BLOCKING IN DESIGN OF EXPERIMENTS

Ingrid KOVACS[1], Bianca FAGARAS[1], Anamaria OROS[1],
Marina Dana ȚOPA[1], Monica RAFAILA[2], Manuel HARRANT[2]
[1] *Technical University of Cluj-Napoca, Basic Electronics, Str. Baritiu nr 26-28, Cluj-Napoca*
[2] *Automotive Power, Infineon Technologies, Neubiberg, Germany*
i.kovacs@bel.utcluj.ro

Abstract: Real life electronic systems must perform within tight limits even if affected by uncertainties such as temperature and supply variations. Design of Experiments (DoE) plays an important role in improving the quality of a system. The present paper focuses on a robust design method based on one of the main principles governing DoE: blocking. The method is automated with the help of MATLAB and involves the following steps: classify the factors, design the experiment, measure, analyze and optimize the response. The method is illustrated on a low-pass filter. The experiment uses a saturated D-optimal design. The metamodel is obtained considering two noise factors, a blocking factor, a small number of runs and replications.

Keywords: *Design of Experiments, Blocking, Randomization, Low-Pass Filter, Real-Time Load Emulation*

I. INTRODUCTION

In the industry of electronics, the verification of a product is becoming time-consuming because of the fast growing complexity of the systems. Due to the expense of electronic systems' fabrication, only a limited amount of experiments can be performed. The difficulty lies in finding a solution to lower product cost and shorten product design and development time. To improve the quality of a product, among the known factors, one should take into consideration several sources of variation that are present in the environment. To estimate this variation noise measurements are necessary. The challenge for engineers is to improve the system by reducing the influence of these sources.

We believe that experiments done on the product have a major role in improving the quality. Using a proper design can increase the efficiency of an experiment and strengthen the obtained conclusions. Also, an optimum design can decrease the number of experimental runs, therefore, costs are reduced.

Design of experiment (DoE) is a concept that can be used to build the design. Using DoE a set of experiments are planned and performed in order to obtain statistical information about the impact of the factors (and their interactions) on the output response [1]. Design of experiments has been successfully applied in many domains as chemistry, industry, agriculture, medicine. In the industry of electronics, computer experiment has been applied later [2]. Nowadays, DoE is applied on real measurements of electronic systems and principles like blocking, randomization and noise treatments are issues which come up in these real-life experiments.

The work presented in this paper focuses on treating variations differently depending on how they occur during product testing and during product use. Experiments were optimally planned and analyzed to get, from a reduced number of tests, estimates for how robust the system is with respect to undesired variations, and how they can be compensated.

The paper is organized as follows: section II depicts the state of the art, section III presents some general concepts and strategies related to DoE and section IV describes the implementation. The results are discussed in section V and in section VI conclusions are drawn.

II. STATE OF THE ART

Sir Ronald A. Fisher, Roger E. Kirk and Douglas C. Montgomery had seminal ideas for experimental design and have popularized their principles in some pioneering works. Sir Ronald A. Fisher has studied the experimental design in the early '20. He stated that experiments should be "carefully planned in advance and designed to form a secure basis of new knowledge" [3]. Fisher introduced statistical principles into experimental investigations, including the factorial design concept. He defined three basic principles for a good experimental design i.e. randomization, replication and blocking. He referred to randomization as the allocation of the experimental factors, in a random way, to ensure that the error effects are statistically independent. Replication was used as an independent repeat of each factor combination and that allowed the experimenter to obtain estimation of measurements error. Blocking isolated the variation that appeared.

Douglas C. Montgomery highlighted in his work [1] the importance of the experimental design in the engineering. He emphasized that the experimental design should be used at the beginning of each product cycle. He also believed that this approach would lead to a good performance of the processes and products, because of the reduction in the development time and cost. As [3], he advised that principles like replication, randomization and blocking are to be used in every experiment. He undelined the fact that replication is not repeated measurements. He stated that randomization was the allocation of the individual experimental runs in a random order and underlined the fact

Manuscript received November 5, 2012; revised March 12, 2013

that complete randomization of factors was almost impossible. Regarding blocking, he affirmed that it can be used to reduce and eliminate the variability that was transmitted by the nuisance factors.

