



HAL
open science

Modeling Pollutant Emissions of Diesel Engine based on Kriging Models: a Comparison between Geostatistic and Gaussian Process Approach

Sébastien Castric, Lilianne Denis-Vidal, Zohra Cherfi, Ghislaine Joly-Blanchard, Nassim Boudaoud

► To cite this version:

Sébastien Castric, Lilianne Denis-Vidal, Zohra Cherfi, Ghislaine Joly-Blanchard, Nassim Boudaoud. Modeling Pollutant Emissions of Diesel Engine based on Kriging Models: a Comparison between Geostatistic and Gaussian Process Approach. 2012. hal-00699171

HAL Id: hal-00699171

<https://hal.science/hal-00699171>

Preprint submitted on 19 May 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Modeling Pollutant Emissions of Diesel Engine based on Kriging Models: a Comparison between Geostatistic and Gaussian Process Approach

S. Castric* L. Denis-Vidal** Z. Cherfi* G. Joly Blanchard**
N. Boudaoud*

* are with Roberval laboratory UMR- CNRS 6253, University of Technology of Compiègne, France (e-mail: sebastien.castric@utc.fr, zohra.cherfi@utc.fr, nassim.boudaoud@utc.fr).

** are with Applied Mathematical Laboratory of Compiègne, University of Technology of Compiègne, France (e-mail: Lilianne.denis-vidal@math.univ-lille1.fr, ghislaine.joly-blanchard@utc.fr)

Abstract: In order to optimize the performance of a diesel engine subject to legislative constraints on pollutant emissions, it is necessary to improve their design, and to identify how design parameters affect engine behaviours. One specificity of this work is that it does not exist a physical model of engine behaviour under all possible operational conditions. A powerful strategy for engine modeling is to build a fast emulator based on carefully chosen observations, made according to an experimental design. In this paper, two Kriging models are considered. One is based on a geostatistical approach and the other corresponds to a Gaussian process metamodel approach. Our aim is to show that the use of each of these methods does not lead to the same results, particularly when "atypical" points are present in our database. In a more precise way, the statistical approach allows us to obtain a good quality modeling even if atypical data are present, while this situation leads to a bad quality of the modeling by the geostatistical approach. This behaviour takes a fundamental importance for the problem of the pollutant emissions, because the analysis of these atypical data, which are rarely erroneous data, can supply precious information for the engine tuning in the design stage.

Keywords: Modelling, Optimization problems, modelling errors, Automotive emissions, Diesel engines

1. INTRODUCTION

The automotive industry faces the competing goals of producing better performing vehicles and keeping development time with low costs. It is crucial for the manufacturers to be able to produce fuel-economic vehicles, which respect pollutant emissions standards, and which meet the customer expectations. Accordingly, the complexity of the engines responses to be optimized and the number of the parameters to be controlled during the design stage, have increased rapidly, in the last years. In order to deliver vehicles, which respond to these requirements, in a reasonable time scale, companies use design of experiments (DOE) in one side, and modeling, in the other side. DOE [Box et al. (2005)] is a powerful tool, but the cost of the experiments and their duration, particularly in the field of pollutant emissions, can be a limit to their use in automotive industry. The engine developers use two main approaches to model engine behaviour. The first one is based on chemical and physical models, via differential systems, see Borg Jonathan et al. (2009) et Brahma et al. (2010). This approach is not the subject of this article, because in our case, we don't have such models. Further-

more, even when these models are available, generally, they are time-consuming, impractical for multi-objective optimisation routines, and fail to capture all the trends in the engine system described by measured data. Statistical modeling based on carefully chosen measured data of engine performance, according to an experimental design, is an important alternative technique. Strategies based on Lolimot [Castric et al. (2007)] (Local Linear Model Tree) and Zeldovich mechanisms Heywood (1988) have been developed in order to predict emissions of NOx. In the first case, the corresponding model can lead to singular points, which reduces the precision of the results. In the second case, the results are not satisfactory enough. Therefore, several statistical modeling methods offer interesting possibilities, see Edwards et al. (1997). The methods based on statistical training such as the neural networks Fang et al. (2006), knew a striking success. At the same time, the use of the computer modeling [Santner et al. (2003) and Bayarri et al. (2009)] is strongly developed. The methods based on the surface responses methodologies are studied and compared, particularly by Jin et al. (2001), taking into account various levels of complexity of the problem in term of nonlinearity and of the number of variables. However,

