Design of experiments for smoke depollution from the output of diesel engine

Matthieu CANAUD[(1*)], François WAHL [(1)], Céline HELBERT[(2)], Laurent CARRARO[(2)]

(*) corresponding author matthieu.canaud@ifp.fr
(1) IFP-Lyon, BP 3, 69390 Vernaison, France
(2) Ecole Nationale Supérieure des Mines de Saint-Etienne, 42023 St-Etienne Cedex 2

Keywords: computer experiments, kriging, non linear trend, derivatives information.

Computer codes based on mathematical models of physical systems are important tools in many fields of scientific research. Complexity of these systems increases due to the higher knowledge of kinetic behaviour and accuracy wanted. In this context, most of the time, usual tools are not sufficient to be as specific as we want. For example, ecological problems and environmental constraints require developing new systems depollution, more and more efficient, involving more and more complexity.

Our work consists in studying a catalytic depollution system, the post-treatment through NOx trap of smoke produced by diesel engines. As a surrogate of the real system, a kinetic model was developed to represent the physico-chemical phenomenon, depending on parameters (pre exponential factors, activation energies, adsorption constants) that cannot be obtained from theoretical considerations. Therefore, experiments are required to calibrate the model.

The final objective of this work is to suggest a criterion for experimental designs adapted to kinetic parameters identification, considering that; firstly, the kinetic model does not fit experimental data well, secondly the model is highly non linear and can be express by $y = f(x, \beta)$, where y is the response vector (for example the content of unburned hydrocarbons), x the experimental conditions and β the kinetic parameters of the model denoted by the function f.

In this paper, it is suggested to consider that differences observed between kinetic model and experimental data can be represented by a Gaussian process realization. Gaussian Process often accounts for correlated errors due to lack of fitness.

Thus, the following model is suggested,
$$y = f(x, \beta) + z_{\sigma^2, \theta}(x),$$
where $z_{\sigma^2, \theta}(x)$ is a centred Gaussian process with Gaussian covariance kernel specified by the variance σ^2 and scale parameters θ (see section 2 for an example of the covariance kernel). This approach is commonly used in the field of computer experiments and is known as kriging (see Fang *et al*, [1], Sacks *et al*, [5], Santner *et al*, [6]). However, in traditional use, the trend is linear and trend parameters estimation and its uncertainty are well known and obtained through an analytical formula.

The first difficulty of our framework is to estimate parameters considering the non linear trend. Similarly to nonlinear regression, the traditional analytical formula for β is replaced by a minimisation procedure. This procedure gives good results but some problems persist, especially high oscillations can be observed where response is less variable. Thus, in this paper, it is suggested to take into account derivatives knowledge. Hence, derivatives are supposed to provide useful information for kriging evaluation in that particular kinetic application, where the process is constant, that's to say at the beginning and at the end of the reaction. Besides, since differentiation is a linear operator, joint distribution between response and derivatives is still Gaussian. Thus, inference and prediction will be conducted in the same way.

Following points will be addressed in further research to those presented in the article.

1

Firstly, approximation of kriging prediction variance will be considered. Generalization of classical formula for prediction variance is not relevant. Alternatively, Bayesian approach to account uncertainty in parameters inference could be a good issue.

Secondly, experimental design adapted to our model and based on calculus of prediction variance is suggested. Our final goal is to combine this knowledge of uncertainty and input variables uncertainties in order to propose an optimal experimental design, which can be compared to classical ones: D-optimal Designs adapted to non linear problems or traditional space filling designs adapted to Gaussian Process.

All these methods are developed and tested on the catalytic system, i.e. on diesel engines smoke post-treatment.

This applied case, the NOx trap system is introduced in Section 1, in particular we present the inputs, outputs and parameters of the non linear model, and global functioning. The theory and mechanism of the kriging extended to a non-linear trend are outlined in Section 2, and the results obtained in our studying case, in Section 3. Section 4 is devoted to kriging with derivatives. A test model is introduced in Section 5. This test model is a similar case of the NOx trap but simpler. The kriging extension to non-linear trend with derivative information is applied on this test model, and results are exposed and discussed in Section 6.