According to Roger E. Kirk [4], the experimental design was a plan in which the experimental units are assigned to treatment levels. The design of experiment involved a number of inter-related activities and indicated the way in which the randomization and statistical aspects of an experiment are to be performed. He also claimed that the main goal of an experimental design was to determine a causal connection between the dependent and independent variables.

III. CONCEPTS

An experiment is a test, or a series of tests, in which the input variables of a process are varied in order to observe and identify the changes of the output response. A classical experiment may be summarized in some main steps: determine the experiment's objective, select the process factors and define the response, plan experiments and execute them, analyze results and draw conclusions. Experiments lead to a design. Literature defines different strategies to design, including best-guess-approach, one-factor-at-a-time or factorial experiments approach. The main types of designs used in an experiment are: central composite design (CCD), Latin hypercube sampling (LHS), full factorial (FF), D-optimal. Table 1 illustrates some basic characteristics of these designs.

Design	Advantages	Disadvantages
FF	Contains all combinations of factors level	Increase of the design space
CCD	Builds a quadratic model of the response, without needing to use a three-level factorial experiment	Unfeasible Design. Regions of interest at particular points cannot be represented.
D-optimal	A computer algorithm chooses the optimal set of design runs	Significant data might be lost, but validation is performed
LHS	For an N point design, projects onto N different levels in each factor	Set of points generated randomly. The one that best satisfies the criteria is chosen.

Table 1: Designs for planning experiments

Experimental designs can be improved using principles like replication, randomization and blocking.

- **Replication** refers to repeating a study using the same methods but with different subjects and experimenters. It offers a more precise estimate of the treatment effects and obtains an estimate of the error effect. The purpose of replication is to assure that the results are reliable and valid. As an example: consider an experiment to study the effect of hardening of oil and saltwater quenching on an aluminum alloy [1]. The objective is to determine which quenching solution produces the maximum hardness for this particular alloy. The replication in this case would consist of treating an alloy specimen by oil quenching and a specimen by saltwater quenching and measure the hardness of the specimens after quenching. Thus, the treatment of five specimens in each quenching medium would represent five replications.
- **Randomization** implies the random allocation of factors and/or the order of experiments.
- **Blocking** is a design technique which refers to arranging similar factors in groups. Other definitions are available in literature. We used blocking to isolate the variability appeared due to the nuisance factors and to find which factors may influence the experimental response, but are of little interest for the experimenter. Blocking improves the precision with which comparisons among the factors of interest are made. It also isolates the systematic effect and prevents obscuring the main effect. For example, let us take a wafer resistivity measurement in which we test the effect on resistivity after the diffusion process, considering three wafers and four material dosages. If the nuisance factor is the furnace run, this could be completely eliminated if we ran twelve wafers in the same furnace run, but unfortunately this is not allowed. A solution using the blocking principle would be to put four wafers with different dosages in each of the three furnace runs [5]. The randomization of this process would consist in the order in which the three sets are assigned to the furnace runs.

Two types of factors are involved in experiments: ***design factors***, i.e. factors selected for study and ***nuisance factors***, i.e. factors which do not present interest in the present experiment. Nuisance factors may be controllable, i.e. we may control their variation, or uncontrollable, i.e. parameters which cannot be controlled during product use, e.g. environment conditions. Table 2 contains their classification and how to deal with them when designing the experiment.