most of these approaches cannot be used in our case, because we are drastically limited by the small number of experiments which our industrial partner is able to do. The advantages offered by the kriging methods [Matheron (1963), Zhu et al. (2010) and Jack (2009)] brought us to choose them for our study. Two approaches of Kriging model are considered in this work. The first one is based on a geostatistical approach. Softwares such as R and Matlab contain a toolbox for the use of this method, but unfortunately, they are restricted to less than 3 dimensions. The method has been adapted to higher dimensions, and more generally an innovative approach for functional data has been proposed. The performance of this model is due to its ability to take into account the spatial dependence of data [Cressie (1991)], and has minimal variance estimators without bias. But this method may be difficult to couple with optimization process of calibration. The second one is a Gaussian process metamodel approach Baillargeon et al. (2004). In this case, the system response in consideration is modeled by the sum of two quantities $f(x)$ and $\delta(x)$, where $f(x)$ corresponds to a deterministic function which is a linear combination of known functions (in our case a one-degree polynomial model) and $\delta(x)$ is the stochastic part and corresponds to a Gaussian stationary process of order 2, see Iooss (2009). The aim of this paper is to bring response to the problem of the polluting emissions. At first, we wish to know if the two kriging models are adapted to the modeling of the three pollutants studied here: NOx, smokes and CO. Then, the question which arises is to know if, according to the values taken by the coefficient used to estimate the quality of the modeling, conclusions can be established as for the presence or not of atypical points. The detection of these atypical points and their interpretation are very important, because these points are often revealing specific information for engine tuning. This paper is organized as follows: In the second section, the Kriging techniques, geostatistical approach and Gaussian process metamodel, are recalled. In the third section, the engine behaviour and the importance of controlling pollutant emissions are described. In the last section, numerical results and some elements of discussion are given.

2. KRIGING TECHNIQUES

Kriging methods are used frequently for spatial interpolation of soil properties, see Krige (1951). Kriging is a linear least squares estimation algorithm, which is used for interpolation. The aim is to estimate the value of an unknown real function y at point x^* , given the values of the function y at some other points $x^{(i)} \in \mathbb{R}$ for each $i \in \{1, \dots, n\}$. The Kriging estimator is defined by

$$\hat{y}(x^*) = \sum_{i=1}^n \lambda_i y(x^{(i)}) \quad (1)$$

In this paper, we will note $y(x^{(i)}) = y^{(i)}$

Where

n is the number of surrounding observations

$x^{(i)}$ is a vecteur composed of the d values $(x_1^{(i)}, \dots, x_d^{(i)})$ of the factors at point i

λ_i is the weight corresponding to the observation $y(x^{(i)})$. The weights are estimated in order to make the estimator unbiased with minimal variance.

The system response is treated as a realization of a random function $y(x)$. This model can be written as:

$$y(x) = f(x) + \delta(x).$$

The deterministic function $f(x)$ provides the mean of $y(x)$ and $\delta(x)$ is the stochastic part which verifies some assumptions depending on the choice of Kriging model.

2.1 Geostatistical approach: Ordinary Kriging

In this approach the weights are determined such that the following Kriging variance

$$\text{var}(\hat{y}(x^*) - y(x^*))$$

is minimal under the unbiased constraint given by

$$\sum_{i=1}^n \lambda_i = 1. \quad (2)$$

It leads to a classical optimization problem with equality constraint. The Lagrange multiplier theory is used in order to work out this problem. This leads to a linear system to be solved, see Davis (1986). This system produces, under certain assumptions specified below, a function called variogram described in the following paragraph.

Variogram The variogram is a function representing the spatial dependency. It is obtained from the stationarity definition. Indeed, this stationarity hypothesis is an indispensable condition for the use of the Kriging method. In the case of ordinary Kriging, the expression of the variogram is obtained from the following definition of intrinsic stationarity:

$$\begin{aligned} E(y(x^{(i)} + h) - y(x^{(i)})) &= 0, \quad 1 \leq i \leq n \\ \text{var}(y(x^{(i)} + h) - y(x^{(i)})) &= 2\gamma(h), \quad 1 \leq i \leq n. \end{aligned}$$

More precisely, the expression of the theoretical variogram is deduced from the second condition of intrinsic stationarity. With the assumption of isotropy, the variation of a data set will be dependent only on distance $\|h\|$ between two locations. To infer the variogram from observed data, we will then use the common formula for the experimental variogram, see Cressie (1991):

$$\hat{\gamma}(r) = \frac{1}{|N(r)|} \sum_{N(r)} [y(x^{(i)}) - y(x^{(j)})]^2 \quad (3)$$

where

$$N(r) = \{(i, j), \|x^{(i)} - x^{(j)}\| = r\}.$$

The quantity $|N(r)|$ is the pair number of $N(r)$ and the function $\hat{\gamma}(r)$ is the experimental variogram.

Variogram Model The experimental variogram presented in equation (3) estimates the theoretical variogram, for only a finite number of distances. Moreover, it does not necessarily form a valid variogram. This means that maybe, it does not concern a negative conditionally function. Indeed, this condition is necessary to ensure the positivity of the variance of a sum of random variables, see Christakos (1984). The experimental variogram is then modelled by a function of negative conditional type and is defined for all distances. It is named the variographic model. This model must be selected among the various

forms of the variogram models, which exist in the literature and adjusted to the experimental variogram Arnaud et al. (2000). It means that the parameters of the model must be estimated. This adjustment is done with an estimation method such as the weighted least squares or maximum likelihood method. Once the variographic model is chosen, and its parameters estimated, the weights λ_i which appear in (1) are computed by solving the following system:

