ANDRZEJ SKOWRONEK*

REDUCING THE NUMBER OF RUNS IN EXPERIMENTAL RESEARCH USING SMART DESIGNS OF EXPERIMENT

ZMNIEJSZANIE ILOŚCI BADAŃ EKSPERYMENTALNYCH Z ZASTOSOWANIEM INTELIGENTNYCH PLANÓW EKSPERYMENTU

Abstract

This paper presents the results of a computer simulation, performed to check how the number of smart designs of experiment units affects the quality of information to be obtained in experimental research conducted on the basis of the analyzed designs. In the simulation, a real research object was replaced by a special testing function whose values were compared to the values predicted by the neural networks trained with the use of data sets based on smart designs of experiments containing various numbers of units.

Keywords: *smart design of experiment, experimental research, neural approximation*

Streszczenie

W artykule przedstawiono wyniki symulacji komputerowej do oceny wpływu liczby układów elastycznych planów doświadczeń na jakość informacji, którą można uzyskać, wykonując badania eksperymentalne na podstawie analizowanych planów. W symulacji rzeczywisty obiekt badań został zastąpiony specjalną funkcją testową, której znane wartości były porównywane z aproksymowanymi wartościami zwracanymi przez sieć neuronową trenowaną z użyciem zbirów danych opartych na analizowanych planach doświadczeń o różnych wielkościach.

Słowa kluczowe: *elastyczny plan eksperymentu, badania eksperymentalne, aproksymacja neuronowa*

* Ph.D. Andrzej Skowronek, Faculty of Mechanical Engineering, Cracow University of Technology.

1. Introduction

In many areas of science, experimental research serves as the main source of obtaining knowledge and information. If it is necessary to limit the cost or time of an experiment, one can use special techniques offered by the theory of experiment, also known as the design of experiment methodology (DoE). Depending on the goal of research, various types of experimental designs can be used [1–3]. The application of the experimental designs with special techniques for the analysis of experiment's results can often facilitate reducing the size of the experiment (number of runs, observations, etc.) and obtaining relevant information from the research, without reducing its quality [4–5].

However, in the traditional designs of experiments, the researcher cannot change the number of design's units, the number of inputs' levels and has to carry out the experiment strictly according to the design which is used. Quite a different approach to the concept of experiment planning is applied in smart designs of experiments [6] which allow the researcher to set the number of design's units and the number of inputs' levels.

2. The idea of smart designs of experiments

Smart designs of experiment are generated in a dedicated computer application, based on three important principles: adaptation, randomness and equipartition [6–7]. The first principle means the possibility of adjusting the design's characteristics to the conditions of the experiment and characteristics of the analyzed object. The researcher is able, for example, to set the number of design's units and the number its levels for each input. The second principle means that smart designs are created in a non-deterministic manner: both the generation of input's levels and the selection of design's units are conducted with using pseudo-random numbers. However, there are some limitations put on the random way of generation of design's units:

– using a parameter called "important difference" (Δx), a minimal permissible distance between the currently generated value and the existing values of each input factor levels (Fig. 1),
– a parameter called "minimal Euclid's distance" (*esmin*) – it is Euclid's distance to the nearest "neighbour-unit" in the input's space, calculated for each design's unit, each unit must fulfill the condition: es $\geq$ *esmin* (Fig. 2)
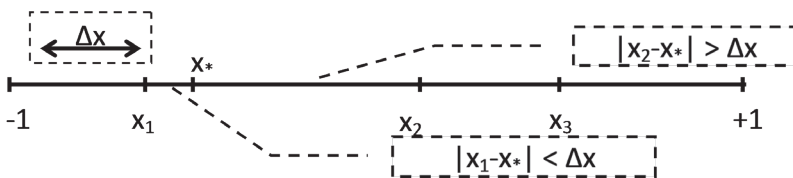

