CS-BIGS 5(2): 115-125

http://www.bentley.edu/csbigs

Comparison of different screening methods

Delphine Dupuy
Ecole des Mines de Saint-Etienne, France

Bernard Corre
Total, France

Magalie Claeys-Bruno
Aix Marseille Université (LISA EA4672), France

Michelle Sergent
Aix Marseille Université (LISA EA4672), France

The aim of this paper is to illustrate sensitivity methods using experimental designs. These techniques are used to deal with simulation models or experimental programs involving a large number of controllable factors. In these cases, one of the objectives is the identification of the subset of factors that have substantial influence on the responses. In this paper we provide a review of different screening methods that are useful in eliminating negligible factors so that efforts may be concentrated upon just the important ones. These different methods are described and are applied on a simulation model in the petroleum industry involving 51 parameters. The comparison of the results is presented with the advantages and disadvantages for each method. The presentation is accessible to readers with an intermediate level of statistics.

Keywords: experimental design, screening, computer experiments, group screening, supersaturated design, sequential bifurcation

Introduction

In the last decade, several disciplines within scientific research and industry have used simulation codes in order to forecast, to optimize and to make good decisions in the context of their studies. These codes which simulate complex phenomena, often badly known and sometimes coupled between themselves, are increasingly realistic and, even if the means of calculation are powerful, the computer time could be expensive: a simulation run may sometimes requires several days of computation. Moreover, these codes take into account a very large number of parameters and many questions remain unanswered, such as the detection of influential parameters corresponding to a screening study, or the study of the impact of small variations in the parameters often called "sensitivity analysis" (Azarian and al. 2011).

To answer these questions, one must find an effective strategy that is as cheap as possible but reliable. It seems obvious that the field of experimental designs, which guides the choice of the best informative experiments, can find a new potential by its transposition to the field of "numerical experiments". Thus, the objective of this study

is to test on a real application different screening methodologies, with different number of simulations, and to compare them.

In the following section, six screening strategies are presented:

- Resolution III Hadamard design
- Resolution IV Hadamard design
- Supersaturated design
- Group screening
- Multiple group screening
- Sequential bifurcation

Before presenting the different methods, we provide more details about the case study.

1. Illustrative example

1.1. Formalization

A phenomenon can be considered as a black box, characterized by:

- variables of environment,uenv, generally badly known and numerous,
- variables of control, ucont, which can be fixed by the user,
- variables of interest, yreal, often called response, with the relation:

$$y_{real} = f_{real}(u_{env}, u_{cont})$$

The function f_{real} is usually approximated by a simulator and in this context a new class of parameters is added, the simulation parameters u_{sim} which can parameterize, and eventually adjust the simulator.

1.2. General framework

The present case study is based on the use of a flow simulator in porous media and concerns an oil field produced by nine wells, five of them being producers and four injectors.

The system is solved by the technique of streamlines. The responses are multiple and are functions of time: in the study in this paper, only the cumulative oil production at a fixed time (2000 days) is considered.

The objective of the study is to simulate as precisely as possible the behavior of the reservoir. The data file is complex and contains several hundreds or thousands of parameters, but only 51 of them, noted U_1, U_2 ..., U_{51}, may vary in this study, taking into account different types of variables:

- environment variables, such as fluid or rock properties (porosity, permeability...)
- control variables, such as location of the wells or injection or production constraints (flow and/or pressure),
- simulation variables, such as the number of streamlines considered in the pressure solver.

These variables may vary within their range of uncertainties and the question is then:

"What is the impact of a small change in input data on the response of the simulator?"

In other words, we want to identify parameters for which a small variation yields a variation in the response. This type of study corresponds to the goal of a screening study that can be defined as a strategy to very quickly identify active factors among many candidate factors.

2. Screening methods

As previously mentioned, at the beginning of a study, we generally do not know which factors have an influence on the studied responses and we want to identify, among a large number of potentially significant factors, those which are effectively active within a fixed experimental domain. This objective falls within the screening strategies, which depend on the "effect sparsity" assumption that says that only a few factors are responsible for most of the effect in a response while most are not. This is equivalent to the Pareto rule used in quality studies. These screening methods are very useful tools for examining simulation models that involve a large number of factors (Barton, 2001; Kleijnen, 2004; Sanchez and al., 2005; Welch and al., 1992). This screening stage should involve a minimum number of experiments, and should not take much computing time.

