

# Sequential and optimal experimental design. An example of application to the synthesis of metal alloys

J.-P. Gauchi

*Rhône-Poulenc S.A., Centre de Recherches d'Aubervilliers, 52 rue de la Haie Coq, 93308 Aubervilliers Cedex, France*

**Abstract.** In this article we shall define and illustrate how to further optimize experimental design methodology through a real example from the field of chemistry. This depends on two principles: the first principle is to provide the experimenter with several optimal experimental designs (based on mathematically defined criteria), rather than the traditional single design. In this way several objectives can be simultaneously reconciled. The second principle, underlying the data analysis stage, is the correct use of statistical tests (in regression as well as variance analysis), particularly to check on their validity after using them. This is the methodology we shall briefly attempt to illustrate here.

**Key words.** Optimal design – multilinear regression – sequentiality – experimental research methodology.

## Introduction

This chemical study consists in determining the optimal calcination conditions for a mixture of three powders *U*, *V* and *Z* to get metal alloys *UVZ* with very specific characteristics. A resolution IV fractional factorial experimental design with 7 two-level experimental factors is made up; it will subsequently be referred to as plan A. When the results are modeled, it turns out that, on the one hand, a strong curvature of the response has been ignored by the significant model (hereafter called model A), while, on the other hand, a “confounding” between two interactions prevents the coefficients of the model from being fully interpreted. A second experimental design is then set up to determine a model B which nests the model A and includes the quadratic effects of the 7 experimental factors. To this end several discrete designs based on the criteria of *D*-, *A*- and *G*-optimality are separately calculated.

## Construction of experimental design A

*Note: For reasons of confidentiality the 7 (continuous) experimental factors that govern the calcination of the mixture and the (continuous) response will be designated*

*respectively  $X_1$  to  $X_7$  and  $Y$ , and no information on the chemical process will be furnished. We trust that the reader will not find the text any more difficult to comprehend for all that.*

Given what is at stake, the researcher agrees in a first phase to carry out some forty experiments to quantify the main effects of the factors and a certain number of double interactions, a goal that is attainable using a resolution IV fractional factorial design. Moreover, six experiments (35 to 40) in the center of the experimental domain are repeated to test the possible overall curvature of the response. Two “extreme” experiments are also repeated, as they are supposed to be delicate to reproduce. Experiment 33 is a repetition of experiment 3 and experiment 34 is a repetition of experiment 32. Altogether,  $32 + 6 + 2 = 40$  experiments are performed. The confounding structure is due to six interactions connected thus:  $X_3 \times X_5 = X_6 \times X_7$ ,  $X_3 \times X_6 = X_5 \times X_7$ ,  $X_3 \times X_7 = X_5 \times X_6$ , where the equals sign is symbolic, actually signifying “varies as”. The formal experimental design, a random drawing of the order of execution of the experiments, is given in table I, along with the results (three repetitions of the measurement, not of the experiment), as well as their mean and standard deviation.

**Table I.** Formal experimental design and results (3 repetitions of the measurement, the mean, and the standard deviation).