Fig. 1. Value *x** fails the important difference condition test and will be removed
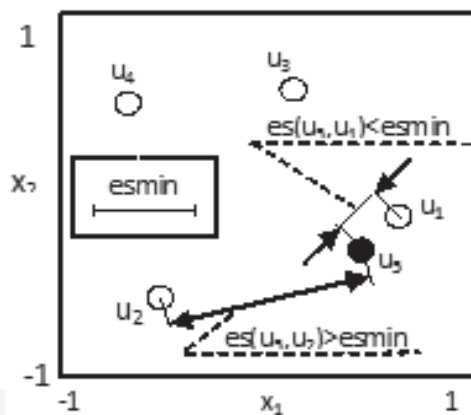
Fig. 2. $es(u_5, u_1) < esmin$, unit u_5 fails the $es \geq esmin$ condition test and will be removed

The conceptions of both parameters described above are based on the conception of Euclid's distance and they use the fact that a set of experimental design units in the space of inputs is equivalent to the set of points in the orthogonal coordinate system as well as the combinations of inputs' levels (which make up the units of designs) are equivalent to the points' coordinates. The Δx and *esmin* parameters support equipartition of the design's units in the inputs' space. If there are no other assumptions, design's units should cover regularly the whole inputs' space (the third rule). To estimate the regularity of the distribution of the design's units, the method of equipartitional analysis (EPA, [6]) is used. The analyzed (created) experimental design is compared with the master-design of the units which are distributed perfectly regularly in the inputs' space ([6–7], Fig. 3). The master-designs have the same numbers of inputs as the
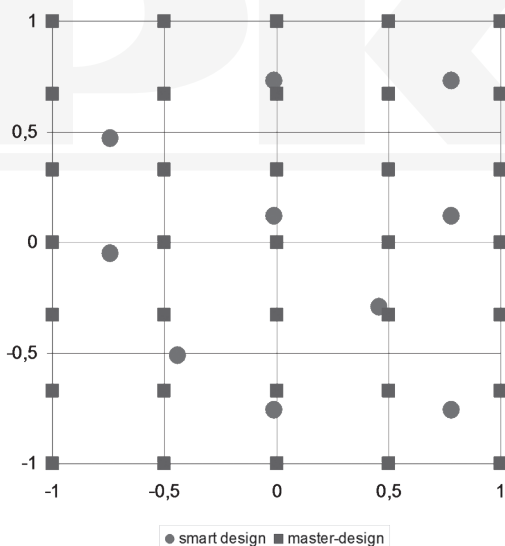

Fig. 3. 2-inputs master-design and smart design

analyzed designs and the same number of various inputs levels, but the number of design's units is often significantly higher and equal to the product of numbers of all input's levels. However, the levels of the master-design are calculated for each input by dividing the length of input range by the number of input's levels (Fig. 3). For each unit of the master-design, one can evaluate Euclid's distance to the nearest unit of the analyzed design. For such a collection (called an equipartitional set), one can evaluate a lot of statistical parameters (e.g. descriptive statistics [8]) or make one of the statistical tests [9]. Each of them could be an equipartition criterion in this analysis. Two parameters have been used: the maximal (*e1max*) and mean (*e1mean*) value of an equipartitional set. The *e1mean* parameter describes the central tendency of an equipartitional set and the *e1max* parameter gives the information whether there are some huge empty areas in the inputs' space (without design's units), which is important when taking into consideration the assumption that the design's units should cover the whole inputs' space. The dependence between both parameters and the design's quality (quality means equipartition, perfect regularity of the design's units in the inputs' space) was verified and described in [7]. The conclusion was: the smaller value of the equipartitional parameter, the more regular distribution of the design's units in inputs' space.

There are three ways to generate inputs' levels [10]. In the first method ("Z" method), inputs' levels are generated as pseudo-random values from the normalized range [−1, 1] and checked if they pass the important difference condition test (see Fig. 1). If a value fails the test, it is removed and the next one is generated to reach the right amount. In the "R" method, the levels of inputs are calculated by dividing the input's ranges by the demanded numbers of input's levels. The smallest level is calculated as the minimum of the input's range, whereas the biggest level is calculated as the maximum of the input's range. In the R2-method, the idea of levels calculation is that each level should be the center point of equal areas of influence. The first and the last levels are not equal to the minimum or maximum of the input's range (Table 1).