$$A\Lambda = B$$

with

$$A = \begin{bmatrix} \gamma(r_{11}) & \gamma(r_{12}) & \dots & \gamma(r_{1n}) & 1 \\ \gamma(r_{21}) & \gamma(r_{22}) & \dots & \gamma(r_{2n}) & 1 \\ \dots & \dots & \dots & \dots & \dots \\ \gamma(r_{n1}) & \gamma(r_{n2}) & \dots & \gamma(r_{nn}) & 1 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix}$$

$$\Lambda = [\lambda_1 \lambda_2 \dots \lambda_n \lambda]^t$$

$$B = [\gamma(r_{01}) \gamma(r_{02}) \dots \gamma(r_{0n}) 1]^t$$

and

$$r_{ij} = \|x^{(i)} - x^{(j)}\|, \quad r_{0j} = \|x^* - x^{(j)}\|$$

where λ is the Lagrange multiplier. The function $\gamma(r_{ij})$ is the variogram model used for adjusting the experimental variogram.

Kriging Emulator Validation The true test of the quality of the fitted emulator model is its ability to predict the response at untried factor values. In order to maximally exploit the data to aid model fitting, the emulators are validated using leave-one-out cross validation. This process involves taking the fitted model and re-fitting it to a subset of non used experimental data. More precisely, for an experiment with d design factors, the set of n experimental design points and corresponding responses, contain the information used to build the Kriging model. A cross validation involves predicting at each design point in turn when that point is left out of the predictor equations. Let $\hat{y}(x^{(i)})$ be the estimate of the $y(x^{(i)})$ based on all the design points except $x^{(i)}$. The prediction error is then calculated by a coefficient of determination R^2 used in multiple regression:

$$R^2 = 1 - \frac{\sum_{i=1}^n |y(x^{(i)}) - \hat{y}(x^{(i)})|^2}{\sum_{i=1}^n |y(x^{(i)}) - \bar{y}|^2} \quad (4)$$

2.2 Gaussian Process Metamodel

In this approach, the deterministic function $f(x)$ is a linear combination of known functions. In this study, a one-degree polynomial model is used and $f(x)$ can be rewritten as follows:

$$f_\beta(x) = \beta_0 + \sum_{i=1}^d \beta_i x_i$$

where $\beta = (\beta_0, \dots, \beta_d)^t$ is the regression parameter vector. The stochastic part $\delta(x)$ is a Gaussian, stationary of order two process, then:

$$E(\delta(x)) = 0$$

$$E(\delta(x), \delta(z)) = \sigma_\delta^2 R(x - z)$$

where σ_δ^2 denotes the variance of Y and R is the correlation function. Among correlation models given in the literature,

see Abr (1997), we choose to use the generalized exponential correlation function:

$$R_{\theta,q}(x - z) = \exp\left(-\sum_{k=1}^d \theta_k |x_k - z_k|^{q_k}\right)$$

or

$$R_{\theta,q}(x - z) = \prod_{k=1}^d \exp(-\theta_k |x_k - z_k|^{q_k})$$

where $\theta = (\theta_1, \dots, \theta_d)^t$ and $q = (q_1, \dots, q_d)^t$ are the correlation parameters (hyperparameters) with $\theta_k \geq 0$ and $0 \leq q_k \leq 2, \forall k \in \{1, \dots, d\}$. This choice is motivated by wide spectrum of shapes that such a function offers. If a new point $x^* = (x_1^*, \dots, x_d^*)$ is considered, the following unbiased linear predictor is obtained:

$$\hat{y} = f_{\hat{\beta}}(x^*) + r(x^*)\Gamma^{-1}(Y - F)$$

where

- Y is the vector of observations $Y = (y^{(1)}, \dots, y^{(n)})^t$,
- $F = (f_{\hat{\beta}}(x^{(1)}), \dots, f_{\hat{\beta}}(x^{(n)}))^t$,
- $\Gamma = \sigma_\delta^2 (R_{\theta,q}(x^{(i)} - x^{(j)}))_{(i,j) \in \{1, \dots, n\}^2}$ is the covariance matrix of $\delta_n = (\delta(x^{(1)}), \dots, \delta(x^{(n)}))^t$,
- $r(x^*) = \sigma_\delta^2 (R_{\theta,q}(x^{(1)} - x^*), \dots, R_{\theta,q}(x^{(n)} - x^*))$, since $r(x^*) = E[\delta(x^*)\delta_n]$.

In the following, unknown regression and correlation parameters $\beta, \sigma_\delta, \theta$ and q are estimated by maximizing a likelihood function, see Fang et al. (2006):

$$\hat{\beta} = (F^t R_{\theta,q}^{-1} F)^{-1} F^t R_{\theta,q}^{-1} Y$$

$$\hat{\sigma}_\delta^2 = \frac{1}{n} (Y - F)^t R_{\theta,q}^{-1} (Y - F)$$

$$(\hat{\theta}, \hat{q}) = \arg \min_{\theta \in \mathbf{R}^{+d}, q \in [1, 2]^d} \left(\hat{\sigma}_\delta^2 |R_{\theta,q}|^{1/n} \right)$$

3. POLLUTANTS EMISSION TESTING

The problem which we handle consists in a contribution to the modeling of the pollutants emissions, with the aim of being able to make their rate predictable, most upstream to the phase of engine development. Obviously, this problem is not new but remains critical, see Christakos (1984). Besides, the good choice of the modeling method and detection of outliers are important as in any domain where data exploitation is used to make predictions, see Zhu et al. (2010).