OBS	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$X_7$	ALEA	$Y_1$	$Y_2$	$Y_3$	YM	SY
1	-1	-1	-1	-1	-1	1	1	3	0.60	0.58	0.61	0.59667	.0001500
2	1	-1	-1	-1	-1	-1	-1	30	0.49	0.48	0.49	0.48667	.0000333
3	-1	1	-1	-1	-1	-1	-1	27	0.38	0.43	0.46	0.42333	.0019000
4	1	1	-1	-1	-1	1	1	7	0.51	0.48	0.55	0.51333	.0005667
5	-1	-1	1	-1	-1	-1	1	19	0.59	0.57	0.54	0.56667	.0005500
6	1	-1	1	-1	-1	1	-1	2	0.57	0.57	0.56	0.56667	.0000167
7	-1	1	1	-1	-1	1	-1	32	0.54	0.54	0.53	0.53667	.0000167
8	1	1	1	-1	-1	-1	1	10	0.54	0.55	0.56	0.55000	.0001000
9	-1	-1	-1	1	-1	-1	-1	18	0.37	0.37	0.36	0.36667	.0000167
10	1	-1	-1	1	-1	1	1	12	0.59	0.58	0.57	0.58000	.0001000
11	-1	1	-1	1	-1	1	1	21	0.65	0.64	0.64	0.64333	.0000500
12	1	1	-1	1	-1	-1	-1	20	0.55	0.56	0.55	0.55333	.0000333
13	-1	-1	1	1	-1	1	-1	9	0.26	0.26	0.26	0.26000	.0000000
14	1	-1	1	1	-1	-1	1	4	0.37	0.36	0.37	0.36667	.0000333
15	-1	1	1	1	-1	-1	1	36	0.43	0.44	0.43	0.43333	.0000333
16	1	1	1	1	-1	1	-1	37	0.25	0.26	0.25	0.25333	.0000333
17	-1	-1	-1	-1	1	1	-1	5	0.62	0.63	0.64	0.63000	.0001000
18	1	-1	-1	-1	1	-1	1	22	0.48	0.49	0.49	0.48667	.0000500
19	-1	1	-1	-1	1	-1	1	15	0.19	0.19	0.20	0.19333	.0000167
20	1	1	-1	-1	1	1	-1	16	0.57	0.56	0.58	0.57000	.0000500
21	-1	-1	1	-1	1	-1	-1	23	0.46	0.46	0.46	0.46000	.0000000
22	1	-1	1	-1	1	1	1	33	0.52	0.52	0.52	0.52000	.0000000
23	-1	1	1	-1	1	1	1	34	0.42	0.43	0.43	0.42667	.0000500
24	1	1	1	-1	1	-1	-1	17	0.57	0.57	0.57	0.57000	.0000000
25	-1	-1	-1	1	1	-1	1	11	0.52	0.52	0.51	0.51667	.0000167
26	1	-1	-1	1	1	1	-1	38	0.64	0.63	0.64	0.63667	.0000333
27	-1	1	-1	1	1	1	-1	13	0.53	0.56	0.57	0.55333	.0005667
28	1	1	-1	1	1	-1	1	25	0.60	0.60	0.58	0.59333	.0000667
29	-1	-1	1	1	1	1	1	39	0.19	0.16	0.18	0.17667	.0003167
30	1	-1	1	1	1	-1	-1	35	0.42	0.41	0.41	0.41333	.0000500
31	-1	1	1	1	1	-1	-1	29	0.50	0.49	0.49	0.49333	.0000500
32	1	1	1	1	1	1	1	31	0.18	0.19	0.16	0.17667	.0001000
33	-1	1	-1	-1	-1	-1	-1	6	0.45	0.43	0.43	0.43667	.0002000
34	1	1	1	1	1	1	1	24	0.13	0.14	0.14	0.13667	.0000500
35	0	0	0	0	0	0	0	26	0.57	0.58	0.57	0.57333	.0000333
36	0	0	0	0	0	0	0	1	0.58	0.59	0.59	0.58667	.0000500
37	0	0	0	0	0	0	0	14	0.59	0.58	0.57	0.58000	.0001000
38	0	0	0	0	0	0	0	8	0.59	0.58	0.60	0.59000	.0000500
39	0	0	0	0	0	0	0	28	0.60	0.60	0.59	0.59667	.0000167
40	0	0	0	0	0	0	0	40	0.60	0.55	0.55	0.56667	.0012500

## Elaboration of Model A

First, the results of the repetitions of the two extreme experiments are averaged after checking that the results are close, and the response is labelled YM. There are now  $32 + 6 = 38$  observations on which the multiple linear regression is to be applied. The matrix of the model is once again perfectly orthogonal. The three next steps are then undertaken using the the SAS/STAT software [1].