Characteristics	Examples	How to treat
Unknown, uncontrollable	Experimenter bias	Randomization
Known, uncontrollable but measurable	Weight, previous learning	Other concepts
Known, controllable	Temperature, batch, time, gender	Blocking

Table 2: Nuisance factors classification. Ways to treat factors in experimental design

The resulted data is used to develop an approximation model which links the outputs and the inputs. This form of approximation is usually found in literature under the name metamodel. A metamodel can be described in terms of regression analysis, from a mathematical point of view. Based on some simulation samples, a matching metamodel can be constructed and can be used later to predict response values at new factor settings. The construction of the metamodel is called fitting or learning. Let us consider a factorial experiment with two treatment factors and a blocking factor [1]; the linear statistical model for this design is described in (1). We also consider that a single replicate of a complete factorial experiment is run for each block.

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \delta_k + \varepsilon_{ijk}; \quad \begin{cases} i = 1,\ 2,\ ...,\ a \\ j = 1,\ 2,\ ...,\ b \\ k = 1,\ 2,\ ...,\ n \end{cases} \qquad (1)$$

where μ is the overall mean effect, τ_i is the effect of the i^{th} level of the first treatment factor, β_j is the effect of the j^{th} level of the second treatment factor, $(\tau\beta)_{ij}$ represents the interaction, while δ_k is the effect of the k^{th} block and ε_{ijk} is a random error component. The order in which the treatment combinations are run is completely randomized within a block.

The metamodel assumes that the blocking factors do not interact with the treatments. The error term includes the fitting residuals and the effects of the noise factors. The fitting residuals are caused by the terms not included in the interactions mentioned above. The effects of the noise factors are unknown and therefore not included in the metamodel. For a further use of the metamodel, to predict new values and optimize the system, a proper error analysis must be performed. After this check, the metamodel can be used to predict extreme values of the response, with respect to design and noise factors.

IV. IMPLEMENTATION

The concepts presented above must be refined considering the constraints in testing electronic control units for smart power automotive applications: medium to high voltages and currents, complex loads with feedback effects. Another important aspect is that different steps of such real-life experiments are sometimes separated by long time periods depending on the resources and on the project planning (it can take weeks, even months since a wafer lot is planned until it is really produced, and it can take similar time for hardware tests to be performed). These steps are sometimes performed by different teams in different locations. Last but not least, the solution must be automated and work properly no matter which is the device under test and the number/types of factors. For this reason MATLAB was chosen as the environment to design and analyze the experiment [2].

The following steps must be performed sequentially:

1. Factors classification: this is not always an easy task and must be performed using the system and domain knowledge. Experience shows the following:

- *Blocking factors* are known and controllable factors only for the sake of the experiment and are allowed to vary during product use. Most of the times they are operating conditions: supply voltage, ambient temperature, etc. These factors are not directly of interest, but have a big impact on the system, so experiments must account them and divide the measurements in blocks, in which they are held fixed. Temperature has another specific characteristic: it cannot be varied so easily from one level to another so batches of measurements with the same temperature must be performed together.
- *Noise factors* are factors which are known and uncontrollable, but measurable, such as tolerances of components. Their impact on the response is expected to be small and can sometimes be compensated by a specific setting of control design factors, which cancel their cumulated effect, by making use of the interaction with the noise factors.
- *Design, Control* factors are controllable factors which can be set by design and which give control over the responses, allow optimization, either of the nominal value or of the range of variation caused by statistical system variations. The unknown sources of noise, such as process or measurement noise are handled by replication.

2. Design the experiment: this step can be automated.

The experiments are generated on normalized factors at first, coded levels, i.e. [-1, 1] range, as Montgomery recommends [1]. The factors are specified by ranges, types and, if existing, by levels. The experiment is an *m* x *n* matrix, where *m* is the number of tests and *n* the number of factors. The experiments depend on the expected complexity of noise factors, the allowed number of runs as well as the number of factors. In our case, a 2-level fractional factorial which can estimate a metamodel with 2-factor interactions is used. The blocks are sets of levels for the blocking factors and, if the effects of the blocking factors must be separately extracted, then we choose a full factorial experiment. Any experimental design must take into account the metamodel which must be obtained in the end.

