XXI IMEKO World Congress "Measurement in Research and Industry"
August 30 – September 4, 2015, Prague, Czech Republic

PROBABILITY OF CONFORMITY ASSESSMENT FOR COMPUTATIONALLY EXPENSIVE SYSTEMS: APPLICATION TO FIRE SAFETY ENGINEERING

Séverine Demeyer [1], *Damien Marquis* [2], *Nicolas Fischer*[3]

[1]Laboratoire National de Métrologie et d'Essais, Trappes, France, severine.demeyer@lne.fr
[2] Laboratoire National de Métrologie et d'Essais, Trappes, France, damien.marquis@lne.fr
[3] Laboratoire National de Métrologie et d'Essais, Trappes, France, nicolas.fischer@lne.fr

Abstract – The use of computational codes has become common practice when experiments are not feasible or when their number is too parsimonious. The statistical modelling of numerical experiments with kriging models yields a probabilistic decision framework to assess the probability of failure of the system and its associated uncertainty. In this work, fast low-fidelity simulations are combined with costly high-fidelity simulations in a co-kriging model and points are sequentially designed to reduce the number of costly simulations. The methodology is applied to a fire engineering case study.

Keywords: multifidelity computer experiments, Gaussian process, probability of failure, sequential planning

1. INTRODUCTION

When dealing with a computational code, one may be interested in propagating the uncertainties associated with the input quantities to evaluate the uncertainty associated with the output quantity.

In the decision theoretical framework pertaining to conformity assessment, one is interested also in the position of the output variable with respect to a given threshold (regulatory threshold...). The problem of knowing whether the output of a computationally expensive model exceeds a given threshold is very common for reliability analysis and safety-critical applications such as aerospace, nuclear power stations and civil engineering (models of bridges and buildings etc.). The threshold can be interpreted as a lower specification limit defining a one-sided tolerance interval.

The statistical methods that are usually used to deal with the propagation of uncertainties may need to be adapted to compute the probability of exceeding a threshold. This is particularly the case when the probability is small or when the code is computationally expensive. For instance, Monte Carlo methods based on a large number of calls of the code may be replaced by Monte Carlo simulations from a metamodel of the code which has just required a few runs to be calibrated (these runs form the training database).

The statistical modelling of numerical experiments with kriging models based on Gaussian processes (Santner, Williams and Notz [1], Rasmussen and Williams [2]), yields a probabilistic decision framework to assess the probability of failure of the system and its associated uncertainty.

In order to decrease the burden of costly simulations when predicting the output of a system, the combination of fast low-fidelity simulations with costly high-fidelity simulations has proved an efficient method (Xiong, Xian and Wu [3], Kennedy and O'Hagan [4]).

Besides, the knowledge encompassed in the metamodel can be increased by sequentially adding points to the training database. These new points are obtained after optimization of a criterion depending on the quantity to be estimated (Chevalier, Picheny, Ginsbourger [5], Bect, Picheny, Ginsbourger and Vazquez [6]). Indeed, for a propagation of uncertainty problem targeting the mean and the variance of an output quantity, a space filling criterion exploring the whole input domain is adequate. In our case, targeting the probability of potentially rare events requires a criterion favouring the exploration of the failure domain.

In this paper, sequential sampling in targeted region is applied to co-kriging metamodel to estimate the probability of conformity and its associated uncertainty. After brief presentation of the concepts, the methodology is illustrated in a one-dimensional toy example and applied to a simplified fire engineering application.

2. CO-KRIGING OF TWO CODES

Denote $F_2(.)$ the higher level code providing accurate but expensive simulations and $F_1(.)$ the lower level code providing cheap approximations of $F_2(.)$. Denote respectively $\tilde{y}_2(.)$ and $\tilde{y}_1(.)$ the statistical surrogates of the codes established from a limited number of simulations at each level on a domain D. Let D_1 and D_2 (with $D_2 \subseteq D_1 \subset D$ for computational convenience) be the training databases of $F_1(.)$ and $F_2(.)$ respectively.