As it has been said previously, our main objective is to establish a comparison between the two Kriging methods, especially for the industrial problem in interest. Are these two approaches equivalent? If they are not, which is the best of them regarding the specificity of our problem? To achieve this goal, we will apply those techniques over data received from a car manufacturer. The company wants to improve its engine responses owing to model based methods. In our case, the goal is to create models representing the pollutant emissions of a diesel engine. In order to be compatible with the engine tuning process, the inputs and the outputs of the models have been imposed by the cars manufacturer. The inputs of our models $x^{(i)} = (x_1^{(i)}, \dots, x_d^{(i)})$ are the following ones:

- the opening percentage of air flux, $x_1^{(i)}$

- the rail pressure, $x_2^{(i)}$
- the mass fuel quantities, $x_3^{(i)}$
- the instants of fuel injection, $x_4^{(i)}$
- the engine speed, $x_5^{(i)}$
- rhe EGR quantity, $x_6^{(i)}$
- the manifold pressure, $x_7^{(i)}$

The outputs of the models are the following ones:

- The NOx quantity in ppm (parts per million)
- The Particule quantity in FSN (Filter Smoke Number)
- The CO quantity in ppm (parts per million)

To solve this problem, we dispose of two databases. The first one contains 317 points. To test our models, we decided to use 267 points to create the models and 50 points to validate them. Consequently, we dispose of a learning database of 267 points and a prediction/validation database of 50 points. The second data base is composed of 90 points that are divided into a learning database of 75 points and a prediction/validation database of 15 points. the second database is smaller than the first one because the engine speed parameter is not taken into account We choose these two bases because they correspond to the two functioning modes possible in the engine tuning. The adjuster builds at first a representative global model of the behaviour of engine on the entire domain which interests him. Then, he builds several models corresponding to particular and more representative functioning points.

4. COMPARAISON OF THE TWO KRIGING APPROACHES

At first time, we want to be sure that both types of kriging, previously presented, can give a response to the modeling of pollutants emitted by diesel engine. To do it, we began with the modeling of NOx and particles. We do not consider CO at first, because, as we shall see it afterward, this measure is rather delicate and the databases can be erroneous. Furthermore, NOx are the pollutants that are modeled the most easily because their formation depends only on the temperature of combustion. It is why they are often used to test the good accuracy of the models which we want to elaborate.

In the Geostatistic Kriging context, we used a variogram whose the formula is the following one:

$$y = b + c * (r)^a$$

Considering the first database, figure 1, 2 and 3 show the results obtained.

