Design and Analysis of Computer Experiments for Bulk Acoustic Wave filters: Comparison of several types of Designs and Comparison of Kriging versus Pseudo-cubic Thin-plate type Spline as Metamodel

François de Crécy[a], Nicolas Durrande[b], Alexandre Reinhardt[a], Sylvain Joblot[c], Céline Helbert[b]
[a]: *CEA, LETI, Minatec, 17 rue des Martyrs, 38054 Grenoble, France*
[b]: *Ecole Nationale Supérieure des Mines de St Etienne, 158, cours Fauriel 42023 St Etienne, France*
[c]: *ST Microelectronics, 850 rue Jean Monnet, 38920 Crolles, France*
mailing address: francois.decrecy@cea.fr

Abstract
In the field of Bulk Acoustic Wave conception, software simulators are used to simulate their electrical and mechanical behaviors. As these simulations are time consuming for Monte Carlo approaches used in yield studies, we need a metamodel instead of the true simulators. We tested three types of Design of Experiments (DOE) (10 factors, 1003 simulations) and three types of metamodels (Ordinary and Universal Kriging, Pseudo-cubic Thin-plate type interpolating Spline) on 4 independent responses. The different approaches are validated on two sets of test points. In our case, the Maximin Latin Hypercube Sampling (LHS) space filling DOE is often the best, and the splines usually lead to smaller Mean Square Error (MSE) on the test points, but suffer from the lack statistical information (no uncertainty around the best estimate value).

1. What is a Bulk Acoustic Wave filter?
Bulk Acoustic Wave (BAW) filters appeared at the beginning of the 1980's as a result of trials for increasing the frequency of quartz resonators [1,2]. Their principle is to exploit the mechanical resonance of the thickness vibration of a piezoelectric layer. Through the capability of these materials to convert mechanical energy into electrical energy, or conversely, the mechanical vibration is excited by an applied radiofrequency signal and is translated into an electrical resonance suitable for the synthesis of low loss and high selectivity band pass filters. Because such devices are fully passive and compact, they find their application in wireless front-ends, for example in mobile communication systems such as GSM or UMTS, where they are used to select one standard in the signal received by the antenna, or transmitted to it. In so-called "full duplex" architectures, as for the UMTS system, transmission and reception need to be simultaneous: therefore the reception filter needs to eliminate all of the unwanted part of the spectrum, including the transmitted signal, a part of which is often redirected towards the reception chain because both reception and transmission chains are connected to the same antenna. This puts a constraint of high selectivity to these filters, and makes them unable to be designed using most of the alternative technologies (lumped elements for example).

Technologically, BAW resonators appear as thin (typically around 1 µm thick) piezoelectric films sandwiched between two electrodes. To avoid leakage of acoustic waves outside the resonator and inside the supporting substrate, two major technologies have been established: one which consists of having the piezoelectric film self standing over an air gap [3], in which acoustic waves cannot be transmitted, or another which makes use of an acoustic Bragg mirror made of layers of alternatively dense and light materials [4] (usually, respectively tungsten and silicon dioxide [5]). Because BAW devices exploit a thickness vibration, the overall thickness of the component is directly related to the final frequency obtained. Thus, to reach the frequency accuracy needed to match specifications, the thickness of the whole material stack needs to be controlled with an extreme precision (up to a few nm for a material stack of a few microns thick [6]). Post-processing techniques have been developed and are now used in industry, but their correction capabilities are still limited and filter responses cannot be perfectly adjusted to compensate for process deviations[6]. This is why, before entering mass production, it is necessary to estimate *a priori* a fabrication yield given

known statistical process dispersions and an evaluation of their impact on the final product.