A co-kriging metamodel of two codes relates their kriging surrogates in a regression model where the higher level of code $\tilde{y}_2(.)$ is explained by the lower level of code $\tilde{y}_1(.)$ as in Kennedy and O'Hagan [4]

$$\tilde{y}_2(x) = \rho\tilde{y}_1(x) + \delta(x) \quad (1)$$

where ρ can be seen as a regression parameter between the surrogate outputs, $\delta(.)$ controls the variability of $\tilde{y}_2(.)$ that is not explained by $\tilde{y}_1(.)$ and is independent from $\tilde{y}_1(.)$.

This model requires hypotheses on the relationship between the codes and prior knowledge on the behaviour of the codes. In particular, if the lower level code is modelled as a Gaussian process $\tilde{y}_1(.)$ and if the regression error is also modelled as a Gaussian Process $\delta(.)$, the higher level code results in a Gaussian process $\tilde{y}_2(.)$. Details about the models, the estimation of parameters and the prediction can be found in Kennedy and O'Hagan [4]. The implementation in an R package can be found in Le Gratiet [7]. Under convenient hypotheses described in Kennedy and O'Hagan [4], predictions at untried inputs in D follow a Gaussian distribution which allows computing pointwise probabilities of exceeding a threshold (also called probability of excursion) based on the Gaussian cumulative distribution function Φ

$$\pi(x) = P(\tilde{y}_2(x) > s) = \Phi\left(\frac{\hat{m}(x) - s}{\hat{\sigma}(x)}\right) \quad (2)$$

where $\hat{m}(x)$ and $\hat{\sigma}(x)$ denote respectively the predicted mean and standard deviation at input x.

3. COMPUTING THE PROBABILITY OF NON CONFORMITY

The targeted probability of exceeding a threshold *s* (non-conformity) is

$$p_f = P(x \in D, \tilde{y}_2(x) > s) \quad (3)$$

The best estimator $\hat{p}_f$ of p_f that minimizes the mean squared error (MSE) is $\hat{p}_f = \int \pi(x) f(x) dx$ where $\pi(x)$ is the probability of excursion at input x defined in (2).

A measure of uncertainty associated to this estimator is the mean squared error (MSE)

$$u^2(\hat{p}_f) = E_{\tilde{y}_2}\left((p_f - \hat{p}_f)^2\right) =: MSE$$

These estimators are computed with Monte Carlo method based on simulated trajectories from $\tilde{y}_2(.)$ as in Oakley and O'Hagan [8].

4. SEQUENTIAL CO-KRIGING

Sequential co-kriging consists in sequentially adding points to the training databases D_1 and D_2 (with $D_2 \subseteq D_1 \subset D$) that will be evaluated by the two code levels, respectively the lower fidelity code and the higher fidelity code, to decrease the relative uncertainty (coefficient of variation $cv(\hat{p}_f) = \frac{u(\hat{p}_f)}{\hat{p}_f}$) of the estimated probability of non-conformity $\hat{p}_f$ with associated uncertainty $u(\hat{p}_f)$. The procedure is repeated until a targeted relative uncertainty is reached (cv.seuil).

Sequentially adding points to a training database (also called sequential planning) according to a space filling design brings a global reduction of uncertainty (see Xiong and Qian [3]). However this approach is not necessarily pertaining to the problem of determining the failure area and its probability accurately. Indeed when the target area has a low probability, points have a very low probability to fall in. This is particularly critical in high dimensional problems. Specific sequential methods targeting the failure area with dedicated criteria (such as the tMSE: targeted Mean Square Error) are required. Such methods have already been developed to estimate a probability of failure based on a kriging model of a computer code (see Bect, Ginsbourger, Picheny, Vazquez [6] and Roustant, Ginsbourger, Deville [9]). For instance, the tMSE criterion yields points with high predicted variance in the target area.

Since the posterior predictive distribution of $F_2(.)$ outputs denoted $\tilde{y}_2(.)|y, \hat{\gamma}$ given observations $\boldsymbol{y}$ and estimated parameters $\hat{\gamma}$ is a Gaussian process such methods can be applied to $\tilde{y}_2(.)|y, \hat{\gamma}$.

