0,285	1,241	0,023
	PRESS RMSE	6,601	1,964	0,576	0,101	0,323	1,388	0,025
gaussian	R <sup>2</sup>	0,911	0,942	0,967	0,995	0,994	0,999	0,985
	RMSE	3,801	0,856	0,175	0,012	0,058	0,079	0,012
	PRESS RMSE	7,098	2,184	0,445	0,109	0,333	1,674	0,028

Table A.4 1500rpmx2bar ppM models: statistical indicators

		bsfc[g/kWh]	NOx[g/kWh]	CO[g/kWh]	HC[g/kWh]	soot[g/kWh]	s415[FSN]	Noise2[dBA]	lambda[-]
2nd order	R^2	0,947	0,958	0,866	0,861	0,909	0,964	0,954	0,993
	RMSE	2,053	0,061	0,608	0,024	0,097	0,222	0,358	0,012
	PRESS RMSE	2,399	0,065	0,699	0,029	0,111	0,245	0,407	0,013
gaussian	R^2	0,99	0,96	0,997	0,99	1	0,994	0,987	0,994
	RMSE	0,805	0,055	0,08	0,0057	0,00647	0,085	0,174	0,00962
	PRESS RMSE	3,018	0,071	0,542	0,028	0,161	0,326	0,394	0,015

Table A.5 1500rpmx5bar pMa models: statistical indicators

		bsfc[g/kWh]	NOx[g/kWh]	CO[g/kWh]	HC[g/kWh]	soot[g/kWh]	s415[FSN]	Noise2[dBA]	lambda[-]
2nd order	R^2	0,963	0,988	0,895	0,849	0,928	0,981	0,902	0,995
	RMSE	1,68	0,042	0,744	0,026	0,135	0,206	0,514	0,011
	PRESS RMSE	1,919	0,047	0,816	0,029	0,155	0,228	0,561	0,012
gaussian	R^2	0,954	0,985	0,972	0,964	0,975	0,978	0,903	0,994
	RMSE	1,674	0,041	0,357	0,012	0,076	0,197	0,473	0,01
	PRESS RMSE	2,23	0,056	0,554	0,023	0,141	0,29	0,609	0,013

Table A.6 1500rpmx5bar ppM models: statistical indicators

		bsfc[g/kWh]	NOx[g/kWh]	CO[g/kWh]	HC[g/kWh]	soot[g/kWh]	s415[FSN]	Noise2[dBA]	lambda[-]
2nd order	R^2	0,965	0,975	0,908	0,86	0,904	0,97	0,922	0,995
	RMSE	1,925	0,061	0,748	0,075	0,156	0,248	0,464	0,01
	PRESS RMSE	2,125	0,068	0,828	0,078	0,173	0,277	0,526	0,012
gaussian	R^2	0,979	0,992	0,995	0,991	0,989	0,986	0,995	0,996
	RMSE	1,355	0,031	0,156	0,019	0,05	0,156	0,112	0,0086
	PRESS RMSE	2,214	0,06	0,662	0,067	0,15	0,268	0,364	0,012

Table A.7 1500rpmx5bar ppMa models: statistical indicators

		bsfc[g/kWh]	NOx[g/kWh]	CO[g/kWh]	HC[g/kWh]	soot[g/kWh]	s415[FSN]	Noise2[dBA]	lambda[-]
2nd order	R^2	0,856	0,92	0,841	0,779	0,823	0,89	0,888	0,977
	RMSE	6,926	0,079	2,052	0,896	0,082	0,258	1,163	0,025
	PRESS RMSE	8,039	0,088	2,385	1,001	0,097	0,287	1,384	0,029
gaussian	R^2	0,979	0,933	0,946	0,996	0,999	0,995	1	0,982
	RMSE	2,358	0,064	1,048	0,095	0,006	0,049	0,051	0,02
	PRESS RMSE	9,083	0,1	2,569	1,081	0,1	0,313	1,487	0,033

Table A.8 2000rpmx2bar ppM models: statistical indicators

		bsfc[g/kWh]	NOx[g/kWh]	CO[g/kWh]	HC[g/kWh]	soot[g/kWh]	s415[FSN]	Noise2[dBA]	lambda[-]
2nd order	R <sup>2</sup>	0,971	0,978	0,937	0,749	0,951	0,974	0,96	0,995
	RMSE	1,75	0,029	0,296	0,054	0,053	0,162	0,341	0,0081
	PRESS RMSE	2,033	0,034	0,356	0,059	0,06	0,186	0,373	0,0091
gaussian	R <sup>2</sup>	0,994	0,998	0,993	0,943	0,994	0,997	0,992	0,999
	RMSE	0,691	0,0079	0,086	0,024	0,016	0,048	0,138	0,0036
	PRESS RMSE	2,166	0,04	0,459	0,05	0,05	0,173	0,39	0,01