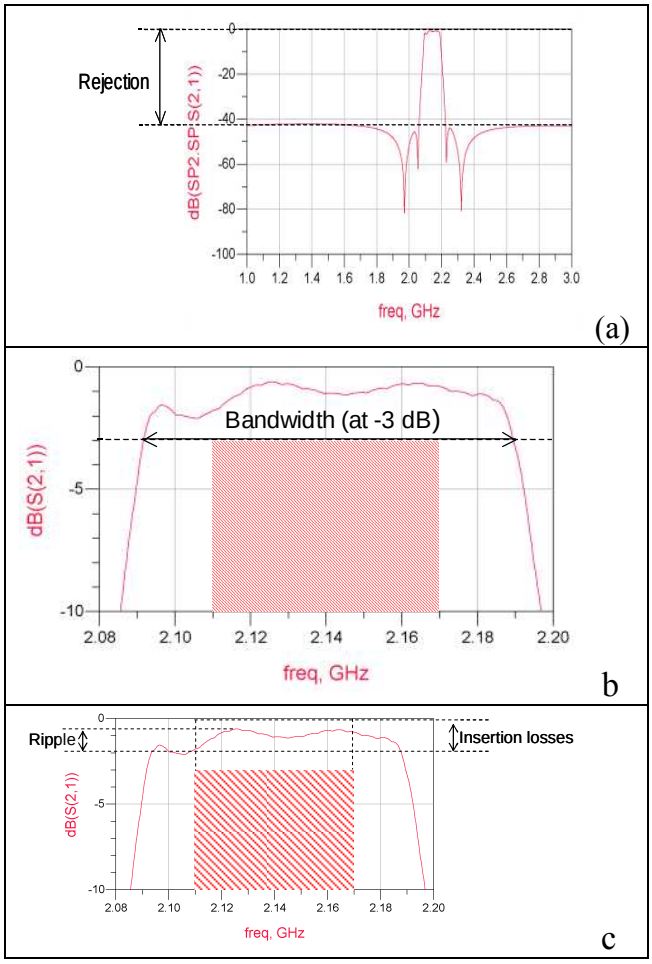


Figure 1 : Illustration of parameters used to evaluate a filter response.

A typical filter response is shown in Fig. 1(b). As explained in the previous paragraph, a filter needs to isolate a part of the spectrum which depends on the exact standard for which it is designed. Therefore, it needs to be centred around this part of the spectrum and to be wide enough to contain it. When comparing two filters, it is often useful to determine their respective **centre frequencies** and **bandwidths**. These quantities are defined with respect to the frequency band for which the transmission of the filter is higher than a given threshold, usually –3 dB. The centre frequency is the middle, whereas the bandwidth is the width of this frequency band, as indicated in Fig. 1(b). Within the frequency band which needs to be selected, one defines the **insertion losses** corresponding to the minimum transmission level inside the targeted frequency band and the **ripple**, which corresponds to the difference between the highest and the lowest transmission levels. These quantities are illustrated in Fig. 1(c). An alternate way of describing a filter is to study not only its transmission spectrum, but also the reflection of electrical waves at its input: the better the filter is electrically matched to its environment, the more electrical power can be transmitted through it instead of being reflected back to the signal source (antenna or power amplifier). So, it may be interesting to study what is called the **adaptation**, which corresponds to the reflection coefficient of electrical waves at the input of the filter, and can often be directly related to insertion losses. Usually, this quantity needs to be as small as possible and specifications provide the maximum level permitted. Outside the frequency band selected by the filter, one evaluates the **rejection** of the signal, which corresponds to the direct attenuation of a signal received outside the frequency band. In industrial specifications, this figure is split among different values for alternate frequency bands, but for the sake of simplicity, we will only consider a general value, as shown in Fig. 1(a).

Why do we need a metamodel?

The final objective of this study is the estimation of a fabrication yield in an industrial context, using a Monte-Carlo approach. The computation time of our simulator forbids us to use it during industrial Monte-Carlo, and it is necessary to replace it by a fast-computing metamodel. As our simulator is perfectly deterministic, we need an interpolating metamodel. It is reasonable to compute around one thousand simulations for building it. We use dimensionless thicknesses of the ten layers as input factors, i.e. the difference between actual and nominal thickness divided by the standard deviation of this thickness as expected from the fabrication process capabilities (process dispersion). Our previous experience indicates that the five first layers are more influent that the five last (Bragg mirror).

2. Three types of DOE

To build the metamodel, we use three different DOE, each one with 1003 simulations.

The first DOE is an interweaving of different "classical" sub-DOE, composed of:

1. the central point,
2. a fractional factorial design $2^{\wedge(10-3)}$ at scale 0.75
3. a fractional factorial design $2^{\wedge(10-3)}$ at scale 1.50, expanded from the new points of a foldover of the previous sub-DOE 2 ,
4. a fractional factorial design $2^{\wedge(10-3)}$ at scale 2.25, the Alias Generator being different from sub-DOE 2 ,
5. a fractional factorial design $2^{\wedge(10-3)}$ at scale 3.00, expanded from the new points of a foldover of the previous sub-DOE 4 ,
6. 10 series of stars points, i.e. point with all factors at 0 (nominal thickness) except the j^{th} factor that varies with step 0.25, from -3.25 to +3.25 for the 1^{st} to 5^{th} layer, from -3.00 to +3.00 for the 6^{th} to 10^{th} layer , *(these series were also performed to facilitate visualizations of 1D effects)*
7. A Box-Behnken DOE at scale 1.50, without central point.
8. 5 series of fractional factorial design $2^{\wedge(5-1)}$ at scale 1.00 for the five first factors except the j^{th} factors at scale 2.5 (j successively from 1 to 5), the five last being at 0.00.