The sequential co-kriging procedure returns an estimate of the probability of failure (non-conformity) of the higher fidelity code based of a limited number of higher fidelity runs thanks to numerous cheap approximations provided a valid relationship between the two codes.

5. ILLUSTRATION OF SEQUENTIAL CO-KRIGING IN A 1-D TEST CASE

This section illustrates in a 1-dimensional example the influence of a lower level code while predicting the output of a given code and the effect of the sequential procedure on the estimation of the probability that the output of higher level code exceeds a threshold. In this example, both codes are supposed known.

Define, for $x \in \left[-\frac{3\pi}{2}, \frac{5\pi}{2}\right]$, $F_2(x) = x\cos(x)$ the higher level code and $F_1(x) = 0.5F_2(x)$ the lower level code, represented respectively as the black and red curves on Fig. 1 and Fig. 2.

Let us assume that 9 points are run with the lower level code $F_1(.)$ and, among those points, 6 points are run with the higher level code $F_2(.)$, displayed respectively as red triangles and black points.

A kriging model is fitted to the 6 black points with a constant mean and Gaussian covariance. Based on the knowledge brought by these points on the code $F_2(.)$, the kriging theory provides the best curve estimating $F_2(.)$ and 95% confidence bounds represented as the grey area on Fig. 1.

When the 9 points evaluated at the lower level are analysed with the higher level points, we observe that the 95% confidence bounds around the estimated curve (represented as the yellow area on Fig. 1 is significantly

reduced and the estimated curve (dashed black curve) is shifted closer to the true $F_2(.)$ function.

Let us consider the event $F_2(.) > 5$ and let us estimate its probability and its associated uncertainty from the co-kriging model based on sections 3 and 4. Fig. 1 shows that the true function may lie anywhere within the yellow area with probability 95%, which brings uncertainty on the probability of exceeding the threshold, as seen in the results $\hat{p}_f = 0.099$, $u(\hat{p}_f) = 0.015$ and $cv(\hat{p}_f) = 0.152$.

In order to reduce the uncertainty in the targeted area, a new point x = 7.847 is added to the database according to the tMSE criterion.

The updated kriging model shows decreased uncertainties in the targeted area with a bias on Fig. 2, that is that kriging alone doesn't allow to estimate the true function. However, co-kriging results displayed on Fig. 2 show no more uncertainty in the target area so that the probability of exceeding the threshold is exactly $\hat{p}_f = 0.1$.

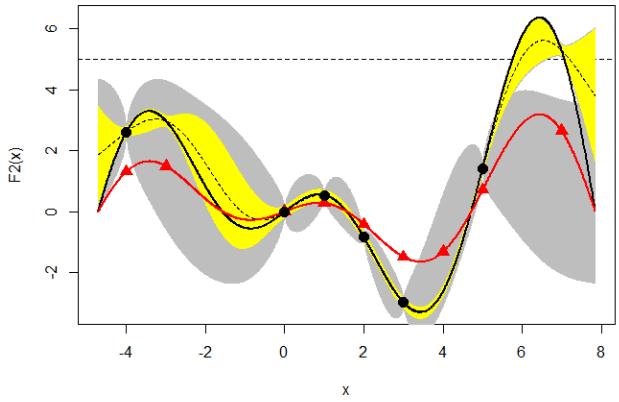


Fig. 1 Plot of the kriging and co-kriging confidence intervals (resp. in grey and yellow) to predict F_2. In red: the F_1 curve and its observations (red triangles). In black: the F_2 curve (plain curve), its best co-kriging estimate (dashed curve) and its observations (black circles).