1. NOx trap system : Presentation

The NOx trap is a catalytic system of smoke post-treatment for diesel engine, built into the exhaust line just at the engine output. The NOx trap has a dual function; firstly, the traditional function of oxidising hydrocarbons, denoted by HC, and carbon monoxide, denoted by CO, (produced by partial combustion due to a lack of oxygen); secondly, the treatment of the nitrogen oxide, NOx.

The NOx trap is a honeycomb structure with porous support impregnated with chemicals (platinum, barium, rhodium) in which exhaust gas pass trough. It works in two phases; the capture phase and the release phase. During the capture phase, the NOx contained in the exhaust gas from a lean-mixture (oxygen excess) are captured and stored on the support. During release phase, a chemical process known as reductive elimination purges the NOx trap of the stored NOx, with the engine operating in rich-burn mode, i.e. when the air-fuel mixture has just enough air for complete combustion of diesel. Nitrogen oxides are converted into neutral gases, mainly nitrogen. In this way the NOx trap is regenerated and is ready to trap more NOx. HC and CO are treated in continuous.
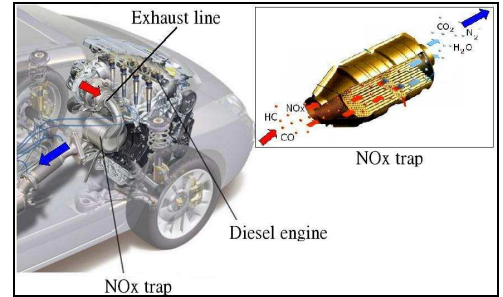

Figure 1 NOx trap in the exhaust line

Our work is done only on the capture phase, to reduce the complexity of the problem. Due to oxygen excess, in this phase, oxidations of HC, CO and NOx are the dominant chemical reactions.

Based on kinetic model retains, a computer model was construct, resulting from a differential equations system, denoted by f. The model can be express by,

$$y = f(x, \beta),$$

where y is the response vector, x the experimental conditions and β the kinetic parameters of the model. Inputs, denoted by x, are selected and controlled by the experimenter. First of all we have the mass flow of gas entering in the NOx trap, denoted by Q, then the mass composition of five species present in the exhaust gas, denoted by $c_i, i=1,...,5$. Those six first inputs are constant during a experiment. The last input is the gas temperature entering, which is not constant during the experiment, but increase as a function of time. The temperature profile is denoted by T. Hence,

$$x = (Q, c_1, c_2, c_3, c_4, c_5, T)^T$$

As output, we focus on the mass composition of the three species mentioned, HC, CO and NOx.

2

Twenty experiments were made, corresponding to different initial composition and mass flow of the gas. From these experiments, kinetic parameters were calibrated by generalized least-squares. Figures 2 and 3 illustrate the results obtained on one case. On the figure 2, the evolution of the concentration of CO (blue), HC (red) and NOx (green) is representing according to the temperature. Solid lines correspond to experimental results and dot lines to model results.
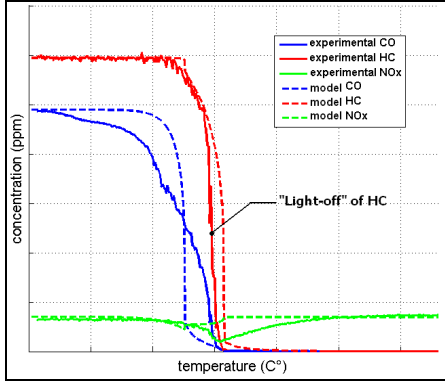

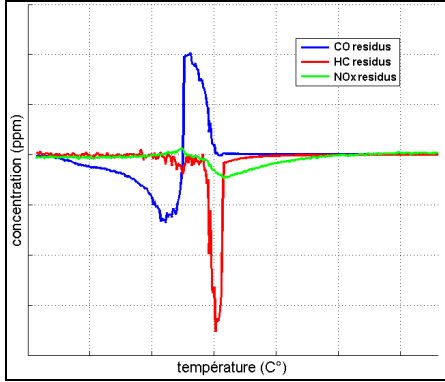**Figure 2 experimental and model CO, HC and NOx emission**

Figure 3 CO, HC and NOx residues' between experimental and model

In spite of calibration, differences between model and experimental results are important. For HC, the overall form of the model is near experiment results, but the light-off temperature, i.e. the temperature where half of the species is converted, is not well determined. In contrary, for CO, the light-off is well determined, but the model's form does not represent the experiment results. Concerning NOx, the model is very bad. To emphasize this difference, Figure 3 presents residuals for each specie. Clearly, we also note that the problem is non-stationary.