### Step 1

The regression model of YM (with an intercept) on the 7 principal effects and the 18 free double interactions is calculated; the three interactions  $X_6 \times X_7$ ,  $X_5 \times X_7$ ,  $X_5 \times X_6$  which constitute the confounding structure are deliberately omitted. The model thus contains 25 potentially explanatory variables. We have 12 degrees of freedom to carry out the first significant high-power statistical tests (100 minus the second-level risk  $\beta$  expressed in %). These tests lead to a step-by-step pooling of the model's different variables. Admitting

a first-level risk  $\alpha$  of 5%, we get a first significant model comprising 6 variables (two principal effects and four double interactions)  $X_3$ ,  $X_4$ ,  $X_2 \times X_4$ ,  $X_3 \times X_4$ ,  $X_3 \times X_6$ ,  $X_4 \times X_6$ .

### Step 2

This step is very important; it makes it possible to verify whether the experimental error hypotheses of Gauss-Markov are reasonably valid. It must be remembered that it is this validity that, *a posteriori*, gives the green light to perform the statistical tests of the previous step. Without getting into mathematical statistics here, we shall merely note that this step can be carried out successfully by working on the residuals of the previous model, which are the only available observations of the random variables of "experimental error". An excellent introduction to the global analysis of residuals can be found in [2]. It should also be added that modern software makes graphing the autocorrelation function easy. Here, no autocorrelation coefficient turns out to be significant. Once these two global tests have been passed, an individual analysis of the residuals is necessary, for

example by calculating the DFITS [3], which measures the impact of a residual on the model. A residual associated with a DFITS value above the threshold  $2\sqrt{(p+1)/n}$ , where  $p$  is the number of variables of the model and  $n$  is the total number of observations, indicates that the mere observation related to this residual can in and of itself influence the model. Therefore, this observation must be investigated if possible. If not possible, we can try replacing the observation by its prediction, a practice justified by empiric experience, and then launching a new pooling. Here the threshold is 0.86, and observations 9 and 19 have respectively a DFITS of 1.22 and 1.62. By replacing observations 9 and 19 by their predictions 0.5151 and 0.3797 we get a new significant model with 5 variables:  $X_3, X_4, X_3 \times X_4, X_3 \times X_6, X_4 \times X_6$ . It is seen that  $X_2 \times X_4$  has disappeared. The new model's degree of explanation is acceptable (79%).

### Step 3

The model obtained in step 2 is a good model of the response under study if the initial postulates truly correspond to the theoretical model, which is generally unknown. It is therefore of the utmost importance to test the model's lack of fit (often written "lof"). It is the role of the experiments in the center of the experimental domain to achieve part of this goal. Here a highly significant curvature test [4] indicates that a curvature effect exists; the null hypothesis to test whether the mean of the results in the center is equal to the mean of the other experiments is rejected with a first-level risk of around  $1.3 \times 10^{-4}$ . As a result the statistical tests on the model's coefficients are no longer valid as they are based on an estimate of the experimental variance made up of an "lof" component and a pure error component. The tests are thus rerun using an estimate based solely on the pure error calculated from the 6 points in the center, that is,  $\sigma_0^2 = 1.23 \times 10^{-4}$ . After calculating the new statistical tests, we end up with model A (see Tab. II), comprising the 5 variables mentioned above. The residuals and the DFITS's are acceptable. Nevertheless, as the (significant) curvature is not taken into account in the model and the  $X_3 \times X_6$  interaction is confounded with the  $X_5 \times X_7$ , model A is not acceptable.

Given this partial conclusion, the researcher must undertake the construction of a second experimental design in order to eliminate the ambiguity of the confounding and determine the factor (or factors) responsible for the curvature. To this end, he postulates a second model here designated model B. Model A is nested in Model B. Model B is made up of 13 variables:  $X_3, X_4, X_3 \times X_4, X_3 \times X_6, X_4 \times X_6, X_5 \times X_7$ , and the quadratic effects of the 7 principal effects (14 terms with the intercept).