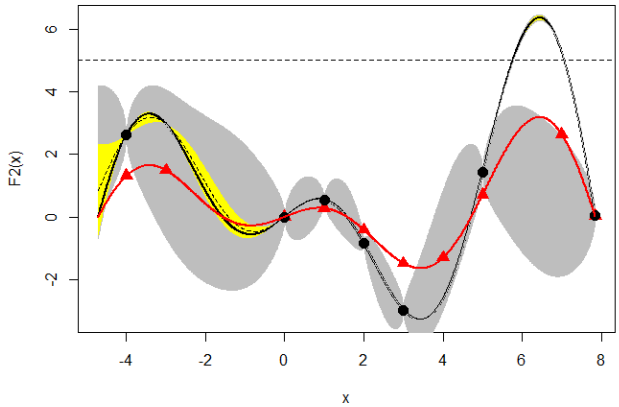


Fig. 2 Plot of the updated kriging and co-kriging confidence intervals (resp. in grey and yellow) after a new point obtained with the tMSE criterion has been added to the databases D_1 and D_2. In red: the F_1 curve and its observations (red triangles). In black: the F_2 curve (plain curve), its best co-kriging estimate (dashed curve) and its observations (black circles).

6. FIRE ENGINEERING APPLICATION

6.1. Conformity in fire safety engineering

The prediction of the evolution of tenability conditions in an environment in fire can be estimated using complex fire models like zone model or computational fluid dynamics tools and requires setting appropriate tenability criteria which ensure that the occupants will not be exposed to untenable conditions. Nowadays, there is no single set that is universally accepted and various tenability criteria are proposed (Purser [10]). The tenability criterion used in the present study comes from the international standard ISO 13571 [11] and is based on the best available scientific judgement of the consequences of human exposure to heat.

According to this standard, the thermal risk is linked to the heat flux associated to the radiation of the smoke and the temperature of the atmosphere. The tenability limit for exposure of skin to radiant heat is approximately 2.5 kW/m^2. Below this incident heat flux level, exposure can be tolerated for 30 min or longer without significantly affecting tenability. The radiant heat limit of 2.5 kW.m^2 may be reached when the hot layer temperature UL rises above 200°C.

The probability of non-conformity arising in this study is thus: P(UL>200°C) considering various fire scenario defined in section 6.3 resulting from a combination of input variables.

6.2. Case study

For the purpose of the study, a known building is considered with dimensions 19.75m (length), 12m (width), 16.50m (height) along with those of the openings with negligible uncertainty. The thickness and the thermal properties of the walls are defined in accordance with the present structures. The test hall is equipped with two doors assumed as open and two natural smoke removal system (see Fig. 3).

The room mesh is defined from the actual dimensions of the test hall, as shown in Fig. 3. In the case of FDS, the grid is uniform, and the cell dimensions are 25 cm on each side. This cell size is a compromise between flow resolution and computational time. The fire source is placed in the center of the test hall.

6.3. Input variables

Apart from changes in environmental conditions (such as outside and inside temperatures, ambient pressure and relative humidity, wind velocity and direction), the properties of the fire (fire source area, fire growth rate and heat release rate (HRR) per unit area) in the hall test is governed by the physical and chemical process evolved.

Multiple interactions between these input variables at different time during the fire may affect the pattern of the fire growth and lead to uncertainties. For this reason, there is a need to determine the uncertainty (probability) with which input variables may affect a real fire in a known building. In the present framework, the environmental conditions and the properties of the fire are therefore randomly determined. The chosen probability distributions for these input variables

are detailed in Table 1. A sensitivity analysis showed that only the fire area A_f and the characteristic heat release rate (HRR) $\dot{Q}''$ have an influence on the exceedance of the threshold. So only these two variables will be analysed and the others are kept constant.