2.1. Screening designs

2.1.1. Hadamard R_{III} et R_{IV}

The most well-known screening designs for estimating the additive (first order) mathematical model are the Hadamard or Plackett and Burman designs (1946). A Hadamard design is a weighing experimental design which contains only elements x_{ij} equal to ±1 corresponding to the two studied levels and for which the resulting information matrix $X'X$ is such that $(X'X) = N$

I_N with X the model design, X' the transpose matrix of X, N the number of simulations and I_N the identical design.

These Hadamard designs exist when N is a multiple of 4 (N = 4, 8, 12, 16, 20, 24, ...). In this set of designs we can consider two subsets: designs where N is a power of 2 (N = 4, 8, 16,...), called regular designs, and non-regular designs where N is a multiple of 4 but not a power of 2 (N = 12, 20, 24, 28, 36, ...). In the case of regular designs, each interaction effect is totally confounded with one of the main effects. For non-regular designs, the aliasing is different: the X_iX_j interaction effect is partially confounded with all of the main effects except those of X_i an X_j. Therefore, with these designs, the risk of drawing a wrong conclusion is reduced, even if some interaction effects exist. In our case, we want to study 51 factors in order to determine those which are influential. For that, we can use a two-level Hadamard design with 56 experiments. This is a non-regular design, where the interaction effects are partially aliased with the main effects: the coefficients of aliasing between interaction and main effects are 0.43 or 0.14. Taking into account the size of the table (51 columns and 56 rows), the experimental design can not be presented in this article but is available as a MS Excel file along with the article.

This design, with N = 56 experiments, is a resolution III screening design. It is possible to generate a resolution IV design by operating a mirror-image foldover: it is obtained by reversing the signs of all the columns of the original matrix. The original design runs are combined with the mirror-image foldover design runs and this combination leads to a R_{IV} design, with 2N experiments, where the main effects are free of all first-order interactions and confounded only with second-order interactions.

2.1.2. Supersaturated designs

A supersaturated design is a design for which there are fewer runs than effects to be estimated. Firstly developed in the 1950's by Satterthwaite (1959) as a random balance and Booth and Cox (1962) in a systematic manner, these designs have recently become increasingly popular.

Many authors have proposed methods for constructing and analysing supersaturated designs (Beth and al., 1999; Butler et al., 2001; Catterjee and Gupta, 2003; Cheng, 1997; Claeys-Bruno and al., 2009, 2011; Deng et al., 1999; Yamada and Lin, 1997; Lin, 1993, 1995; Liu and Hickernell, 2002; Liu and Dean, 2004, Lu and Meng, 2000, Nguyen, 1996; Tang and Wu, 1997; Wu, 1993). The most well-known method proposed by Lin in 1993 consists in considering a non-regular Hadamard design, H_N, with N experiments, allowing to study up to (N-1) factors. The design H_N can be written differently by considering one of the columns of the design, called the branching column, and by grouping the N experiments into two groups: a group (H_1) with a + sign in this column and a second group (H_2) with a – sign in this column. Using this notation, the Hadamard design can be written as follows:

$$H_N = \begin{bmatrix} 1 & H_1 \\ -1 & H_2 \end{bmatrix}$$
$$\uparrow$$
branching column

If we consider H_1 or H_2, we obtain designs with N/2 rows and (N-2) distinct columns, called supersaturated designs allowing the study of up to (N-2) factors in $N/2$ experiments.

In this particular case, we selected column 55 in the Hadamard design with N = 56 as the branching column; the rows with -1 in this column provide a supersaturated design allowing the study of 51 factors in 28 experiments.

2.1.3. Group screening (Dorfman, 1943; Finucan, 1964; Hunter and Mezaki, 1964; Patel, 1962)

Watson (1961) suggested an alternative idea by grouping factors into a smaller number of groups and treating each group of factors as a "factor" which is called a grouped factor to distinguish it from the original factors. A much smaller design can then be used to study the grouped factors. First, the experimenter must use experience and knowledge of the problem to arrange the factors into logical groups. A grouped factor is said to be at its "high" level if each of its individual factors is at a level that gives a higher response. Analogously, a group is at its "low" level if each individual factor gives a lower response. An adequate design is then run on the groups. The results of the first stage are analyzed and used to establish the second stage; this screening method is iterative. Upon identification of important grouped factors in the first stage, the factors within these groups are separated into smaller groups (multiple stage procedure) or individual factors (2-stage procedure) and a new design on the subgroups is run until the active factors are identified.

This method can lead to a strategy with few experiments but requires important assumptions to ensure that the effect of a grouped factor is significant if and only if the effect of at least one of its factors is significant. Furthermore, all the factors of the groups with a non-significant effect are eliminated at the first stage.