### Construction of the optimal designs

As model B is a unique polynomial model, there obviously are no experimental designs perfectly adapted to it to be found in the tables of extant designs. It is thus logical to use algorithmic methods to tailor one's own designs to satisfy previously specified optimal criteria. Taking several criteria into consideration, instead of just one, is one of the two aspects of optimality in the broad sense evoked in this article. Too often, in fact, the experimenter confines himself to the criterion of *D*-optimality, probably to facilitate programming, a consideration that is now obsolete. In our example, to satisfy the priorities of the experimenter, i.e. to set up a model with precise and weakly correlated coefficients while still guaranteeing a satisfactory level of predictive accuracy, we have opted to construct designs based on the criteria of *D*-, *A*- and *G*- optimality. This study will not detail the definitions and mathematical properties of these criteria, nor will it develop the approach based on approximate designs (designs that come close to theoretical designs of probability measure). The reader is directed to a synthesis of the question [5], while the designs will here be worked out within the context of so-called discrete-exact designs. However, for purposes of clarity, the goal of these three criteria will be briefly recalled. The criterion of *D*-optimality aims to minimize the volume of the confidence ellipsoid of the estimates of the coefficients of the postulated model (here model B). This minimization is expressed concretely by an increase in the overall precision of these estimates. The criterion of *A*-optimality minimizes the sum of the variances of these estimates, which also raises their overall

**Table II.** Model A, analysis of variance and regression coefficients (parameter estimates).

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F value	Prob > F
Model	5	0.43232	0.08646	>10000	<0.0001
Pure Error	32	$1.23 \times 10^{-4}$	$3.84 \times 10^{-6}$		
Squared correlation coefficient = 0.7862					
Variable	Parameter Estimate	Standard Error	T for H0: Parameter = 0	Prob >  t	
Intercept	0.498814				
$X_3$	-0.061300	$1.96 \times 10^{-3}$	-31	< 0.0001	
$X_4$	-0.034931	$1.96 \times 10^{-3}$	-18	< 0.0001	
$C_{34}$	-0.067777	$1.96 \times 10^{-3}$	-35	< 0.0001	
$C_{36}$ "or" $C_{57}$	-0.052867	$1.96 \times 10^{-3}$	-27	< 0.0001	
$C_{46}$	-0.033819	$1.96 \times 10^{-3}$	-17	< 0.0001	

precision, but in a different way, by lowering the disparity in these variances. Finally, the criterion of *G*-optimality is to minimize the maximum of the variance of the prediction obtained with this model, which often turns out to be a very fruitful objective for the experimenter. All of these constructions can be carried out using the OPTEx procedure of the SAS/QC software program, version 6.12 [6].

The three criteria selected thus depend on a predetermined goal. Other criteria are possible if the objectives of the researcher are different from those described here. Thus, among other possibilities, we can:

- give precedence to one or several coefficients of the model; an appropriate criterion would then be the criterion of *D<sub>s</sub>*-optimality [5],
- try to discriminate several models [5,7],
- study the competition between the bias and the variance, using criterion *J* [5,8].

Moreover, according to common practice, along with the three criteria, the following usual characteristics are calculated:

- the variance inflation factors (VIF) of the model's coefficient estimates [9]; ideally (an orthogonal design), they are all 1, and the value 10 is never to be exceeded;
- the maximum absolute value of the coefficient of linear correlation between the variables of the model;
- the Turing criterion [5] at which the ideal value 1 conveys the sphericity of the confidence ellipsoid and thus the equality of the estimate variances;
- the values of the normalized variance function [5,10] for each experimental point of each design; this function is equal at each of these points to the number of terms in the model [5], here 14, when the design is *D*-optimal and

*G*-optimal at once; this simultaneity is not systematic when the designs are discrete-exact.