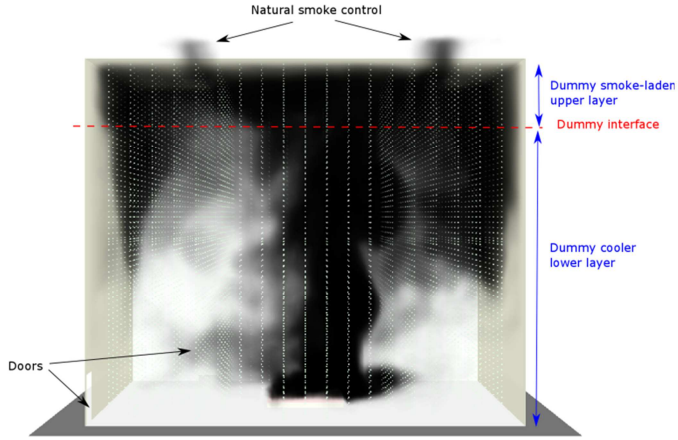


Fig. 3 Simulation of fire plume in a building. Numerical room with the CFD code Fire Dynamics Simulator. Smoke plume obtained from a fire area of 13.75m2 and a heat release rate of 487kW.m-2. Dummy regions and layer height are represented.

Table 1 Description of the input variables of the fire code (N: normal, U: uniform).

Variable	Name	Unit	Range	Distribution
P_{atm}	Atmospheric pressure	Pa	$[98000, 102000]$	N
T_{ext}	External temperature	K	$[263.15, 303.15]$	N
T_{int}	Inside temperature	K	$[290, 303.15]$	N
α	Fire growth rate	$kW.s^{-2}$	$[0.011338, 0.20]$	U
A_f	Fire area	m^2	$[1, 20]$	U
$\dot{Q}''_f$	Characteristic HRR per unit area	$kW.m^{-2}$	$[300, 500]$	U
q''_f	Design fire load density per unit area	$MJ.m^{-2}$	$[300, 600]$	U

6.4. Numerical codes

Two numerical tools are considered in this application: CFAST ($F_1(.)$) and FDS ($F_2(.)$) used respectively as the lower level and the upper level code.

The Consolidated Model of Fire and Smoke Transport, CFAST, is a fire model which relies on the assumption that a volume is subdivided in two zones.

Upper level CFD simulations were performed with the version 6.1.2 of the Fire Dynamics Simulator (FDS) code developed by the NIST Building and Fire Research Laboratory (USA) and the VTT technical research centre of Finland. FDS is far more complex because there are not two distinct zones, but rather a continuous profile of temperature (see McGrattan, McDermott, Weinschenk, and Overholt [12]) as illustrated on Fig. 4.

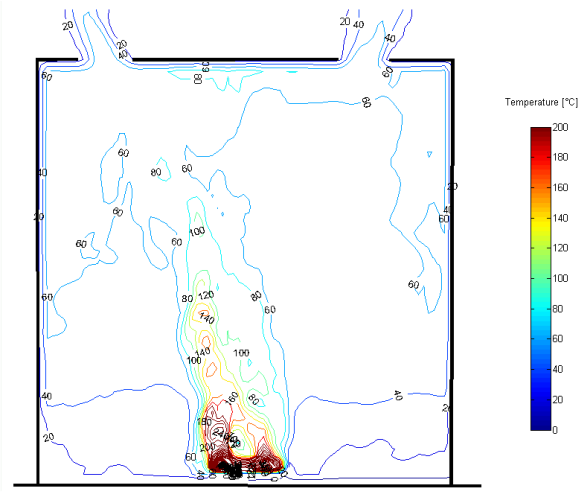


Fig. 4 Iso-temperature located at the term source (y = 0) at t= 200s. Results obtained from a fire area of13.75m^2 and a heat release rate of 487kW.m^{-2}.

6.5. Initial databases

The initial database evaluated by CFAST ($F_1(.)$) comprises 9 points displayed in Fig. 5, among them 5 points (in red) are evaluated by FDS ($F_2(.)$). Voluntarily, no FDS points exceed the threshold. The domain of variation of A_f and $\dot{Q}''$ is [1-20]m² x [300-500] kW/m² respectively.