2. Kriging with non-linear trend

Considering that differences observed between computer model and experimental results can be represented by a Gaussian process, the following model is suggested,

$$y = f(x, \beta) + z_{\sigma^2, \theta}(x),$$

where $f(x, \beta)$ is a trend vector, compute by computer model, $x \in \mathbb{R}^k$ is the inputs and β kinetic parameters and where $z_{\sigma^2, \theta}(x)$ is a centred Gaussian process, with covariance kernel defined by,

$$Cov(z(x), z(x+h)) = \sigma^2 R_\theta(h),$$

where σ^2 is the process variance and θ is the scale parameter.

Note that Gaussian spatial correlation is used, defined by,

$$\forall h \in \mathbb{R}^k, R_\theta(h) = \exp\left\{ -\sum_{i=}^{k} \theta_i h_i^2 \right\}.,$$

This approach is commonly used in the field of computer experiments and is known as kriging. However, in traditional use the trend is linear. The difficulty is to estimate parameters considering the non-linear trend.

Let m be the number of design points, $Y = (Y_1, \ldots, Y_m)^T$ is the output observed at location $S = (s_1, \ldots, s_m)^T$, $s_i \in \mathbb{R}^k$. Similarly to non linear regression (see Seber and Wild [7]), the traditional analytical formula for β is replaced by a minimisation procedure. Using maximum likelihood estimation, expression of the kriging predictor, $\hat{y}$, and prediction variance, φ, at a new location x_0 are given by,

$$\hat{y}(x_0) = rR^{-1}Y - (F^T R^{-1} r - f)^T (F^T R^{-1} F)^{-1} F^T R^{-1} Y$$

and

$$\varphi(x_0) = \sigma^2 \left(1 + \left\| F^T R^{-1} F \right\|_{(F^T R^{-1} r - f)} + \left\| R \right\|_r \right)$$

where

$$\left\| A \right\|_u = u^T A^{-1} u$$

$$r = \left(R_{\hat\theta}(x_0 - s_1), \ldots, R_{\hat\theta}(x_0 - s_m) \right)^T$$

$$R = \left(R_{\hat\theta}(s_i - s_j) \right)_{1 \le i, j \le m}$$

$$F = f(S, \hat\beta) = \left(f(s_1, \hat\beta), \ldots, f(s_m, \hat\beta) \right)^T$$

$$f = f(x_0, \hat\beta)$$

where parameters are obtained by solving recursively the following simultaneous equations:

$$\hat{\beta} = \min_{\beta} \left(Y - F \right)^T R^{-1} \left(Y - F \right)$$

$$\hat{\sigma}^2 = m^{-1} \left(Y - F \right)^T R^{-1} \left(Y - F \right)$$

$$\hat{\theta} = \arg\min \left[\hat{\sigma}^2 \left| R^{-1} \right|^{\frac{1}{m}} \right]$$

Like universal kriging, kriging predictor and prediction variance expressions cannot be interpreted as conditional expectation and variance, for the same reasons outlined in Helbert *et al*, [2]. Besides, prediction variance expression depends on the estimation of β.

The algorithm determines $\hat{\theta}$, minimising the objective function $\hat{\sigma}^2 \mid R^{-1} \mid^{1/m}$, through modified Hooke and Jeeves method, described in Kowalik and Osborne, [3]. For more detailed, see Lophaven *et al*, [4].