Thus, for the three criteria of *D*-, *A*-, and *G*-optimality, all the discrete-exact optimal designs are successively calculated for 14 to 80. It is known that an optimal research design can be sought simply by imposing or proposing the experiments of a first design [11]. It is not here worthwhile to undertake such a procedure since model B postulates the quadratic effects of all the factors, whereas model A postulated none. Here, the discrete-exact optimal designs are calculated by exchange algorithms [5] from a set of candidates experiments. This set is the full factorial design 3<sup>7</sup> (2187 experiments). Meanwhile, the characteristics of the approximate *D*- and *G*- optimal designs (here denoted *D*\*) are calculated, using, for example, the FEDOR software program [12]. This calculation is necessary to determine the characteristics of the different discrete-exact designs. So, for *D*\* the raw determinant of the information matrix is found to be 6.235 × 10<sup>40</sup>, the normed determinant of the information matrix is 1.0958 × 10<sup>-6</sup>, the information matrix inverse trace is 0.0211, and the Turing criterion is 1.237. We also make sure that the normalized variance function for each experimental point is 14 (for a 14-term model).

### Choice of design B

This choice is facilitated by plotting the graphs of change in the relative increase of the raw determinant, the normed determinant, the trace of the variance-covariance matrix, and the maximum of the variance function, according to the three criteria. For lack of space all these graphs cannot be provided here. We shall confine ourselves to summarizing the essential points in tables III, IV, and V respectively for the three criteria *D*, *A* and *G*. Some very interesting zones are detected in these graphs: a zone A for 21- to 24-experiment designs, and a zone B for 30- to 40-experiment

**Table III.** *D*-optimal designs, essential characteristics.

Number of exper.	<i>D</i> -eff (%)	Trace of the inv. of the inf. matrix	Turing Criterion	Max of ABS( <i>r</i> )	Max of VIF	Variance Fonction		
						min	max	std
21	57.86	2.414	1.302	0.23	1.23	11.28	17.20	1.57
22	66.09	2.199	1.257	0.29	1.18	11.64	17.32	1.41
23	67.58	2.118	1.268	0.23	1.13	11.51	18.16	1.54
24	72.20	2.046	1.279	0.25	1.17	12.13	16.31	1.24
30	82.21	1.578	1.252	0.19	1.09	12.60	16.32	0.88
31	84.64	1.523	1.250	0.24	1.10	12.65	15.77	0.77
32	82.59	1.472	1.248	0.17	1.07	11.66	16.39	1.09
33	86.63	1.413	1.236	0.27	1.11	11.94	15.46	0.82
34	85.30	1.393	1.254	0.19	1.08	11.77	15.83	1.00
35	86.53	1.329	1.235	0.24	1.08	12.40	16.42	0.94
36	90.46	1.298	1.242	0.28	1.10	12.44	15.83	0.91
37	88.38	1.273	1.251	0.22	1.07	12.54	15.15	0.68
38	88.67	1.226	1.241	0.28	1.11	12.64	16.09	0.73
39	89.77	1.205	1.250	0.21	1.07	12.23	15.48	0.68
40	91.23	1.161	1.234	0.25	1.09	12.77	15.85	0.74

**Table IV.** A-optimal designs, essential characteristics.

Number of exper.	D-eff (%)	Trace of the inv. of the inf. matrix	Turing Criterion	Max of ABS(r)	Max of VIF	Variance Function		
						min	max	std
22	53.42	2.107	1.201	0.20	1.12	11.32	18.80	1.85
23	54.38	1.992	1.189	0.22	1.10	11.43	16.81	1.64
30	65.92	1.488	1.170	0.17	1.05	10.74	17.65	1.95

**Table V.** G-optimal designs, essential characteristics.

Number of exper.	D-eff (%)	Trace of the inv. of the inf. matrix	Turing Criterion	Max of ABS(r)	Max of VIF	Variance Function		
						min	max	std
21	65.17	2.34	1.27	0.28	1.16	10.82	17.36	1.72
22	69.00	2.19	1.26	0.33	1.18	11.03	17.09	1.41
27	81.48	1.71	1.22	0.23	1.10	11.98	16.04	1.01
30	89.81	1.57	1.25	0.25	1.09	12.68	15.57	0.71

designs. The designs in zone A are obviously more economical and thus helpful in meeting budgetary constraints, but their characteristics are generally inferior. We have nonetheless examined in detail all the optimal designs of both zones with the objectives put forth in the preceding paragraph kept in mind.