Table 1

Examples of levels in R and R2 methods

Method	Values used in the case of 5 levels				
R	−1	−0.5	0	0.5	1
R2	−0.8	−0.4	0	0.4	0.8

The generating of the smart designs of experiment consists of the following steps [7]:
- defining characteristics of the design: the number of inputs (factors), the number of designs units, the number of inputs' levels;
- generating the inputs' levels according to the method chosen;
- generating the sets of levels of inputs' factors;
- generating the set of all possible design's units by permuting all inputs' levels;
- completing the design by selecting from the set of all possible design's units only the ones which fulfill the *esmin* condition;
- equipartitional analysis to evaluate the quality of the design (quality means regular and equipartitional distribution of design's units in inputs' space).

The smart design's generator in the current version has implemented the functionalities which support the selection of the optimal values of important generation parameters – the important difference (Δx, used in Z-method of levels' generating) and the minimal Euclid's distance (*esmin*, used to ensure high regularity and equipartition of design's units in the inputs' space) [7]. Using previous versions of the generator, a researcher must set it by himself. If he does not know well the principles of how the smart designs generator functions, or he does not have some intuition or experience in the designs' generating process, it is likely that the generated design will not be optimal – designs' units will not cover equally the whole inputs' space. In the case of setting too small values of generation parameters, equipartition decreases or it is not possible to get the experimental design with the assumed number of units otherwise. To increase the probability of obtaining high-quality designs, they are generated in series and each design has to fulfill the *esmin* condition. If at least one design can be created, the *esmin* value is automatically increased and new designs are generated again. If any design is created, the *esmin* value is automatically decreased and designs are generated again. In the case of an increasing *esmin* value, if the newly generated design has a better quality than the previous one, the *esmin* value increases again and new designs are generated once again. If the new generated design is of worse quality than the previous one, the new design is not generated again and the previous design is saved. In the case of decreasing the *esmin* value, if at least one new design can be generated, new designs are not generated again. If any new design can be generated, the *esmin* value is automatically decreased again and new designs are generated once again. The design of better quality is selected on the basis of equipartitional parameters: *e1max* and *e1mean*.

The smart designs of experiment are multiple-generated [11]. The reason is the application of pseudo-random numbers in the algorithm of designs generating. Designs generated with the same seed of a pseudo-random number generator, the same parameters of generation (Δx, method of input's levels generating, *esmin*) and the same design's characteristic (the number of inputs, the number of input's levels, the number of design's units) will be identical. But if the seed value is changed or just if one tries to generate it next time even with the same generation parameters, they could be different and the difference of the design's quality could be sometimes significant. To avoid such problems, it seems necessary to generate several designs and choose one, based on EPA-parameters (*e1max* and *e1mean*). That is the idea of multi-generated smart designs of an experiment. The researcher can set a pseudo-random number generator seed value by oneself or can generate it, based on the real-time clock. To generate the identical design again, only its seed value must be known. The researcher can select the EPA-parameter with which he prefers to choose the best design. Each design generating is repeated up to 20 times to get 10 designs.

3. Computer Simulation

To study how the number of the smart designs of experiment units affects the quality of information to be obtained in experimental research which is conducted on the basis of the analyzed designs, a computer simulation was performed, where a real research object was simulated by a special testing function – Rosenbrock's Function [12]:

$$f(x) = \sum_{i=1}^{N-1}\left((1-x_i)^2 + 100 \cdot (x_{i+1} - x_i^2)^2\right) \tag{1}$$

The idea of using the testing function simulating the real object is easy to evaluate its values for each point in the input's space and compare them to the values predicted by neural networks trained with the use of data sets based on the smart designs of an experiment. It was assumed that the values predicted by the neural network will be significantly influenced by a training set which was built on the basis of the analyzed smart design. The comparison of the values predicted by the neural network with the values of the testing function calculated for a special testing set allow indirectly evaluate the quality of information gain in a computer simulation of an experiment conducted according to the smart design. By comparing the results obtained for the designs consisting of various numbers of units, one can study the impact of their size (number of units) on the simulation errors. It is very likely that the same effect could be observed in the real scientific experiment carried out with the use of smart designs.