3. Application of kriging with a non linear trend to the case study

Probabilistic model introduced in section 2 is estimated on the learning set composed of nineteen experiments. The last one is kept outside learning set to test results from estimation. Twelve points have been taken uniformly along the temperature, on each experiment. Hence, the design matrix has 228 points. HC and CO are treated independently.
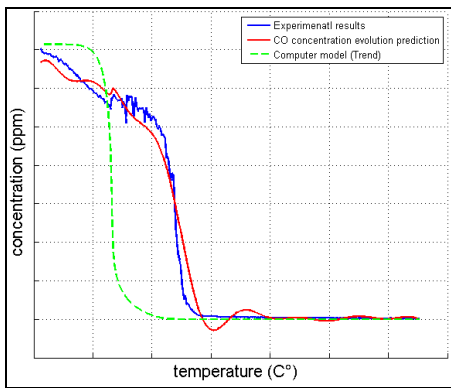

Figure 4 CO concentration evolution prediction by non linear trend kriging, with 12 points taken uniformly along temperature on the first nineteen experiments

Figure 4 presents predicted curve of CO concentration with temperature. Computer model

result (function $f(x, \beta)$) is drawn in green dot line, experimental result in blue and prediction in red. Figure 5 present HC concentration evolution prediction.
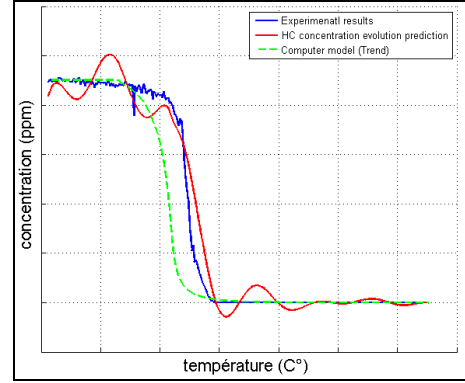

Figure 5 HC concentration evolution prediction by non linear trend kriging, with 12 points taken uniformly along temperature on the first nineteen experiments

Even if computer model does not fit experimental results well, predictions obtained by non linear trend kriging, for each output, are relatively close to observed results, i.e. Z plays a complement role between kinetic model an response model of the system.

However, oscillations appear on initial and final stages. Besides the fact that these oscillations do not correspond to the results, this leads to negative concentrations, which has no physical meaning.
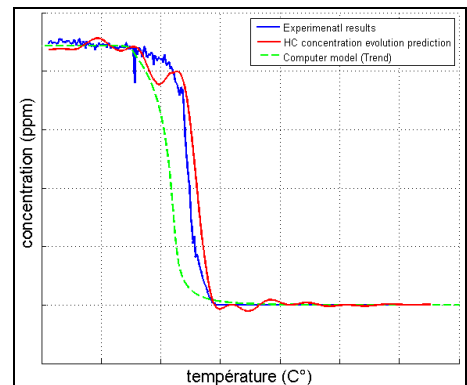

Figure 6 HC concentration evolution prediction by non linear trend kriging, with 15 points taken uniformly along temperature on the first nineteen experiments

To limit this phenomenon, it is possible to add points in the learning set by more discretizing in temperature direction. To illustrate this, 15 points have been taken,

4

instead of 12, to predict HC concentration evolution. Figure 6 outlined results.

Oscillations are less strong, but still present. Besides, these added points lead to numerical problems. Therefore, an envisaged solution to avoid these oscillations is to conditionalize process to derivatives as well as to responses. Theory is introduced in the next section.

4. Conditioning by derivatives : a test improvement for kriging with non linear trend

From the fact that both initial and final stages have a zero derivatives, conditioning kriging by derivatives could be a good issue to prevent oscillations. Santner *and al*, [6], introduced derivatives conditioning theory for universal kriging.