### Zone A offers

- a G-21 solution, i.e. a 21-experiment G-optimal design with 65% D-efficiency, a correlation coefficient maximum ( $r_{\max}$ ) of 0.28, a variance-covariance matrix trace of 2.34, and a standard deviation of 1.72 of the variance function ( $sd_{vf}$ ) at the experimental points;
- a D-24 solution with 72% D-efficiency, an  $r_{\max}$  of 0.25, a trace of 2.05, and an  $sd_{vf}$  of 1.24;
- an A-22 solution with 53% D-efficiency, an  $r_{\max}$  of 0.20, a trace of 2.11, and an  $sd_{vf}$  of 1.85.

### Zone B offers

- a G-30 solution with 90% D-efficiency, an  $r_{\max}$  of 0.25, a trace of 1.57, and an  $sd_{vf}$  of 0.71;
- a D-32 solution with 83% D-efficiency, an  $r_{\max}$  of 0.17, a trace of 1.47, and an  $sd_{vf}$  of 1.09;
- a D-39 solution with 90% D-efficiency, an  $r_{\max}$  of 0.21, a trace of 1.21, and an  $sd_{vf}$  of 1.09;
- an A-30 solution with 66% D-efficiency, an  $r_{\max}$  of 0.17, a trace of 1.40, and an  $sd_{vf}$  of 1.95.

After all these values had been compared, the compromise choice was the D-24 design, as it gives acceptable values for the criteria for a reasonable number of experiments. The 24 experiments of the D-24 design are to be selected from the 24 sets found in table VI. This design thus unveils an ultimate flexibility, since the experimenter is left to a choice between several equivalent experiments. The reason for this ultimate flexibility is based on the fact that the model B is

not full: for example, the factors  $X_1, X_2, X_5, \dots$ , are not present. So, we can find equivalent experiments (relatively to the present factors and interactions) determined by different levels of these absent factors in the whole factorial set. No special algorithm is necessary to do that.

**Table VI.** Sets of possible experiments for the D-24 design.

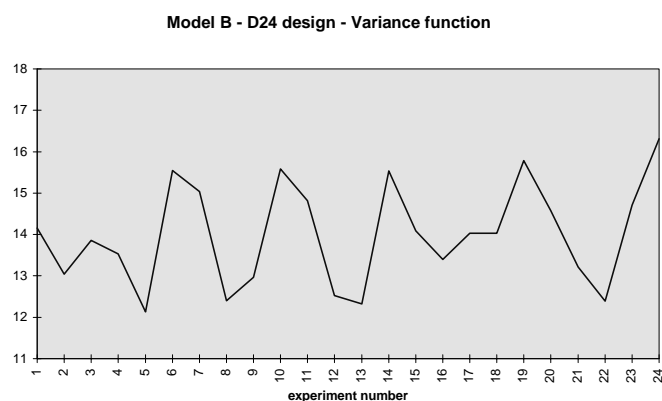
Experiment 1: one of the following 4							
OBS	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$X_7$
1	-1	0	-1	-1	0	1	-1
2	-1	0	-1	-1	0	1	1
3	1	0	-1	-1	0	1	-1
4	1	0	-1	-1	0	1	1
Experiment 2: one of the following 8							
OBS	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$X_7$
1	-1	-1	-1	-1	-1	1	0
2	-1	-1	-1	-1	1	1	0
3	-1	1	-1	-1	-1	1	0
4	-1	1	-1	-1	1	1	0
5	1	-1	-1	-1	-1	1	0
6	1	-1	-1	-1	1	1	0
7	1	1	-1	-1	-1	1	0
8	1	1	-1	-1	1	1	0
Experiment 3: one of the following 4							
OBS	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$X_7$
1	0	-1	-1	-1	-1	-1	1
2	0	-1	-1	-1	1	-1	-1
3	0	1	-1	-1	-1	-1	1
4	0	1	-1	-1	1	-1	-1
Experiment 4: one of the following 2							
OBS	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$X_7$
1	0	0	-1	-1	-1	-1	-1
2	0	0	-1	-1	1	-1	1
Experiment 5: one of the following 2							
OBS	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$X_7$
1	-1	0	-1	0	0	1	0
2	1	0	-1	0	0	1	0

Table VI. continued.