In fact, this DOE emphasizes more the most external regions than the two other DOE.

The second DOE is a space filling design based on a MaxiMin Latin Hypercube Sampling (LHS). A latin hypercupe sampling [7] is a design where the points are all distinct after been projected on each axis. Thus, the chosen values for input parameters are all different. This kind of design is interesting in computer experiments where the result is determinist, i.e. there is no white noise usually present in physical experiments. Thus, there is no need to repeat two simulations at the same value. However among latin hypercubes sampling, the points are not necessarily well distributed into space. Therefore in order to construct a space filling, we add the "maximin" criterion to the construction. The design is searched so that the minimum distance between two sampling points is maximized. This design has been generated in $[0,1]^{10}$ by the "lhsdesign" Matlab procedure of the statistics toolbox

The third DOE is also a space filling design, but based on a Halton's low discrepancy sequence. The discrepancy is a mesure of the distance between empirical and uniform distributions. The d dimensional Halton sequence is the generalisation of the Van der Corput sequence which is the basic one dimensional low discrepancy sequence. The construction is based on the schema that each point is repelled from the others (see [8] for details). In our application the sequence has been obtained in $[0,1]^{10}$ by the R software.

To get locally accurate metamodels in the center of the domain (which is the region of highest interest) for the two space filling designs we use a continuous transformation toward the $[-4; +4]^{10}$ region of $[0,1]^{10}$ that leads to a higher point density in this central region. Naturally, the inversed cumulative standard normal distribution (noted ICSND on Figure 2) should be applied to generate a normal Design from uniform LHS or a Halton sequence. However, another function has been used (ICSNDo**g**) in order to little reduce the concentration of points in the central region (see Figure 2 for a comparison between ICSND and the chosen function).

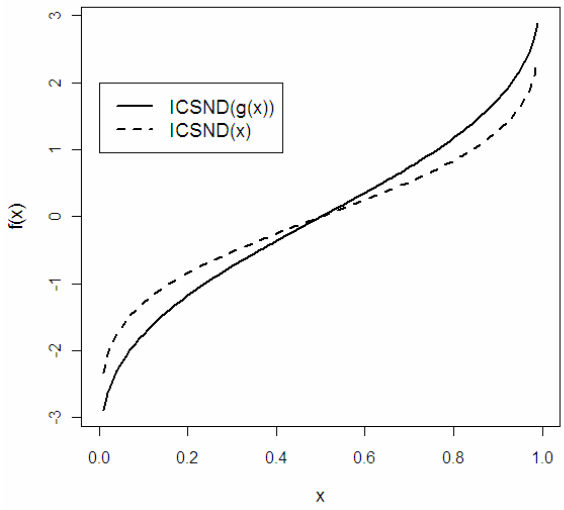


Figure 2 Transformations from [0,1] toward [-4,4]

The quality of prediction was tested using two independent sets of 500 test points each. The

first set is normally distributed over $\mathbb{R}^{10}$ (mean 0, unity covariance matrix), the second set is uniformly distributed into the $[-3 ; +3]^{10}$ region of $\mathbb{R}^{10}$. Of course, these test simulations are never used to determine the metamodels. The purpose of these two different test sets is to be able to test independently the quality of metamodels focusing either on the most probable region (normally distributed) or on the full range of interest (uniformly distributed).

When analyzing the 5 responses, it immediately appears that "insertion losses" and "ripple" are highly linearly correlated. It is not worthwhile to keep these two responses and we choose "ripple".

"Ripple" and "adaptation" exhibit large and steep variations in a small range of parameters, in the most external region of $[-3 ; +3]^{10}$, with very large ripple and very low adaptation. This is due to some quasi-discontinuities in the BAW behavior and will cause some difficulties. In some cases, we built a metamodel without these **"special points"** or used the uniformly distributed test set without these "special points" *(there is no "special points" in the normally distributed test set)*.