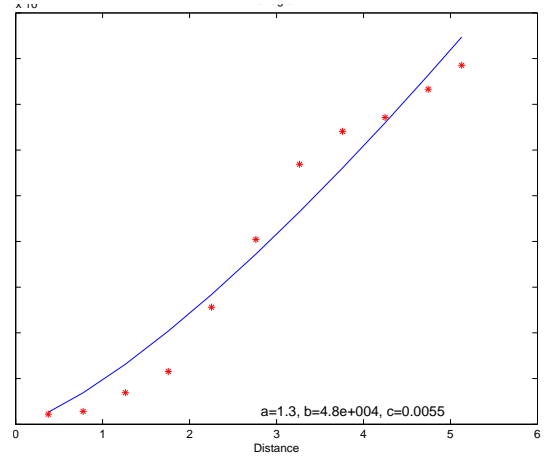


Fig. 1. adjusted variogram model

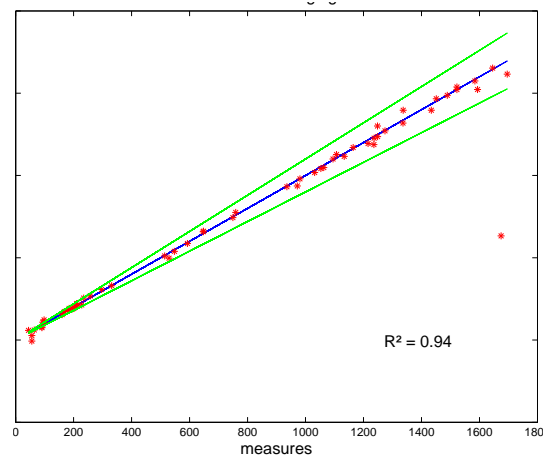


Fig. 2. geostatistical kriging model - NOx measures vs estimations

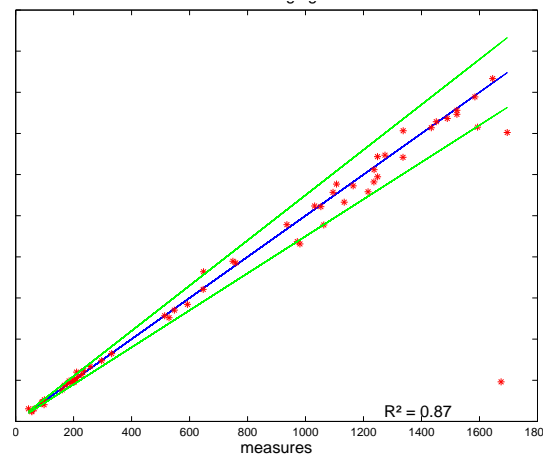


Fig. 3. statistic kriging model - NOx measures vs estimations

We can see that whatever the Kriging method used, the R^2 coefficient is satisfactory and is greater than 0.85. However, in the figures 2, 3, for the NOx, the R^2 coefficient obtained by the geostatistical approach is about 0.94 and is better

than the R^2 for Gaussian Kriging model ($R^2 = 0.87$). Moreover, we can notice that the fit of the variogram is correct with this function.

The results obtained for the particles are quite similar.

We reproduce the same reasoning with the second database which contains fewer test points than the first one. In the same way, we consider a power variogram model for the geostatistical approach: $y = b + c * (r)^a$.

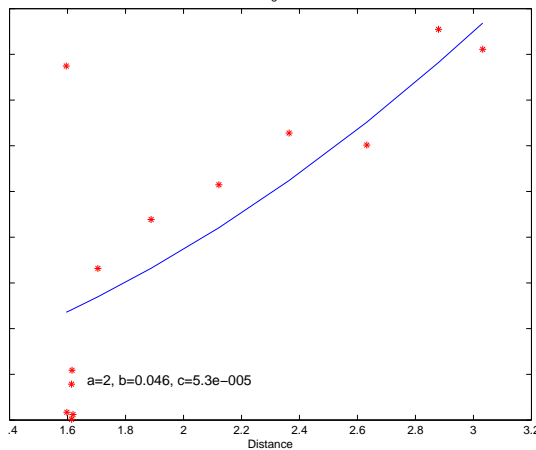


Fig. 4. adjusted variogram model

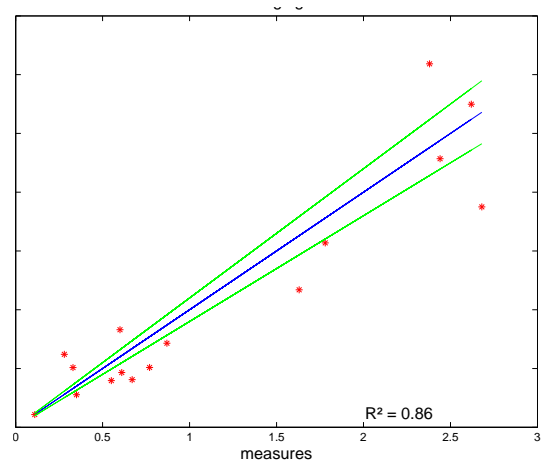


Fig. 6. statistic kriging model - particles measures vs estimations

For the NOx, which are not represented here, the results are quite satisfactory and completely equivalent with $R^2 = 0.97$ for the geostatistical kriging, and $R^2 = 0.95$ for the Gaussian Kriging. However, for this pollutant, the variogram is less adjusted. This can be explained by the small size of the database which we have and which limits the adjustment of the experimental variogram. As for particles, the results are just satisfactory with a value of R^2 around 0.85. This is explained partially by the insufficient size of the database but is especially due to the complexity of the physical phenomenon. We can conclude that the Kriging methods offer a good response to the problem of modelling pollutants emissions such as NOx and particles. Nevertheless, it often arrives that the measures contain errors. This is due to the context in which these measures are made and to the necessity of keeping instruments of post-treatment on the engine in spite of the modifications which it leads on the conditions of the combustion.

5. CO MODEL AND ATYPICAL POINTS

In what follows, we are interested in the pollutant CO by using the data contained in the first database.