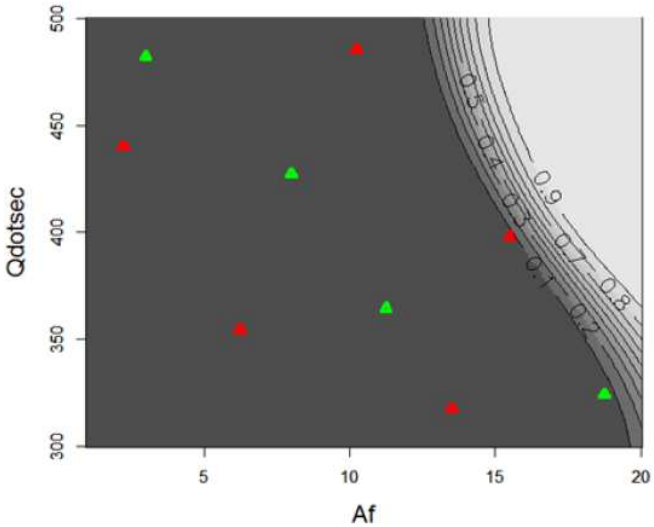


Fig. 5 Initial databases: in red: points evaluated by FDS and CFAST (D_2), in green: points evaluated only by CFAST (D_1). The contour plot of the probability of excursion function is also displayed.

6.6. Analysis

Numerical interpretation of the sequential co-kriging results

Remark: In this paper, the additional FDS points required to carry out the sequential co-kriging method have been replaced by the predicted mean of the current surrogate of FDS at each iteration to avoid modelling numerical instability.

A comparison of kriging and co-kriging methods has been carried out on the fire engineering case study to show the influence of the number of points in the database and the influence of their location on the estimates of the probability of conformity and its accuracy (relative uncertainty). Results are displayed in Table 2.

For the same number of FDS runs (5 runs) co-kriging reduces the relative uncertainty of the point estimate. A relative uncertainty divided by 3 ($cv(\hat{p}_f) = 0.13$ versus $cv(\hat{p}_f) = 0.3968$) was obtained at the cost of 4 cheap CFAST simulations. This shows the positive impact of combining expensive runs with cheap runs.

Final results of the sequential co-kriging show a dramatic decrease of the relative uncertainty with respect to co-kriging results ($cv(\hat{p}_f) = 0.0078$ versus $cv(\hat{p}_f) = 0.13$) at a cost of only 6 smartly chosen new expensive runs.

Table 2 Summary results of the probability of non-conformity obtained with various kriging based methods.

number of FDS runs	type of analysis	$\hat{p}_f$	$u(\hat{p}_f)$	$cv(\hat{p}_f)$
5	kriging	0.1634	0.0648	0.3968
5	co-kriging	0.172	0.022	0.13
$5 + 3 \times 2$	sequential co-kriging	0.1710	0.0013	0.0078

Graphical interpretation of the sequential co-kriging results

Results obtained at each iteration of the sequential co-kriging procedure (see section 4) are displayed in Table 3. The first line (iteration 0) gives the co-kriging based Monte Carlo estimates obtained with the initial database (9 points including 5 expensive FDS simulations) displayed at line 2 of Table 2.

The contour plot of the posterior probability of excursion function on the full domain provides a graphical tool to assess the efficiency of the iterative algorithm.