3. Three types of Metamodels

We use three types of metamodels: Ordinary Kriging (OK) and Universal Kriging (UK) with Gaussian type covariance function and Pseudo-cubic Thin-plate type interpolating multidimensional Splines.

3.3.1 Kriging

This section describes OK and UK [9]. OK assume that the function y is a realization of a Gaussian process $\left(Y(x)\right)_{x\in D}$ such as $E[Y(x)]=\mu$ and $Cov[Y(x^{(1)}),Y(x^{(2)})]=k(x^{(1)},x^{(2)})$, where μ is the mean of the process and k the kernel that stands for the covariance structure. Let $\mathrm{X}=\left(x_1 \ \ldots \ x_n\right)^T$ be the matrix of experiments and $\mathrm{Y_{obs}}=\left(y_1 \ \ldots \ y_n\right)^T$ the vector of the observations at the design points. Then, the kriging mean is the best linear unbiased predictor in the sense of the mean squares.

For our study, we consider the Gaussian kernel

$$k(x^{(1)},x^{(2)})=\sigma^2\exp\left(-\sum_{i=1}^{k}\left(\frac{(x_i^{(1)}-x_i^{(2)})}{\theta_i}\right)^2\right).$$

The parameter σ^2 is the process variance and θ is a vector called the range parameter that adjusts the smoothness of the response in each direction. Since the parameters μ, σ and θ are unknown, they are determined using the maximum likelihood estimation. Let $\hat{\mu}$, $\hat{\sigma}$ and $\hat{\theta}$ be their estimators. Then we have the following results:

$$\hat{\mu}=\frac{\mathbf{1}^T K^{-1}Y_{obs}}{\mathbf{1}^T K^{-1}\mathbf{1}}$$

$$\hat{\sigma}^2=\frac{\left(Y_{obs}-\mathbf{1}\hat{\mu}\right)^T R^{-1}\left(Y_{obs}-\mathbf{1}\hat{\mu}\right)}{n}$$

$$\hat{\theta}=\arg\min\left(\frac{n}{2}+\frac{1}{2}\log\left(\det\left(K\right)\right)\right.$$

where $k(x)=(k(x,x_1)\ \ k(x,x_2)\ \ \ldots\ \ k(x,x_N))^T$, $K=(k(x_i,x_j))_{1\le i,j\le n}$, where R is the correlation matrix ($\mathrm{R}=\mathrm{K}/\sigma^2$) and $\mathbf{1}=(1\ 1\ 1\ \ldots\ 1)^T$.

The matrix K is symmetric, real and invertible but it is sometimes necessary to add a little value on the diagonal in order to prevent numerical instability. This transformation, quite common in kriging, is called adding a nugget effect. In our study, we had to use this method to make easier the inversion. Finally, the new covariance matrix is $K_N=K+5.10^{-3}.Id$.

This transformation implies that the kriging mean does not interpolate anymore since the variance at a training point is not equal to zero (but in our case, $(5.10^{-3})^2$ is an upper bound).

The expressions of kriging mean and kriging variance are then the following:

$$m_{ok}(x)=\hat{\mu}+k^T(x)K^{-1}\mathrm{Y_{obs}},$$

$$s_{ok}^2(x)=k(x,x)-k^T(x)K^{-1}k(x)+$$
$$(1\text{-}k^T(x)K^{-1}\mathbf{1})(\mathbf{1}^TK^{-1}\mathbf{1})^{-1}(1\text{-}k^T(x)K^{-1}\mathbf{1})^T \quad ,$$

where $\hat{\sigma}$ and $\hat{\theta}$ are plugged into K and $k\left(x_1,x_2\right)$.

OK is a useful tool, but it assumes that $E[Y(x)]=\mu$ for all $x\in D$, which may not be realistic in many cases. In order to free of this

constraint, one may consider that $\left(Y(x)\right)_{x\in D}$ is composed of the sum of a trend and a Gaussian process and thus consider UK. The assumptions of UK are:

$$E[Y(x)] = f(x)\beta$$

$$Cov[Y(x^{(1)}), Y(x^{(2)})] = k(x^{(1)}, x^{(2)})$$

where $f(x) = \left(f_0(x) \ \ldots \ f_p(x)\right)$ is a known trend vector, $\beta = \left(\beta_0 \ \ldots \ \beta_p\right)^T$ is the vector of trend parameters. Under those assumptions $\hat{\beta}$, $\hat{\sigma}$ and $\hat{\theta}$ become:

$$\hat{\beta} = \left(\mathrm{F}^T K^{-1}\mathrm{F}\right)^{-1} \mathrm{F}^T K^{-1}\mathrm{Y}_{\mathrm{obs}}$$

$$\hat{\sigma}^2 = \frac{\left(\mathrm{Y}_{\mathrm{obs}} - \mathrm{F}\hat{\beta}\right)^T R^{-1}\left(\mathrm{Y}_{\mathrm{obs}} - \mathrm{F}\hat{\beta}\right)}{n}$$

$$\hat{\theta} = \arg\min\left(\frac{n}{2} + \frac{1}{2}\log\left(\det(K)\right)\right),$$

where $\mathrm{F} = (f_j(x_i))_{i,j}$ avec $1 \le i \le n$, $1 \le j \le p$.

Finally, the kriging mean and the kriging variance are:

$$m_{uk}(x) = f(x)\hat{\beta} + k^T(x)K^{-1}\left(\mathrm{Y}_{\mathrm{obs}} - \mathrm{F}\hat{\beta}\right), \text{ and}$$

$$s_{uk}(x) = k(x,x) - k^T(x)K^{-1}k(x) +$$
$$(f(\mathrm{x}) - k^T(\mathrm{x})K^{-1}\mathrm{F})(\mathrm{F}^T K^{-1}\mathrm{F})^{-1}(f(\mathrm{x}) - k^T(\mathrm{x})K^{-1}\mathrm{F})^T$$

One can notice that OK is a particular case of UK considering $f(x) = 1$.

The probabilistic interpretation of these results can be obtained by a Bayesian approach (see Helbert [10])

3.3.2 Pseudo-cubic Thin-plate type Spline

The Pseudo-cubic Thin-plate type multidimensional Spline is a family of spline proposed by Duchon [11], and it may be smoothing or interpolating spline. In its interpolating version, it is the function f(**t**) defined over $\mathbb{R}^d$ which goes through all the design points exactly and which minimizes a curvature energy defined by:

$$E_c(f) = \int_{R^d} \sum_{k=1}^{d}\sum_{p=1}^{d} \left[Four\left(\frac{\partial^2 f}{\partial t_k \partial t_p}\right)(u)\right]^2 \|u\|^{d-1} du_1\, du_2 \ldots . du_d$$

where $Four\left(\frac{\partial^2 f}{\partial t_k \partial t_p}\right)$ is the Fourier transform of a second derivative of the function f(t) and u is the variable in the frequency space. In its smoothing version, it is the function which minimizes a total energy defined as :

$$E_{total} = E_c(f) + \rho * \sum_{i=1}^{n} \omega_i \left(y_i - f(t_i)\right)^2$$

where n is the number of points, y_i the value of the response for the i^{th} point, ω_i is the weight associated with this i^{th} point *(in our case, ω_i =1 for any points)*, t_i is the vector of the values of the d factors at this i^{th} point.

The last term is a residual variance multiplied by the positive parameter ρ, inverse of the smoothing parameter. In our interpolating case, ρ = ∞ and the residual variance vanishes. In this energy analogy, this residual variance could be seen as the energy of n springs connecting the points to the surface (the plate). The smoothing parameter characterizes the relative stiffness of the plate and the springs.

The expression of the spline is very simple:

$$f(t) = \sum_{i=1}^{n} \lambda_i \, H(t, t_i) + \sum_{k=1}^{d} \alpha_k t^{(k)} + \alpha_o$$

with $H(t,t_i) = \left(\sum_{k=1}^{d} dil_k^2 \left(\frac{t^{(k)} - t_i^{(k)}}{\sigma_k} \right)^2 \right)^{3/2}$ is a radial basis function, $t^{(k)}$ is the k-th component of t, σ_k is the standard deviation of the k-th variable, and dil_k is the scale dilatation of the k-th variable (the product of all the scale dilatations is imposed to 1.0), which plays a similar role to scale parameter in kriging, n is the number of points in the learning base, and d the dimension of space (d=10 in this study). There is no constraint on the position of the design point in the parameter range, and the spline is defined over the full $\mathbb{R}^d$. Any extrapolation becomes rapidly very close to a local hyperplane.

It should be noticed that while spline is based on energetic considerations and kriging is more based on statistic considerations, the two overall mathematical objects are not very different: both are based on radial basis function.