Let m_1 the design point's number where function value is known, m_2 the design point's number where partial derivative according to the first direction is known and exponent (1) denoted derivative output. Hence,

$$Y = (Y_1, \ldots, Y_{m_1}, Y_1^{(1)}, \ldots, Y_{m_2}^{(1)})^T,$$

and

$$F = (f(s_1, \beta), \ldots, f(s_{m_1}, \beta), f^{(1)}(v_1, \beta), \ldots, f^{(1)}(v_{m_2}, \beta))^T$$

The pairwise joint correlation function of $Y(\cdot)$ and $Y^{(1)}(\cdot)$ is given by,

$$Cov(Y(s_i), Y^{(1)}(v_j)) = -2\theta_1(s_i^1 - v_j^1) R_\theta(s_i - v_j)$$
$$Cov(Y^{(1)}(v_i), Y^{(1)}(v_j)) = (2\theta_1 - 4\theta_1^2(v_i^1 - v_j^1)^2) R_\theta(v_i - v_j)$$

Hence,

$$R = \begin{pmatrix} R_{00} & R_{01} \\ R_{01}^T & R_{11} \end{pmatrix},$$

where R_{00} is the $m_1 \times m_1$ matrix of correlations among the elements of $Y_i, 1 \le i \le m_1$, R_{01} is the $m_1 \times m_2$ matrix of correlations between Y_i and $Y_j^{(1)}, 1 \le j \le m_2$, and R_{11} is the $m_2 \times m_2$ matrix of correlations among the elements of $Y_j^{(1)}$.

Finally, the correlation vector, r, between the new location, x_0, and the observed data is defined by $r = (r_0, r_1)^T$, where $r_0 = (R_\theta(x_0 - s_i))_{1 \le i \le m_1}$ and $r_1 = (-2\theta_1(x_0 - v_j)R_\theta(x_0 - v_j))_{1 \le j \le m_2}$.

Kriging equations are the same as the ones introduced in section 2.

In order to improve this method, a test model has been constructed, which is presented in the next section. The results are presented and discussed in section 6.

5. Presentation of test model

In order to test our approach, a model, close to our studying case but simpler, is built. Model's characteristics are shown to be similar to our case, it means initial and final constant stages, non linear problem and a computer model which does not fit well "experimental results". Thus, we can make analogy between our studying case and this one.

First, consider a simple kinetic system of the form,

$$A \rightarrow B$$
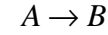
where A and B correspond to chemical species that react in catalyst presence. This system is governed by the following equations, depending on Langmuir-Hinshelwood formalization,

$$\begin{cases} \dfrac{d[A]}{dt} = -\dfrac{k_0 \exp\left\{-\dfrac{E}{RT}\right\}[A]}{1 + b_0 \exp\left\{-\dfrac{\Delta H}{RT}\right\}[A]} \\ [A]_0 = A_0 \end{cases}$$

where $[A]$ is the concentration of species A, $[A]_0$ its initial value, T the temperature, R the perfect gas constant and $\zeta = \{k_0, E, b_0, \Delta H\}$ the kinetic parameter vector, chosen to conduct to very different concentration evolution of A depending on temperature.

Hence, inputs are temperature, time and initial concentration. Output is concentration evolution of A according to the temperature.

To make the analogy with studying case, results of this system, noted *Y*, are named "experimental results".

Secondly, another kinetic system is determined, which represent "computer model",

$$\begin{cases} \dfrac{d[A]}{dt} = -k_0' \exp\left\{-\dfrac{E'}{RT}\right\}[A] \\ [A]_0 = A_0 \end{cases}.$$

Let $f(x,\beta)$ the model governed by this system, where β is kinetic parameter vector. Parameters on the Gaussian process $z_{\sigma^2,\theta}(x)$ have to be determined, such as,

$$y = f(x,\beta) + z_{\sigma^2,\theta}(x).$$

In order to display that this case is similar to studying case, figure 7 expose simulations obtained through both model.
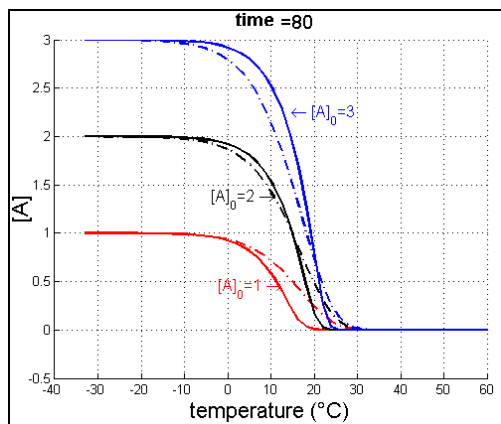

Figure 7 "Experimental results" in solid line, and "Computer model" results in dot line according to temperature for 3 initial concentration and time=80sec

Despite of kinetic parameter calibration, computer model does not fit experimental results well. Besides, the shapes of results correspond to studying case, i.e. initial and final constant stages, with decreases more or less important.