Although many methods of approximation are known in the simulation, the neural approximation method was applied due to the easiness of implementation and the lack of additional assumptions, which is important especially when the function of a real object being the subject of experimental research is poorly known, or not known at all. The neural networks were created in the Statistica Automated Neural Networks module. For all the neural networks, the same methods of learning were applied, so the same influence on the results is assumed. The Automated Network Search tool was used which enabled the automated evaluation and the selection of multiple network architectures. The random selection of learning cases and the default settings were used: 70% cases as a learning set, 15% cases as a testing set and 15% cases as a validation set. The multi-layer perceptron type of nets, with between 3 and 10 hidden neurons, was used. The automated learning tool used the BFGS algorithm. While searching for the best net, various activation functions were checked: linear, logistic, hyperbolic tangent and exponential. 20 nets with various settings (the number of hidden neurons, activation functions) were trained and the best 5 were saved. Finally, the best nets were selected on the basis of nets quality parameters (sum of squares errors) which were calculated for learning, testing and validation sets.

In the simulation the author studied 3-inputs designs with 5 (first input), 7 (second input), 9 (third input) levels generated according to R and R2 methods. The numbers of design's units were set as 10% (32), 25% (79), 50% (158), 75% (237) of the full design units' number (315). The full design consists of units which are all possible combinations of all inputs' levels. In the case of 5, 7 and 9 levels for 3 inputs there are 315 combinations. The real and approximated values were calculated with the use of a special testing set, consisting of 1331 units and built as all combinations for 11 levels of 3 inputs. The levels were calculated by dividing regularly the input range [–1, 1] into 10 subranges: –1, –0.8, –0.6 ... 1. For each error set (a collection of differences between real and approximated values), statistical parameters were calculated: the maximal and average error, standard deviation, the number of errors' values higher than 0.1 (10% of output range [0, 1]).

4. Results of Simulation

For each error set (the collection of differences between testing function values and approximated values) statistical parameters were calculated: the maximal error, the average error, standard deviation, the number of absolute errors in values higher than 0.1 (10% of the length of the standardized output range 0..1). Additionally, the nonparametric two-samples Kolmogorov-Smirnov test ($\alpha = 0.05$) [9] was conducted to compare approximated values sets and testing function sets. The test, which is very popular to determine whether distributions of two random samples differ significantly, is sensitive to any kind of distributional difference and there is no need to make an assumption about, for example, normality or homogeneity of sample variance. The simulation errors, obtained for the errors sets consisting of values calculated for a special testing set as absolute differences between testing function values and neural approximated values, are shown in Table 2.

Table 2

Simulation errors obtained for the R- and R2-method of smart designs generating

units	errors for R-method					errors for R2-method				
	max	average	err > 0.1	std. dev.	p-value (K-S test)	max	average	err > 0.1	std. dev.	p-value (K-S test)
32	0.45	0.04	100	0.05	< 0.001	0.46	0.05	117	0.05	< 0.001
	0.70	0.11	700	0.09	< 0.001	0.57	0.04	117	0.06	< 0.001
79	0.13	0.01	1	0.01	< 0.01	0.30	0.03	54	0.03	< 0.1
	0.06	0.01	0	0.01	< 0.005	0.38	0.03	53	0.04	> 0.1
158	0.18	0.01	4	0.01	> 0.1	0.12	0.02	7	0.02	> 0.1
	0.05	0.01	0	0.01	> 0.1	0.07	0.01	0	0.01	> 0.1
237	0.04	0.01	0	0.01	> 0.1	0.18	0.02	23	0.02	> 0.1
	0.19	0.02	5	0.02	< 0.025	0.14	0.02	4	0.02	> 0.1
315	0.02	0.01	0	0.01	> 0.1	0.13	0.01	3	0.01	> 0.1
	0.07	0.02	0	0.02	< 0.05	0.10	0.01	1	0.02	> 0.1