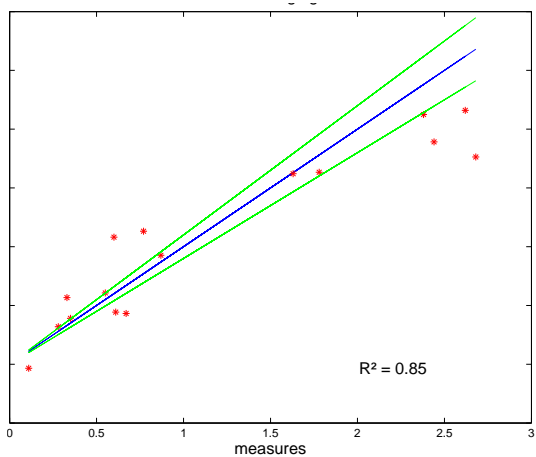


Fig. 5. geostatistical kriging model - particles measures vs estimations

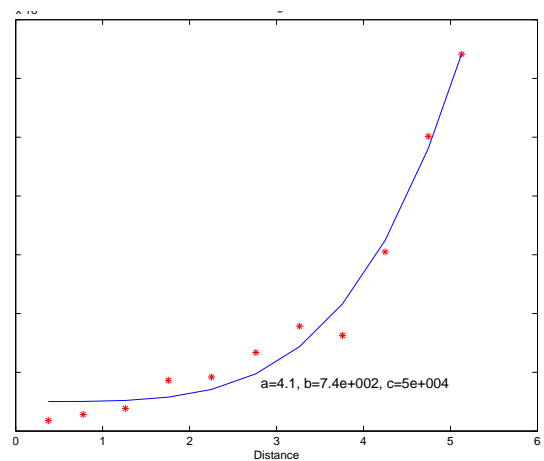


Fig. 7. adjusted variogram model

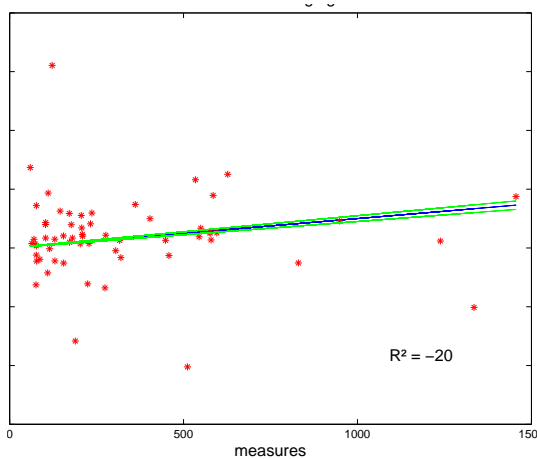


Fig. 8. geostatistical kriging model - CO measures vs estimations

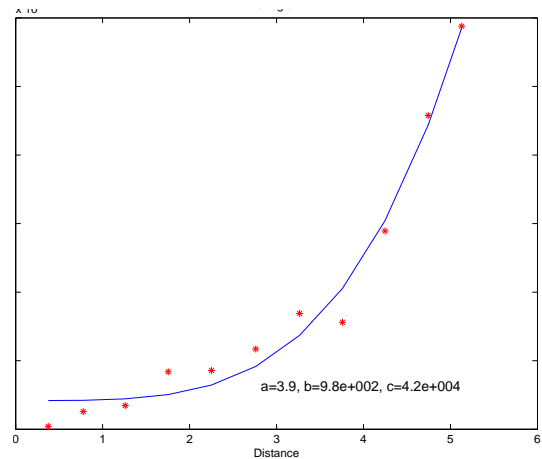


Fig. 10. adjusted variogram model

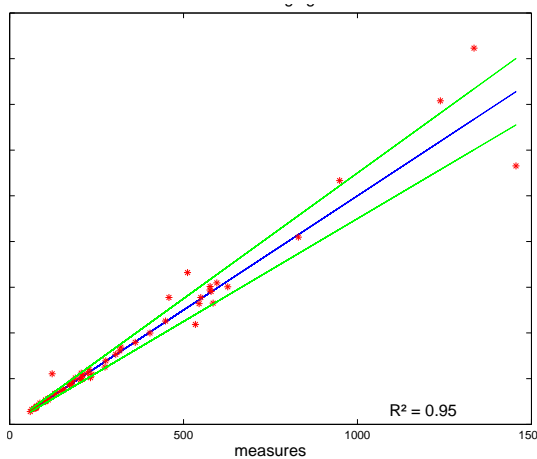


Fig. 9. statistic kriging model - CO measures vs estimations

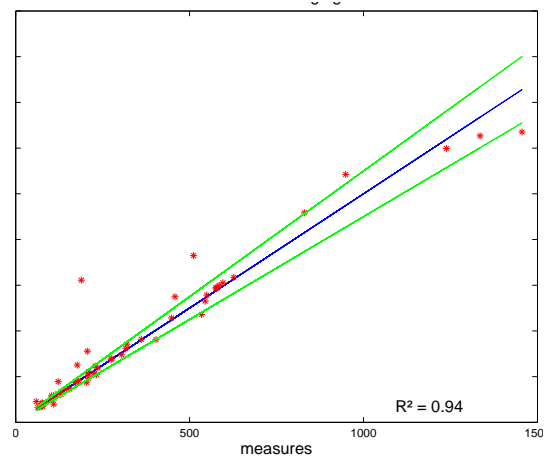


Fig. 11. geostatistical kriging model - CO measures vs estimations

We notice that the model obtained by the geostatistical kriging does not allow a correct prediction of the CO, in spite of the fact that the variogram fits correctly the experimental data. We notice that the gaussian kriging gives a very good value of R^2 equal to 0.95 (figure 9). Indeed, the results are linked to the presence of an atypical point in the learning database. We could consider this point as an outlier. However, nothing can allow considering it as erroneous datum of a physical viewpoint. It is possible that it corresponds to a particular mode of engine functioning or to the fact that the instruments of post-treatment, need a cleaning as for example for the Anti-particles Filter (FAP) which requires regularly a regeneration operation. To validate this supposition, we eliminate this point of our sample and we build the model again. The following results are obtained.