Non linear trend kriging with and without derivatives information is applied on this test model in order to conclude about relevance of derivatives knowledge.

6. Non linear trend kriging with derivatives information results

For sake of convenience, function and derivative values are supposed to be available for each design point, i.e. $s_i = v_i, 1 \le i \le m_1 = m_2$.

Evolution prediction is done on experimental design construct as follow. Three initial concentrations are chosen, eight fixed time and ten temperatures on each simulation are determined uniformly. Hence, experimental design has 240 points.

Kriging extended to non linear trend is first realized, and then derivative information is added, in order to conclude about derivative information contribution

Figure 8 presents results obtained. Red lines correspond to prediction, dot one to kriging extended to non linear trend, solid one to kriging with derivative information. Blue solid line corresponds to real result, and black points are experimental points taken in this experimental condition.
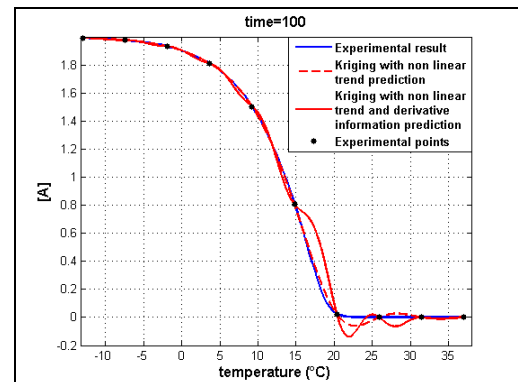

Figure 8 Test model prediction by kriging extended to non linear trend (red dot line), and kriging extended to non linear trend with derivative information (red solid line) compare to real results (blue line)

Oscillations problem appears to be stronger. Due to derivative information, prediction is too constraint, involving this phenomenon.

Taking derivative information on all design points is not relevant, but in our studying case, initial and final constant stages only have to be considering for derivative conditioning.

At the moment of redaction of this paper, results are not yet available, further investigations have to be done, and this approach should be test on our studying case.

7. Conclusions

Two points have been described in this paper. Firstly, kriging extended to non linear trend and then a test improvement of this method through derivative information.

Kriging with non linear trend gives good results but some problems occur. First, numerical problems when number of design points increases, secondly, oscillations appear where response is less variable.

Use of derivative information where the process is constant, that's to say at the beginning and at the end of the reaction seems to be relevant to correct this problem, but results obtained, for the moment, on test model, tend to demonstrate that this is not the case. Anyway, further investigations have to be conduct to be sure of our conclusion, meanly conditioning by derivative information only on both stages. For the moment, results are not available.

Reference

[1] Fang KT, Li R and Sudjianto A. *Design and Modeling for computer experiments.* Chapmann & Hall/CRC, 2006.

[2] Helbert C, Dupuy D and Carraro L. "Assessment of uncertainty in computer experiments: from universal kriging to bayesian kriging", *Applied Stochastic Models in Business and Industry*, 25, 2009, 99-113.

[3] Kowalik J and Osborne MR. Methods for unconstrained optimization problems. *Elsevier*, 1968. New-York, USA.

[4] Lophaven SN, Nielsen HS and Sondergaard J. *Aspect of the Matlab toolbox DACE.* Informatics and mathematical modelling, DTU, 2002, 44 pages.

[5] Sacks J, Welch WJ, Mitchell TJ and Wynn HP. Design and analysis of computer experiments. *Statistical Science* 1989; 4:409-435.

[6] Santner T, Williams B, Notz W. *The Design and Analysis of Computer Experimetns.* Springer: Berlin, 2003.

[7] Seber GAF and Wild CJ. *Nonlinear regression.* Applied probability and statistics John Wiley & Sons, 1989. New-York.