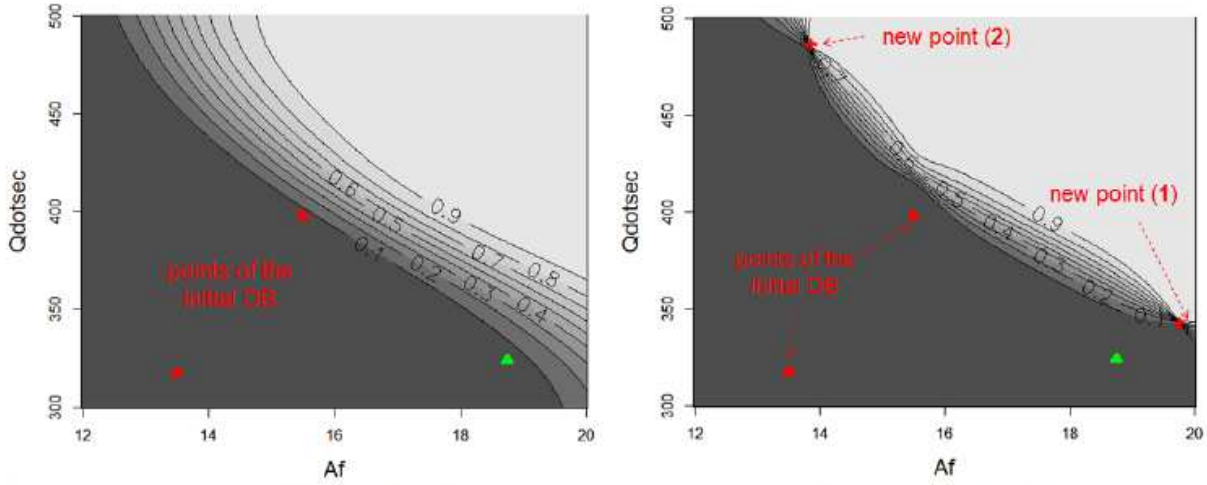
The initial contour plot displayed on Fig. 6a shows a large dispersion of the level lines, representing the uncertainty on the predicted frontier between non-conformity (white area) and conformity (black area). For instance the predicted 0.5 line indicates that there is 50% chance that the non-conformity domain lies above this line and the predicted 0.9 line indicates that there is 90% chance that the non-conformity domain lies above this line.

As points are sequentially added to the database, the uncertainty is reduced. The first two points added (see Fig. 6b) (red crosses) have a high predicted variance in the target area, as they are far from FDS points (red triangles). Their predicted mean is close to the threshold so that the level lines interpolate the points. The lines are also narrowed in between, close to an initial FDS point, which shows the effect of this point given the additional knowledge.

Table 3 Summary results of the probability of non-conformity obtained at each iteration of the sequential procedure. Iteration 0 gives the baseline results obtained with the initial database (9 points).

iteration	number of added points	$\hat{p}_f$	$u(\hat{p}_f)$	$cv(\hat{p}_f)$
0	-	0.172	0.022	0.13
1	2	0.1707	0.0076	0.0446
2	2	0.1709	0.0023	0.0132
3	2	0.1710	0.0013	0.0078

The next two points added (see Fig. 7Fig. 6a) have also a predicted mean close to 200°C and yield reduced uncertainties so that 1-probability of non-conformity zones appear.



(a) Baseline : points of the initial database (b) Iteration 1: add 2 points

Fig. 6 Zoomed in. Contour plot of the initial probability of excursion (a) and the updated probability of excursion after 2 points have been added at the first iteration of the sequential co-kriging procedure (b).