The coefficients λ_i and α_j are the solutions of the n+d+1 symmetric linear system:

$$\sum_{j=1}^{n} \lambda_j H(t_i, t_j) + \frac{\lambda_i}{\rho\, \omega_i} + \sum_{k=1}^{d} \alpha_k \frac{dil_k}{\sigma_k} t_i^{(k)} + \alpha_o = y_i \quad \text{for i=1, …, n}$$

$$\sum_{j=1}^{n} \lambda_j \frac{dil_k}{\sigma_k} t_j^{(k)} = 0 \quad \text{for k=1, … , d}$$

$$\sum_{j=1}^{n} \lambda_j = 0$$

With this type of spline, there is no constraint on the repartition of the learning points in $\mathbb{R}^d$, and the linear system to solve is usually very stable and well conditioned. We never have to use the equivalent of kriging's nuggets (for other studies, splines for up to 3000 design points in a 17 dimension space have no problems).

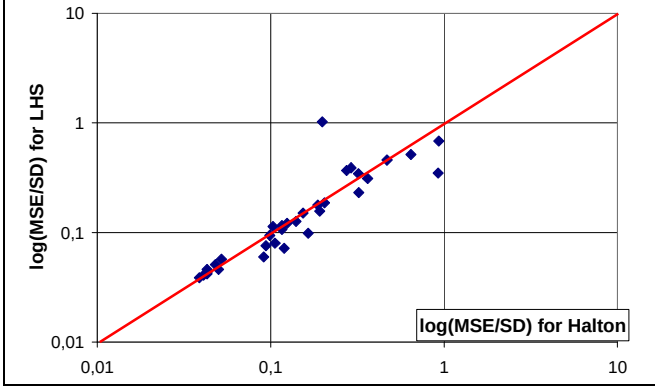
The choice of the d scale dilatations is inspired by BootStrap: a partition of the n learning points is randomly made into Q subsets (typically Q from 10 to 20), and for each subset, the spline is computed without the points of this subset and used only on this subset to calculate the mean square difference between predicted and actual values. The scale dilatations are iteratively determined to minimize the sum of all mean square differences for all the Q subsets. To enhance the convergence of these iterations, it is convenient to follow the path determined by the same mean square second derivatives with respect to dimensionless factors $t^{+(k)}$ defined by : $t^{+(k)} = \frac{dil_k}{\sigma_k} t^{(k)}$ The scale dilatations are chosen such that the sum of the squares of second dimensionless derivatives on the whole set of learning points, $\sum_{i=1}^{n} \left(\frac{\partial^2 f}{\partial t^{+(k)\,2}} (t_i) \right)^2$, tends to have the same value whatever the factor k.

4. Results of the comparison

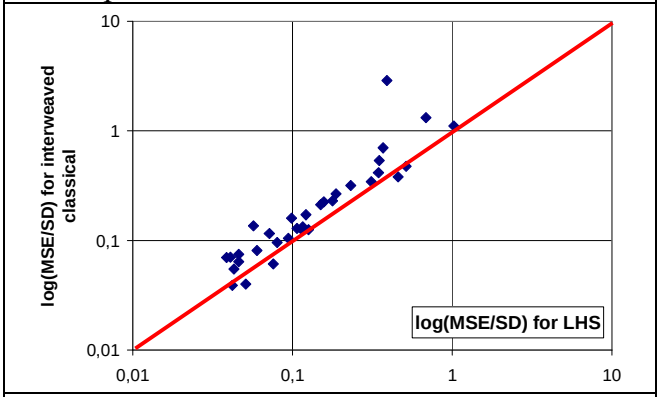
4.1 MSE as one of the criteria of comparison

For each DOE (with or without "special" points), for each response, for each type of metamodel, and for the two test sets, we compute the MSE, Mean Square Error, as: $MSE = \sqrt{\frac{\sum_{i=1}^{m} (y_i - f(t_i))^2}{m}}$, where the summation is done over the m=500 test points. This MSE, divided by the standard deviation (SD) of the response on this test set is used as one of the criteria to compare the DOE and metamodels.