Table A.9 2000rpmx5bar pMa models: statistical indicators

		bsfc[g/kWh]	NOx[g/kWh]	CO[g/kWh]	HC[g/kWh]	soot[g/kWh]	s415[FSN]	Noise2[dBA]	lambda[-]
2nd order	R <sup>2</sup>	0,951	0,97	0,938	0,722	0,969	0,986	0,902	0,997
	RMSE	1,544	0,035	0,218	0,018	0,065	0,132	0,491	0,0057
	PRESS RMSE	1,705	0,038	0,24	0,019	0,077	0,153	0,571	0,0064
gaussian	R <sup>2</sup>	0,971	0,979	0,955	0,695	0,978	0,994	0,996	0,995
	RMSE	1,092	0,027	0,166	0,017	0,049	0,075	0,09	0,0069
	PRESS RMSE	1,863	0,041	0,284	0,022	0,088	0,183	0,664	0,0088

Table A.10 2000rpmx5bar ppM models: statistical indicators

		bsfc[g/kWh]	NOx[g/kWh]	CO[g/kWh]	soot[g/kWh]	s415[FSN]	Noise2[dBA]	lambda[-]
2nd order	R <sup>2</sup>	0,974	0,984	0,842	0,894	0,976	0,99	0,993
	RMSE	1,415	0,068	0,319	0,138	0,197	0,155	0,0097
	PRESS RMSE	1,576	0,074	0,36	0,157	0,232	0,178	0,01
gaussian	R <sup>2</sup>	0,988	0,99	0,991	1	1	0,994	0,997
	RMSE	0,872	0,049	0,071	0,0038	0,022	0,104	0,0062
	PRESS RMSE	1,95	0,08	0,423	0,171	0,269	0,243	0,016

Table A.11 2500rpmx8bar pMa models: statistical indicators

		bsfc[g/kWh]	NOx[g/kWh]	CO[g/kWh]	HC[g/kWh]	soot[g/kWh]	s415[FSN]	Noise2[dBA]	lambda[-]
2nd order	R <sup>2</sup>	0,935	0,991	0,864	0,715	0,923	0,975	0,878	0,964
	RMSE	2,597	0,046	1,025	0,015	0,239	0,21	0,454	0,013
	PRESS RMSE	2,817	0,053	1,163	0,017	0,282	0,239	0,493	0,014
gaussian	R <sup>2</sup>	0,971	0,994	0,997	0,849	1	0,993	0,879	0,97
	RMSE	1,591	0,032	0,127	0,0098	0,014	0,1	0,415	0,011
	PRESS RMSE	3,55	0,076	1,065	0,016	0,288	0,383	0,529	0,014

Table A.12 2750rpmx12bar pMa models: statistical indicators

## Bibliography

- [1] G. Vitale et al., *DoE and beyond: the evolution of the model-based development approach how legal trends are changing methodology*, 6. Internationales Symposium fur Entwicklungsmethodik, Wiesbaden, 2015.
- [2] D. Montgomery, *Design and Analysis of Experiments, Introduction*, Eighth edition.
- [3] Matlab documentation, *Create a space filling design*.
- [4] M. Koegeler, *Doe Principles: optimization of injection and combustion*, AVL, 2013.
- [5] P.F. de Aguiar, B. Bourguignon, M.S. Khots, D.L. Massart, R. Phan-Than-Luu, *D-optimal designs*, 1994
- [6] AVL cameo 4, *User's guide advanced*, 2020
- [7] AVL cameo 3, *Solution sheet: Active DoE*, 2017
- [8] M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006
- [9] B. Berger, *Modeling and Optimization for Stationary Base Engine Calibration*, 2012
- [10] Montgomery, Peck, Vining, *Introduction to linear regression analysis*, fifth edition, Wiley, 2012
- [11] Montgomery, Douglas, *Statistica per l'ingegneria*, 2004
- [12] Matlab documentation, *Stepwise regression*
- [13] Matlab documentation, *Prediction error variance view*
- [14] D'Ambrosio, Mancarella, Iemmolo, Vitolo, *Preliminary optimization of the PCCI combustion mode in a diesel engine through a design of experiments*, 71st Conference of the Italian Thermal Machines Engineering Association, 2016, Turin
- [15] D'Ambrosio, Ferrari, *Potential of double pilot injection strategies optimized with the design of experiments procedure to improve diesel engine emissions and performance*, 2015
- [16] D'Ambrosio, Ferrari, *Potential of multiple injection strategies implementing the after shot and optimized with the design of experiments procedure to improve diesel engine emissions and performance*, 2015