Moreover, the levels of each factor are chosen so that the effects, if the factor is active, are positive.

Since the method is iterative, the different steps will be explained in the following section ("Results").

2.1.4. Multiple group screening (Morris, 1987)

Sequential group screening methods may be operationally impractical, for example, when the execution of each run requires substantial time but many runs can be executed simultaneously. In the multiple group procedure, each factor is assigned to more than one group at the first step and a factor is potentially influential if all groups containing this factor are active. The individual factors followed up in the second stage are those for which all types of groups are active.

2.2. Sequential bifurcation

Sequential bifurcation (SB) is a group-screening technique proposed by Bettonvil (1990) in 1990 and developed by Bettonvil and Kleijnen (1997).

The Sequential Bifurcation (SB) procedure is sequential: it consists of a sequence of steps.

The first step aggregates all factors into a single group and tests whether or not that group has an important effect by comparing the two extreme factor combinations: all factors low and the other one, all factors high. If the group has indeed an important effect, then the second step splits the group into two subgroups – bifurcates – and tests each of these subgroups for importance. SB analysis compares the new observation with two old observations, to infer the two sums of the effects in the two smaller sub-subgroups.

The next steps continue in a similar way: SB splits active subgroups into smaller subgroups and discards non active subgroups. It is clear that at the end of the bifurcation individual factor effects are estimated.

The most important assumption used by the SB, as in all group-screening techniques, is that the experimenter knows whether a specific individual factor has a positive or negative effect on the response: the factor effects must have known signs.

3. Results and interpretation

3.1. Classical study

Scientists often perform one-factor-at-a-time (OFAT) experiments, which vary only one factor or variable at a time while keeping others fixed. In our case, the OFAT method needs one simulation for each level for all variables, so 102 simulations. The results are presented on Figure 1.

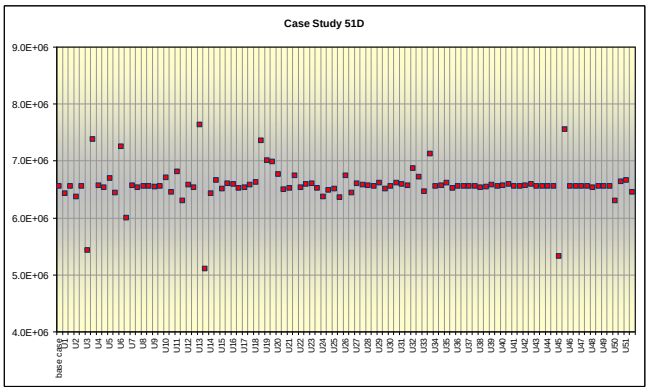


Figure 1. OFAT results

3.2. Screening designs

3.2.1. R_{III} and R_{IV} Hadamard designs

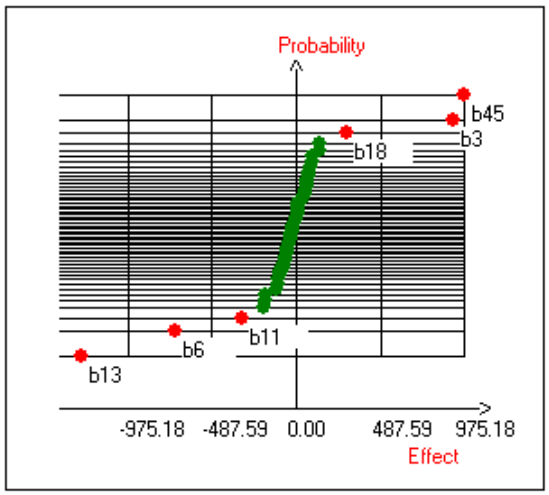
The 56 and 112 simulations of the experimental designs of R_{III} and R_{IV} respectively, have been achieved using the simulator and we have the production at 2000 days $\times 10^{-3}$ for each experiment.

R_{III} Hadamard design:

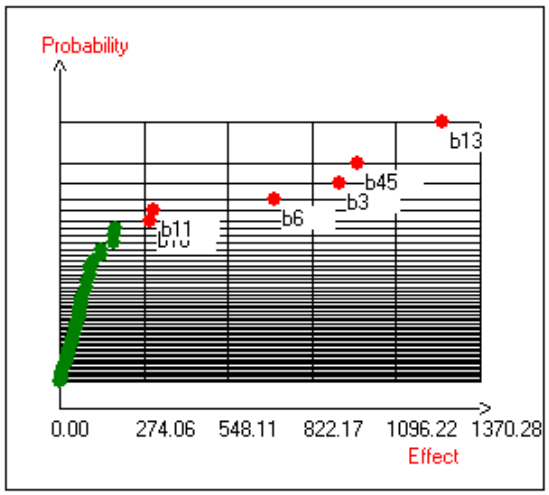
We postulate that the result of each experiment (η = production 2000 days $\times 10^{-3}$) is a linear combination of the effect of each coded variables X_1, X_2, ... , X_{51}. A first order polynomial model for the 51 variables is proposed. This model is a "model of the model simulation" and is often called a meta model. It is valid only for the 2 levels of the factors, and therefore cannot be used for any interpolation or extrapolation.