3. Measure

First the experiment must be translated into real factor levels. The central (nominal) point will be replicated to ensure that the noise of the measurement is small enough (compare replications with measurements with different factors).

4. Analyze

The metamodel must treat the factors' categories differently. The noise factors are expected to have small effects, but interactions between them can also occur, while the blocking factors have bigger effects, but additive, so the first assumption is of no interaction with the noise factors. This results into a set of effects corresponding to these assumptions on the metamodel, and which will be used to perform a regression on the results. The output is a vector of corresponding coefficients, which together with any set of normalized factors can estimate the value of the measured output.

The errors between the measurements and the predictions are used to evaluate if the metamodel is good enough. This is done by comparing the normed residuals to a predefined threshold, which is calculated as:

$$\max\left(normed\_residuals\right) < 0.1 \tag{2}$$

$$\frac{\max\left(abs\left(predicted\_response - measured\_response\right)\right)}{\max\left(measured\_response\right) - \min\left(measured\_response\right)} < 10\% \tag{3}$$

5. Optimize

It is important to approximate the extreme values of the response, by using its analytical form. Noise factors introduce small ranges of variation, while blocking factors will have a significant impact. But since the effect of the blocking factor is additive, the best/worst point for the response with respect to the noise factors will be the same, no matter which is the value of the blocking factor.

V. EXPERIMENTAL RESULTS

The methods presented in the previous section were applied on a low-pass filter. We want to estimate the robustness of the system with respect to the produced current. The applied voltage *V* is varied in a given range. The passive components *R*, *C* have deviations from the nominal values, as follows:

$$\begin{cases} R = R_0 + \Delta R \\ C = C_0 + \Delta C \end{cases} \tag{4}$$

Having these influence factors, the measured output is the maximum value of the current, which must be kept under a specified maximum value.

An experiment was planned and tested with an application emulation test system [11] used for considering application variances and parameter spread during post-silicon verification of automotive smart power products [11]. Statistical distributions for different factors (see Table 3) were executed in an automated way and the current consumption was monitored. The resistor and capacitor are listed as a noise factor because they are easily affected by their noise, and their interaction cannot be estimated.

Factor name	Factor type	Range
V (Supply voltage)	blocking	[8V; 16V]
R (Filter resistance)	noise	[9.5 kΩ ; 10.5 kΩ]
C (Filter capacitance)	noise	[9μF; 11μF]

Table 3: Classification of the factors for the experiment planning

For validation purposes, measurements were performed for a three-level full factorial, with replications. The replications were checked for low noise and then averaged for further analysis.

Based on the metamodel presented in (1) with two noise factors and one blocking factor, we can define a corresponding set of effects, for each coefficient and design a saturated D-optimal experiment [1], which can estimate all these coefficients with not more than six runs. MATLAB [10] offers functions as *cordexch*, *daugment*, which can create or augment D-optimal experiments.

Table 4 illustrates these design points considering the normalized values of the factors, i.e. -1 for the minimum values, 0 for nominal values and 1 for maximum. Thus, the values of the factors used for fitting the metamodel are scaled as:

$$normedV = \frac{V - \frac{V_{\max} + V_{\min}}{2}}{\frac{V_{\max} - V_{\min}}{2}} \tag{5}$$

where V_{max} and V_{min} denote the maximum and minimum voltage levels. The normalized values of the R and C factors are obtained similarly. The last column of Table 4 describes the measured current I.