<i>Experiment 6: one of the following 8</i>							
OBS	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>
1	-1	-1	-1	0	-1	0	-1
2	-1	-1	-1	0	1	0	1
3	-1	1	-1	0	-1	0	-1
4	-1	1	-1	0	1	0	1
5	1	-1	-1	0	-1	0	-1
6	1	-1	-1	0	1	0	1
7	1	1	-1	0	-1	0	-1
8	1	1	-1	0	1	0	1
<i>Experiment 7: one of the following 4</i>							
OBS	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>
1	0	-1	-1	1	-1	1	-1
2	0	-1	-1	1	1	1	1
3	0	1	-1	1	-1	1	-1
4	0	1	-1	1	1	1	1
<i>Experiment 8: one of the following 2</i>							
OBS	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>
1	0	0	-1	1	-1	0	1
2	0	0	-1	1	1	0	-1
<i>Experiment 9: one of the following 8</i>							
OBS	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>
1	-1	-1	-1	1	-1	-1	1
2	-1	-1	-1	1	1	-1	-1
3	-1	1	-1	1	-1	-1	1
4	-1	1	-1	1	1	-1	-1
5	1	-1	-1	1	-1	-1	1
6	1	-1	-1	1	1	-1	-1
7	1	1	-1	1	-1	-1	1
8	1	1	-1	1	1	-1	-1
<i>Experiment 10</i>							
OBS	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>
1	0	0	-1	1	0	-1	0
<i>Experiment 11: one of the following 2</i>							
OBS	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>
1	0	-1	0	-1	0	0	0
2	0	1	0	-1	0	0	0
<i>Experiment 12: one of the following 4</i>							
OBS	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>
1	0	0	0	0	-1	-1	1
2	0	0	0	0	-1	1	1
3	0	0	0	0	1	-1	-1
4	0	0	0	0	1	1	-1
<i>Experiment 13: one of the following 4</i>							
OBS	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>
1	0	0	0	0	-1	-1	0
2	0	0	0	0	-1	1	0
3	0	0	0	0	1	-1	0
4	0	0	0	0	1	1	0
<i>Experiment 14: one of the following 8</i>							
OBS	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>
1	-1	-1	0	1	0	1	-1
2	-1	-1	0	1	0	1	1
3	-1	1	0	1	0	1	-1
4	-1	1	0	1	0	1	1
5	1	-1	0	1	0	1	-1
6	1	-1	0	1	0	1	1
7	1	1	0	1	0	1	-1
8	1	1	0	1	0	1	1

Table VI. continued.