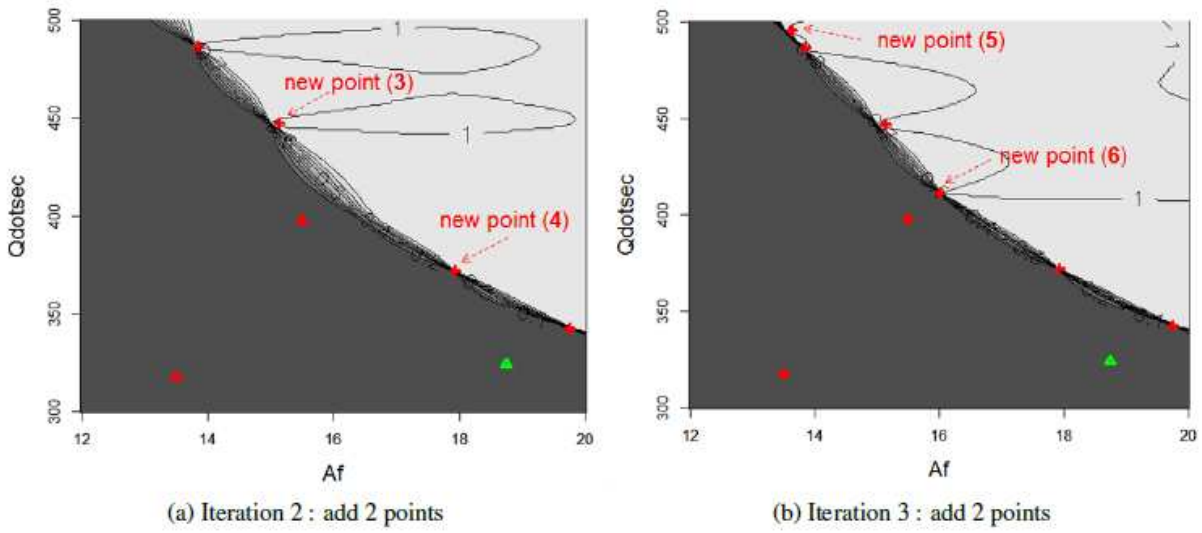


Fig. 7 Zoomed in. Contour plot of the updated probability of excursion at the second iteration of the sequential cokriging procedure (a) and at the third iteration (b).

Finally, two more points are needed (see Fig. 7b) to reach cv =1%, with the effect of creating a 1-probability of non-conformity zone, which means that fires starting with coordinates lying in this zone will lead to non-conformity (UL > 200°C).

CONCLUSION

The methodology presented in this document to assess the conformity of a characteristic given a specified threshold, using an expensive computational code, has proved efficient to overcome too parcimonious evaluations from an expensive code when fast approximations are available. This method builds on the kriging models usually used to model code outputs, to improve the predictions of the kriging model of the expensive code. Another desirable feature is that the method allows smart sampling of new points targeting the non-conformity domain in order to further reduce the uncertainty of the probability of non-conformity. Although demonstrated in a simple but realistic case study, the method is flexible to take into account more complex relationships between the two codes in higher dimensional problems.

ACKNOWLEDGMENTS

This work is supported by the European Metrology Research Programm (EMRP), which is jointly funded by the EMRP participating countries within EURAMET and the European Union.

REFERENCES

[1] T J Santner, B J, Williams and W I Notz, *The design and analysis of computer experiments*, Springer Verlag, Springer Series in Statistics, 2003. 234, 2003.

[2] C. E Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006.

[3] S. Xiong, P.Z.G. Qian, and J. Wu. “Sequential design and analysis of high-accuracy and low-accuracy computer codes”. *Technometrics*, 55:37–46, 2013.

[4] M. C. Kennedy and A. O’Hagan. “Predicting the output from a complex computer code when fast approximations are available”. *Biometrika*, 87:1–13, 1998.

[5] Chevalier, C, Picheny, V, Ginsbourger, D, The KrigInv package : An efficient and user-friendly R implementation of Kriging-based inversion algorithms, Computational Statistics and Data Analysis, 71, 1021-1034, 2014.

[6] J. Bect, D. Ginsbourger, L. L, V. Picheny, and E. Vazquez. “Sequential design of computer experiments for the estimation of a probability of failure”. *Statistics and Computing*, vol 22, pp. 773–793, 2012.

[7] Loic Le Gratiet. MuFiCokriging: Multi-Fidelity Cokriging models, 2012. R package version 1.2.

[8] J. Oakley and A.O’Hagan. “Bayesian inference for the uncertainty distribution of computer model outputs”. *Biometrika*, 89:769–784, 2002.

[9] O. Roustant, D. Ginsbourger, and Y. Deville. “DiceKriging, DiceOptim: Two R packages for the analysis of computer experiments by kriging-based metamodeling and optimization”. *Journal of Statistical Software*, vol 51, nº. 1, pp.1–55, 2012.

[10] D. A. Purser. ‘Assessment of hazard to occupants from smoke, toxic gases and heat’. In SFPE handbook of fire protection engineering, pages 2.96–2.193. 4th edition

[11] NF EN ISO 13571 - Guidelines for the estimation of time to compromised tenability in fires. volume ICS: 13.220.01. 2012.

[12] K. McGrattan, R. McDermott, C. Weinschenk, and K. Overholt. Fire dynamics simulator – technical reference guide, volume 1 mathematical model. NIST SP - 1018, page 175 pp, 2014.