Taking into consideration the maximal error, the values evaluated for both generating methods (R and R2) in the case of 32 design's units are definitely too high. The output range for the testing function was normalized to the 0..1 range. The maximal error, for example 0.70, means therefore the difference between the testing function value and the approximated value is on the level of 70% of the length of the output range and it disqualifies the design to be used instead of, for example, a full-design. Furthermore, other errors (e.g. the number of cases where the absolute error is higher than 0.1) calculated for 32-unit designs are high and significantly worse than in the case of other unit numbers. However, it must be noted that 32 units are only 10% of full-design units' number, so the increase of approximation errors (which means a possible loss of information gain in a real experiment conducted according to the analyzed smart design) is very likely. Comparing the errors obtained for the cases of 79, 158 and 237 units, the errors in the case of the R2 generating method are a little bit higher. But the fact is that error values are less varied

when comparing two results obtained for the neural approximation made twice in each case of unit number. Taking into consideration average errors, their low values and low variety are noticed. Additionally, by analyzing the number of cases where absolute errors are higher than 0.1, it can be concluded that the cases of really high absolute errors are not too frequent, which could be important, especially if the average error is preferred instead of the maximal error in order to evaluate a smart design.

Let us focus now on p-values obtained in the two-samples Kolmogorov-Smirnov (K-S) test, conducted using $\alpha = 0.05$ significance level. While comparing the distributions of testing function values sets and approximated values sets, a lot of cases where the difference is significant can be noticed for the R-method. In the case of the R2-method of design generation, the situation seems to be clear – for designs with 79 or more units there are no significant differences between testing function values samples and approximated values samples.

5. Conclusions

The simulation shows that it is possible to reduce the number of experiments (observations, runs) performed in the experimental research. The application of smart designs of experiment could facilitate their reduction, even to 25–50% of the number of full design's units. It could sometimes mean reducing the time or the cost of research significantly. The simulation was conducted for 3-input smart designs. However, the procedures of generating and analyzing smart designs are universal for all designs' characteristics, so the conclusions should also be true for cases of other numbers of inputs, levels or design's units, and could be used in a broad area of engineering.

Having compared the two analyzed methods of generating design's units, the R2-method appears to be more recommended, especially because of the Kolmogorov-Smirnov test results and a smaller variation of errors. The same conclusions might be drawn by analyzing the results of the simulation in which the influence of generation methods on the quality of smart designs of an experiment was studied [13].

References

[1] Polański Z., *Design of experiment in technology*, PWN, Warsaw 1984 (in Polish).
[2] Montgomery D.C., *Design and Analysis of Experiments*, John Wiley & Sons, 2012.
[3] Hinkelmann K., Kempthorne O., *Design and Analysis of Experiments. Volume 1. Introduction to Experimental Design*, John Wiley & Sons, Hoboken 2008.
[4] Hinkelmann K., Kempthorne O., *Design and Analysis of Experiments. Volume 2. Advanced Experimental Design*, John Wiley & Sons, Hoboken 2005.
[5] Goos P., Jones B., *Optimal Design of Experiments: A Case Study Approach*, John Wiley & Sons, Hoboken 2011.
[6] Polański Z., *Empirical research – methodology and computer aiding*, [in:] *Modern metrology*, ed. J. Barzykowski, WNT, Warsaw 2004 (in Polish).
[7] Skowronek A., *A method of self-adaptation calculating the value of the smart designs of experiment generation's parameter*, Pol. J. Environ. Stud. 18 (3B), 2009, 327-331.

[8] Johnson R.A., Bhattacharyya G.K., *Statistics: Principles and Methods*, John Wiley & Sons, Hoboken 2010.

[9] Sheskin D.J., *Handbook of parametric and nonparametric statistical procedures*, 3rd ed., Chapman&Hall/CRC, 2004.

[10] Skowronek A., The *smart designs of experiment generated with circular input's ranges*, [in:] *Problems of Modern Techniques in Engineering and Education*, ed. P. Kurtyka, P. Malczewski, K. Mroczka, K. Ziewiec, Pedagogical University of Cracow, Kraków, 2009, 123-129.

[11] Skowronek A., *A conception of multi-generated smart designs of experiment*, [in:] *Computer Methods in Mechanics CMM-2009: short papers*, ed. M. Kuczma [et al.], Zielona Góra, Univ. Ziel. Góra Press, 2009, 419-420.

[12] http://www-optima.amp.i.kyoto-u.ac.jp/member/student/hedar/Hedar_files/TestGO_files/Page364.htm, retrieved November 2013

[13] Skowronek A., *Evaluating the influence of generation methods on the quantity of the smart design of experiment*, Technical Transactions, 1-M/2013, 357-364.