We notice that R^2 has a highest value 0.94 (figure 11) and is comparable to that one obtained with the Gaussian Kriging. Then, two hypothesis are possible. The first one is that the point is an outlier. Thus we have to wonder why the Gaussian model did not react to its presence while the geostatistical model is very sensitive. In the second hypothesis, this value is an atypical one, revealing an important phenomenon. In that case, the capacity of the geostatistical Kriging to have a different behaviour following the presence or no of atypical data can be very useful to inform us about particular behaviour of the engine. These hypothesis are very important specially if the tests must be exploited very quickly (even in real time) to detect points of particular functioning during engine tuning and calibration. The hypothesis was tested over the second data base, still on CO modelling:

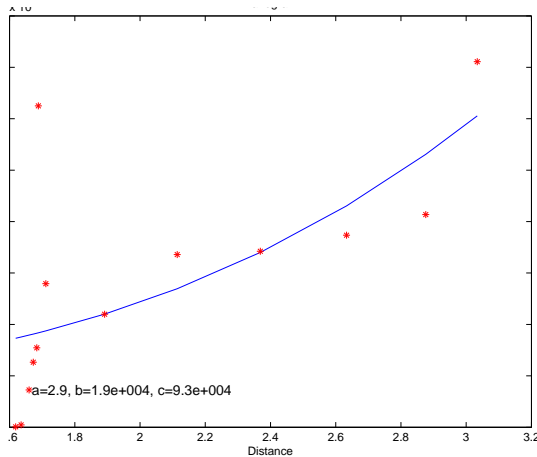


Fig. 12. adjusted variogram model

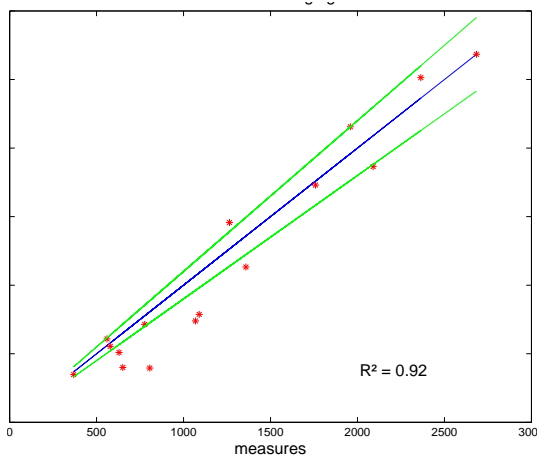


Fig. 13. geostatistical kriging model - CO measures vs estimations

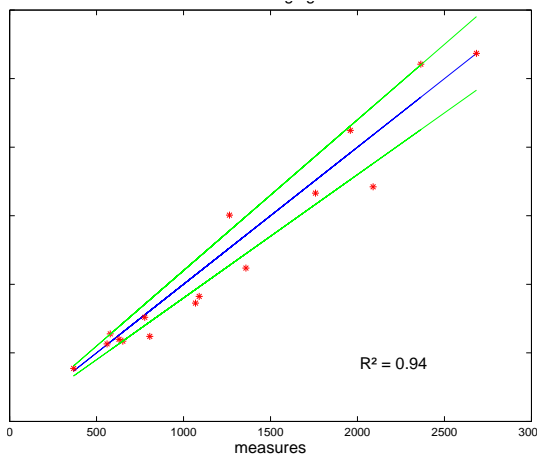


Fig. 14. statistic kriging model - CO measures vs estimations

The figures show that over this database it seems that there is no atypical point and in this case the two approaches produce quite the same results. The last test done is the introduction in this database of an outlier, which is

illustrated by the two figures 15 and 16.

In this case, the R^2 of the two models are degraded even if the statistical approach is still better than the geostatistical one. Thus, the comparison between the two approaches can provide an indication about the nature of a point: if it is an outlier or if it is an atypical point. If the statistical kriging gives acceptable results whereas the geostatistical one not, it can be assumed that there is an atypical point that must be studied.