Run	normed$V$	normed$R$	normed$C$	I [A]
1.	-1	-1	-1	0.45177
2.	-1	-1	1	0.90704
3.	-1	1	-1	0.48715
4.	0	0	0	0.68591
5.	1	-1	1	0.82178
6.	1	1	1	0.90595

Table 4: The saturated D-optimal experiment and the response

The next step is to extract a metamodel from the collected data using the regression. The fitted metamodel can be approximated analytically by a function as:

$$\begin{aligned} I &= f(V) + f(R,C) + err = \\ &= c_0 + c_V \cdot \text{normed}V + c_R \cdot \text{normed}R + \\ &+ c_C \cdot \text{normed}C + c_i \cdot \text{normed}R \cdot \text{normed}C \end{aligned} \tag{6}$$

where c_0 represents the free term coefficient, c_V, c_R, c_C are coefficients of the main effects and c_i represents the interaction coefficient.

In order to check if the metamodel is properly fit, a residuals analysis was performed. We compared the normed residuals to a normal distribution. The maximum residual is small enough i.e. less than 10%, and the distribution is approximated normal with mean 0. These values indicate the fact that we can use the metamodel further on.

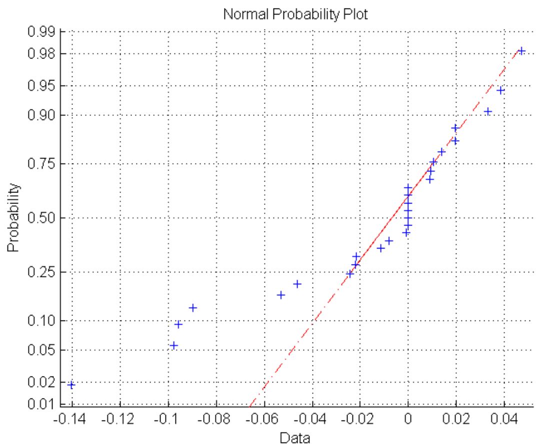

Figure 1: The normed residuals plot.

The metamodel can be visualized in one dimension in Figure 2, one factor being varied over the complete range and the others being kept to a fixed value. The plots were built in MATLAB using the *rstool* plot [1]. The dashed lines show the 95% confidence intervals for the response predictions. The factors were used in scaled values for consistency of the fit. As seen in Figure 2, the blocking factor that is the voltage is indeed important.

As previously stated the response studied can be approximated as a sum of two functions: one depending on the blocking factor (supply voltage) and the other one containing the noise factors (R and C) and their interactions. Considering the function (6), the current variation is described as:

$$\begin{aligned} \Delta I\,|_{R,C} &= I_{\max}(V)\,|_{R,C} - I_{\min}(V)\,|_{R,C} = \\ &= f_{\max}(R,C) + f(V) - f_{\min}(R,C) - f(V) = \\ &= c_R \cdot \Delta\text{normed}R + c_C \cdot \Delta\text{normed}C + \\ &+ c_i \cdot \Delta\text{normed}R \cdot \Delta\text{normed}C \end{aligned} \tag{7}$$

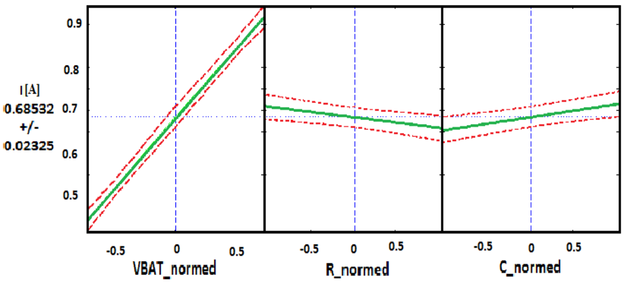

Figure 2: The metamodel using the normed values of the factors

Extreme values	normed V	normed R	normed C	Output Current I[A]
Minimum	-1	1	-1	0.48715
Maximum	1	-1	1	0.82178

Table 5: Factor values which determine the extreme values of the output current

where ΔnormedR and ΔnormedC represent the variations of the normalized R and C values. Table 5 shows the extreme values of the current response and the corresponding factor levels.