<i>Experiment 15: one of the following 4</i>							
OBS	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>
1	-1	0	0	1	-1	0	-1
2	-1	0	0	1	1	0	1
3	1	0	0	1	-1	0	-1
4	1	0	0	1	1	0	1
<i>Experiment 16: one of the following 2</i>							
OBS	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>
1	0	0	1	-1	0	-1	-1
2	0	0	1	-1	0	-1	1
<i>Experiment 17: one of the following 8</i>							
OBS	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>
1	-1	-1	1	-1	-1	-1	0
2	-1	-1	1	-1	1	-1	0
3	-1	1	1	-1	-1	-1	0
4	-1	1	1	-1	1	-1	0
5	1	-1	1	-1	-1	-1	0
6	1	-1	1	-1	1	-1	0
7	1	1	1	-1	-1	-1	0
8	1	1	1	-1	1	-1	0
<i>Experiment 18: one of the following 4</i>							
OBS	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>
1	-1	0	1	-1	-1	0	1
2	-1	0	1	-1	1	0	-1
3	1	0	1	-1	-1	0	1
4	1	0	1	-1	1	0	-1
<i>Experiment 19: one of the following 4</i>							
OBS	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>
1	0	-1	1	-1	-1	1	-1
2	0	-1	1	-1	1	1	1
3	0	1	1	-1	-1	1	-1
4	0	1	1	-1	1	1	1
<i>Experiment 20: one of the following 8</i>							
OBS	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>
1	-1	-1	1	0	0	1	-1
2	-1	-1	1	0	0	1	1
3	-1	1	1	0	0	1	-1
4	-1	1	1	0	0	1	1
5	1	-1	1	0	0	1	-1
6	1	-1	1	0	0	1	1
7	1	1	1	0	0	1	-1
8	1	1	1	0	0	1	1
<i>Experiment 21: one of the following 4</i>							
OBS	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>
1	0	-1	1	1	-1	-1	1
2	0	-1	1	1	1	-1	-1
3	0	1	1	1	-1	-1	1
4	0	1	1	1	1	-1	-1
<i>Experiment 22: one of the following 2</i>							
OBS	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>
1	0	0	1	1	-1	-1	-1
2	0	0	1	1	1	-1	1
<i>Experiment 23: one of the following 2</i>							
OBS	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>
1	0	-1	1	1	0	0	0
2	0	1	1	1	0	0	0
<i>Experiment 24: one of the following 4</i>							
OBS	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>
1	-1	0	1	1	-1	1	0
2	-1	0	1	1	1	1	0
3	1	0	1	1	-1	1	0
4	1	0	1	1	1	1	0



**Fig. 1.** *D*-24 design: values of the normalized variance function at each of the support points of the design.

## Conclusion

This study contributes no special innovation in defining new criteria, but it does, through the real example drawn from chemistry, emphasize two aspects: the methodology used to elaborate a statistically significant model in the context of experimental design, and a sequential design technique based on several criteria of optimality in the broad sense. It is this double approach that we call broad-based optimality. It should make possible the construction of several optimal experimental designs, i.e. based on different criteria of optimality that enable the researcher to respond to distinct objectives. Rather than having just one design at his disposal, he can choose among several designs, and this power of choice should make him reflect more deeply about his priorities. It is too often forgotten that this broad-based optimality should

also comprise a rigorous analysis of the results, with careful respect for the rules of modelization that encompasses the variability inherent in all experimental science.

## References

1. Logiciel SAS/STAT version 6.12. SAS Institute Inc. Cary, North Carolina.
2. Draper, N. R.; Smith, H. Applied Regression Analysis, Second Edition, Wiley, New York, 1981; pp 141-192.
3. Belsley, D. A.; Kuh, E.; Welsch, R. E. Regression Diagnostics, Wiley, New York, 1980; p 15.
4. Montgomery, D. C. Design and Analysis of Experiments, Wiley, New York, 1984; p 306.
5. Gauchi, J. P. Plans d'expériences optimaux pour modèles linéaires. In: Plans d'expériences- Applications à l'entreprise, Driesbeke, G.; Saporta, J.; Fine Eds., Technip, 1997; pp 325-400.
6. Logiciel SAS/QC version 6.12. SAS Institute Inc. Cary, North Carolina.
7. Mathieu, D. Contribution de la méthodologie de la recherche expérimentale à l'étude des relations structure-activité. Thèse Université Aix-Marseille III, 1981.
8. Khuri, A. I.; Cornell J. A. Response Surfaces: designs and analyses. Marcel Dekker, 1987, p 209.
9. Tomassone, R.; Lesquoy, E.; Miller, C. La régression, nouveaux regards sur une ancienne méthode statistique, Masson, Paris, 1983; p 147.
10. Box, G. E. P.; Draper, N. R. Empirical Model-Building and Response Surfaces, Wiley, New-York, 1987; p 481.
11. Gauchi, J. P. Application de la *D*-optimalité à un problème de réparation de plan d'expériences; in: Plans d'expériences- Applications à l'entreprise, Driesbeke, Saporta, G.; Fine, J. Eds., Technip, 1997; pp 439-450.