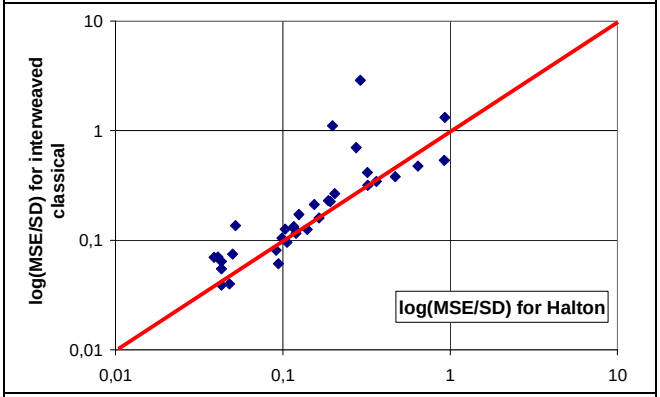
4.2 Comparison of the DOE



a: comparison of maximin LHS vs Halton



b. Comparison of "interweaved classical" vs LHS



a. Comparison of "interweaved classical" vs Halton

Figure 2: Comparison of different DOE

Figure 2 compares the different DOE. Each point represents a combination of a response, a metamodel, a test set.

Fig. 2.a shows that MSE with LHS is generally lower than with Halton.

Fig. 2.b shows that MSE with LHS is generally lower than with "interweaved classical" DOE.

Fig. 2.c shows that MSE with Halton is generally lower than with "interweaved classical" DOE.
So, we can conclude that, in our case, Maximin LHS is better than Halton, and Halton is better than "interweaved classical" DOE.

4.3 Comparison of the metamodels

The comparison of ordinary and universal kriging is done only on the normally distributed test points and without the "special" points in the learning data sets. In our case (see figure 3), UK is always better than or equivalent to OK.

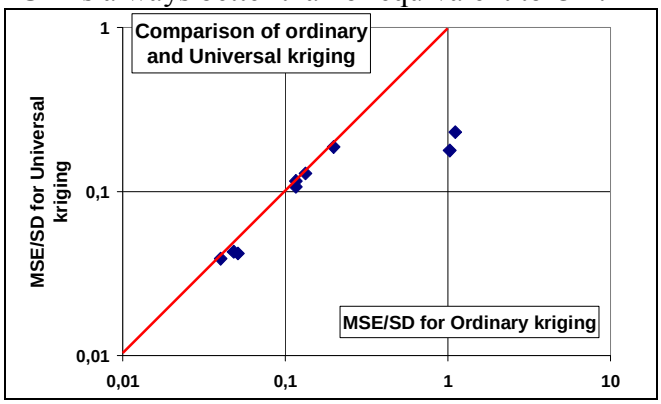


Figure 3: Comparison of OK and UK. *Each point represents a combination of a response and a DOE.*

The comparison of universal kriging and spline (see figure 4) is done on all the responses, DOE and test sets.

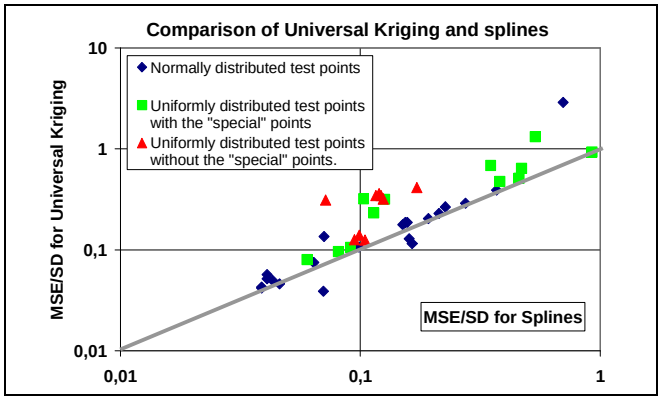


Figure 4: Comparison of Universal kriging and spline. *Each point represents a combination of a response, a test set and a DOE.*

On this figure 4, we separate the test sets. With normally distributed test points, UK and splines are more comparable, splines being often slightly better than UK. With uniformly distributed test points, splines are always better than UK and often comparable. This means that splines are more capable than UK to simulate the most external part of the $[-3; +3]^{10}$ region of $\mathbb{R}^{10}$, where normally distributed points are very rare.

4.4 Confidence intervals

The real advantage of kriging over splines is its capability of providing a confidence interval around the best estimate value. We tested this confidence interval with the UK based on maximin LHS for the adaptation response. For each point of the two test sets, but without the "special points", we compute the quantity:

$t_i = \frac{f(x_i) - y_i}{e_i}$, where e_i is the estimated uncertainty at the point x_i and y_i is the true value of the simulated response.