The minimum current is obtained for a maximum value of the resistance and minimum value of the capacitance. The robustness is measured by the maximum current variation for any value of the supply voltage $\Delta I_{\max} = 0.5394$ A .

The variations determined by the tolerances in the components have the maximum impact on the current which was computed. If this impact is considered too big we have two options for optimization. The first solution is to reduce the variation of the R and C factors in order to reduce the overall current variation. The second solution is to use the interaction between R and C on I to determine optimal, new nominal values which will compensate the overall current distribution. This is subject to further research.

VI. CONCLUSION

Electronic systems should behave the same no matter what variations appear during functioning. As the design space became n-dimensional, the verification is practically hard to cover. Planning key experiments is a valid solution for this particular problem.

The paper proposes a five step method that covers this solution. The method is automated with the help of MATLAB and involves the following steps: classify the, design the experiment, measure the response, analyze the response (metamodeling) and optimize the response.

The method was applied on a low-pass filter. The experiment was planned and tested on a Hardware Emulation System. Our purpose was to approximate the robustness of the system with respect to the consumed current when applied a certain voltage.

Normally, the first step is not an easy task, but can be performed if the system is known. The step that involved planning experiments was fulfilled using the concepts of DoE. Principles like blocking and randomization were used for refinement because of their capability to characterize what happens in real-life situations.

The experiment was conducted according to a D-optimal design with 6 design points. For validation purposes measurements were performed for a three-level full factorial with replications and involved 27 runs. There was one blocking factor (V) and two noise factors (R and C). The metamodel was obtained by regression on the D-optimal design and validated using residual analysis on all available measurements.

The one-dimension plots of the metamodel showed that the supply voltage V has an important impact on the response, being indeed a blocking factor. To decrease the variation of the response (I) and so the enlarge the robustness of the system there are two solutions: either reduce the variation of the R and C factors or use the interaction between R and C on I to determine optimal, new nominal values which will compensate the overall current distribution. This is subject to further research.

We treated the variations that occur on the real measurements. The variations were handled differently, depending on their impact on the output response. The resulting output was a metamodel and two solutions for optimization were proposed. The system's robustness with respect to the undesired variation was calculated. Using this electronic system we proved the applicability of the concept of DoE in real measurements.

REFERENCES

[1] D. Montgomery, *Design and Analysis of Experiments*, John Wiley & Sons, 2001.

[2] M. Rafaila, *Planning Experiments for the Validation of Electronic Control Units*, Doctoral Thesis, Faculty of Electrical Engineering and Information Technology, Vienna, 2010.

[3] A Fisher, *The Arrangement of Field Experiments*, Springer New York, 1992.

[4] R. E Kirk, *Experimental Design*, In R. Millsap & A. Maydeu-Olivares (Eds.), *Sage handbook of quantitative methods in psychology,* (pp. 23–45, 2009.

[5] R. A. Bailey, *Design of Comparative Experiments*, Cambridge University Press, Cambridge Series in Statistical and Probabilistic Mathematics Series, 2008.

[6] T. Donnelly, "Design of Experiments for Real-World Problems", *NDIA 26th Annual National Test & Evaluation Conference*, March 2, 2010.

[7] J. K. Teford, *A Brief Introduction to Design of Experiments*, Johns Hopkins University Applied Physics Laboratory, Laurel, United States 2007.

[8] G. W. Ochbert, *A First Course in Design and Analysis of Experiments*, 1st edition, W. H. Freeman, 2000.

[9] K. Hinkelmann, O. Kempthorne, *Design and Analysis of Experiments, Introduction of Experimental Design*, John Wiley & Sons, 2nd edition, 2008.

[10] MATLAB, Statistics Toolbox http://www.mathworks.com/help/toolbox/stats/.

[11] M. Harrant,T. Nirmaier,G. Pelz, F. Dona, C. Grimm, *Configurable load emulation using FPGA and power amplifiers for automotive power ICs,* Specification and Design Languages (FDL), 2012.