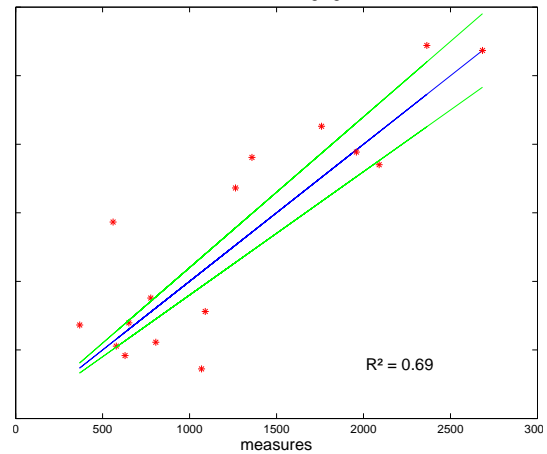


Fig. 15. geostatistical kriging model - CO measures vs estimations

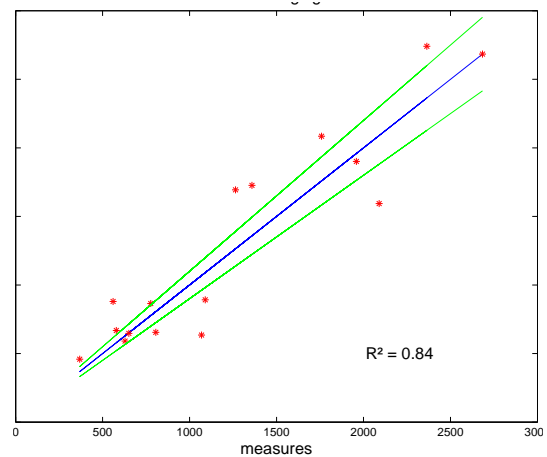


Fig. 16. statistic kriging model - CO measures vs estimations

5.1 Conclusions

The results we have described below, allow us to assert that in the field of diesel pollutants emissions, the kriging approaches can be mobilized. However, as outliers or atypical data are present or not, both methods have different performances. The geostatistical approach is recommended especially in our case, where often an atypical point corresponds to a particular behavior of the engine. These first results encourage us to continue the study, first to demonstrate its validity for modelling of other

pollutants, secondly by studying other models of kriging, such as cokriging, for example.

REFERENCES

- P. Abrahamsen *A review of Gaussian random fields and correlation functions* second edition Technical Report Norwegian computer center, 1997.
- M. Arnaud and X. Emery Estimation et interpolation spatiale *Hermes Science Publications*, Paris, 2000.
- S. Baillargeon, J. Pouliot, L.P. Rivest, V. Fortin and J. Fitzback interpolation statistique multivariable de donnees de precipitations dans un cadre de modélisation hydrologique *Colloque Géomatique 2004: un choix stratégique*, Montréal, 2004.
- M. J. Bayarri, Jim Berger, David M. Steinberg Special issue on computer modeling *Technometrics*, vol51, n4, nov 2009.
- M. Borg Jonathan and C.Akidas. Alex On the application of Wiebe functions to simulate normal and knocking spark-ignition combustion *Int Journal of vehicle design 2009*, vol 49, n1-3, p52-59, 2009.
- G. E. Box, W.G. Hunter and J.S. Hunter *Statistics for Experimenters: Design, Innovation, and Discovery* 2nd Edition, Wiley, ISBN 0471718130, 2005.
- E.H. Brahmi, L. Denis-Vidal, Z. Cherfi, V. Talon Identifiabilité et estimation des paramètres du modèle de combustion d'un moteur diesel *CD, Conférence Internationale Francophone d'Automatique, CIFA*, Nancy, 2010.
- S. Castric, V. Talon, Z. Cherfi, N. Boudaoud and P. Schimmerling Diesel engine combustion model for tuning process and a calibration method *IMSM07 The Third International Conference on Advances in Vehicle Control and Safety AVCS'07*, Buenos Aires, Argentina, 2007.
- G. Christakos On the problem of permissible covariance and variogram models *Water Resources Research*, 20(2), p 251-265, 1984.
- N. A. C. Cressie *Statistics for spatial data. Wiley Series in Probability and Mathematical Statistics Applied Probability and Statistics*. John Wiley and Sons Inc., New York. Revised reprint of the 1991 edition. A Wiley-Interscience Publication, 1991.
- J.C. Davis *Statistics and Data Analysis in Geology 2nd ed. John Wiley and Sons, New York*, 1986.
- S.P. Edwards, S. Michon, and G. Fournier The optimization of common rail FIE equipped engines through the use of statistical experimental design, mathematical modelling and genetic algorithms *S.A.E paper*, vol. 106, no3, p 505-523, 1997.
- K.T. Fang, R. Li R. and A.Sudjanto Design and modeling for computer experiments *Chapman and Hall/CRC*, 2006.
- J. Heywood *Internal combustion engine fundamentals London: Mac Graw-Hill*, 1988.
- B. Iooss Contribution au traitement des incertitudes en modélisation numérique: propagation d'ondes en milieu aléatoire et analyse statistique d'expériences simulées HDR, University of Toulouse III Paul Sabatier, 2009
- P.C. K. Jack Kriging metamodeling in simulation: A review *European Journal of Operational Research*, No 192, p 707716, 2009.
- R. Jin, W. Chen and T. W. Simpson Comparative studies of metamodeling techniques under multiple modeling criteria *Journal of Structural Optimization*, Vol. 23, No. 1, p 1-13, 2001.
- D.G. Krige A statistical approach to some basic mine valuation problems on the Witwatersrand *J. of Chem. Metal. and Mining Soc. of South Africa*. Vol. 52 p 119-139, 1951.
- G. Matheron Principles of Geostatistics *Economic Geology*, v. 58, no 8, (December 1963) p 1246-12688, 1963.
- T.J. Santner, B.J. Williams and W.J. Notz The design and analysis of computer experiments Springer, 2003.
- O.Zhu and H.S.Lin Comparing Ordinary Kriging and Regression Kriging for Soil Properties in Contrasting Landscapes *Pedosphere*, 20(5), p 594-606, Elsevier 2010.