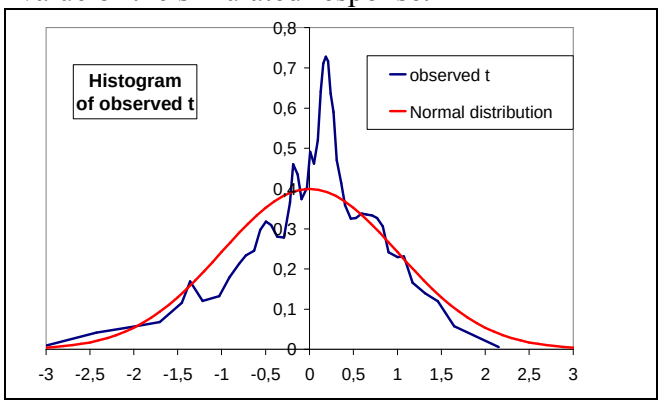


Figure 5: Comparison of observed t with a normal distribution.

Figure 5 shows that the distribution of observed t is in reasonable agreement with a normal distribution.
This estimated uncertainty can be used for Monte-Carlo yield estimations. In fact, the acceptance criterion for a BAW filter may be summed up as simultaneous conditions of type $R_k < T_k$ and/or $R_k > T'_k$ where R_k is one of the p responses (adaptation, ripple, centre frequency, bandwidth, insertion losses, …) and T_k and T'_k are fixed thresholds. At a given Monte-Carlo run x, the quantity $t_k(x) = \frac{R_k(x) - T_k}{e_k(x)}$ is transformed into a probability p_k of acceptation for this response and criterion via the cumulative distribution function of a normal distribution. The overall probability of acceptance of this Monte-Carlo point is the product of these probabilities p_k and the yield is estimated by:

$$yield = \frac{1}{N}\sum_{i=1}^{N}\left(\prod_{k=1}^{P} p_k(x_i)\right)$$

where the summation is done over the N runs x_i of Monte-Carlo with the appropriate random repartition.

5. Conclusion

In our case, space filling design, especially Maximin LHS, are better than "interweaved classical" designs and pseudo-cubic thin plate type interpolating splines are more precise (when tested with independent test points) and seem more stable than ordinary or universal kriging. Nevertheless, the confidence interval around the best estimate value of kriging may be an important advantage for many studies, as yield estimation.

Bibliography

[1] T.W. Grudkowski, J.F. Black and T.M. Reeder, *Fundamental mode VHF/UHF bulk acoustic wave resonators and filters on silicon*, in 1980 IEEE Ultrasonics Symposium Proceedings, pp. 829-833.

[2] K.M. Lakin and J.S. Wang, *UHF composite bulk wave resonators*, in 1980 IEEE Ultrasonics Symposium Proceedings, pp. 834-837

[3] H. Satoh, Y. Ebata, H. Suzuki and C. Narahara, *An air-gap type piezoelectric composite thin film resonator*, in Proceedings of the 39th Annual Frequency Control Symposium, pp. 361-366 (1985).

[4] K.M. Lakin, K.T. McCarron and R.E. Rose, *Solidly mounted resonators and filters*, in 1995 IEEE Ultrasonics Symposium Proceedings, pp. 905-908.

[5] D. Petit, N. Abelé, A. Volatier, A. Lefevre, P. Ancey, J-F Carpentier, *Temperature compensated bulk acoustic wave resonator and its predictive 1D acoustic tool for RF filtering*, in 2007 IEEE Ultrasonics Symposium Proceedings, pp. 1243-1246.

[6] A. Reinhardt, S. Joblot, N. Buffet, A. Shirakawa, J.B. David, G. Parat, M. Aïd and P. Ancey, *Simulation of BAW resonators frequency adjustment*, in 2007 IEEE Ultrasonics Symposium Proceedings.

[7] M. Stein, *Large Sample Properties of Simulations Using Latin Hypercube Sampling*, Technometrics. 29, 143–151 (1987).

[8] H. Niederreinter, *Random Number generation and Quasi-Monte Carlo Methods* Siam, (1992).

[9] T. J. Santner, B. J. Williams, W. I. Notz, *The Design and Analysis of Computer Experiments,* Springer, 2003

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

[11] J. Duchon, "*Splines minimizing rotation-invariant semi-norms in Sobolev space*", Lecture Notes in Mathematics, vol 571, pp85-100, 1977