$$\eta = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + ... + \beta_{51} X_{51}$$

We can then estimate the model coefficients b_i using multiple linear regression. It is expected that the large effects, whether positive or negative, are the statistically significant ones. We can use various statistical tools(Mathieu et al, 2007) for identifying active factors which are commonly used such as Daniel's plots (Figure 2) where the probability P of the occurrence of some values b_i is plotted against b_i. Most of the points appear to lie on a straight line, corresponding to non-significant effects, and6 points deviate from it corresponding to the active factors U_{13}, U_{45}, U_3, U_6, U_{11} and U_{18}.



(a)



(b)

Figure 2. Daniel's plots for the response "production at 2000 days $\times 10^{-3}$" for the Hadamard R_{III} design (a) Effects values b_i and (b) absolute effects values $|b_i|$ have been plotted on normal probability paper.

R_{IV} Hadamard design:

The same statistical tools have been used in order to determine active factors with the resolution IV Hadamard design. The coefficient values can be represented on an effects plot (Figure 3) where magnitudes and signs of each effect of the variables are shown. For more readability, only the twenty most important effects have been reported on this figure. A significant limit could be calculated by Lenth's method (1989). This graph shows that the factors U_{13}, U_{45}, U_3, U_6, U_{11} and U_{18} are the most important for the production at 2000 days.

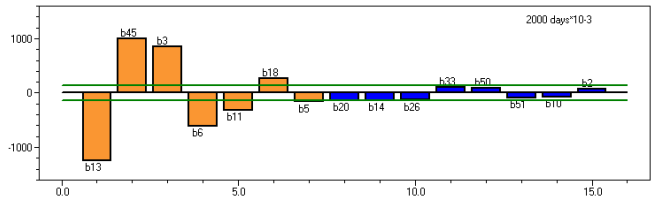


Figure 3. Effects plot for the response "production at 2000 days$\times 10^{-3}$"for the Hadamard R_{IV} design. The importance (coefficient values b_i) of each input factor x_i is proportional to the length of the bar. (Only the first twenty values have been reported). The significance limits are drawn as horizontal continuous lines. Only factors whose effects are clearly outside this limit are considered as influential.

The results confirm the influence of the factors U_{13}, U_{45}, U_3, U_6, U_{11} and U_{18}.

3.2.2. Supersaturated designs

Standard methods for analysing data, such as multiple linear regression, cannot be used in the case of supersaturated designs since the least squares estimates are not unique. There is no way to estimate all the main effects simultaneously. In this example, we can use a combination of a stepwise regression and an all subsets selection procedure as proposed by Abrahamand Chipman (1999) but the high number of factors makes the conventional all-subsets regressions infeasible. We have applied a strategy developed by Lu and Wu (2003) in 2003and more recently modified by R. Phan-Tan-Luucalled sequential approach (Celaand al., 2007). In a first step, the stepwise regression reported in Figure 4a clearly shows that five variables are sufficient to obtain an R^2 greater than 0.90. So, in can be anticipated that the number of active factors is probably not higher than 5.

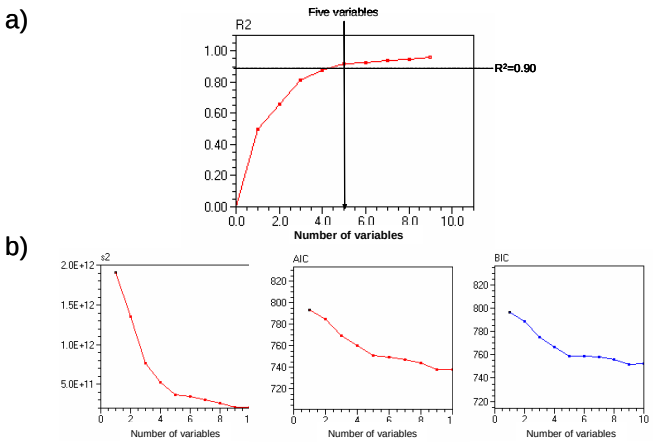
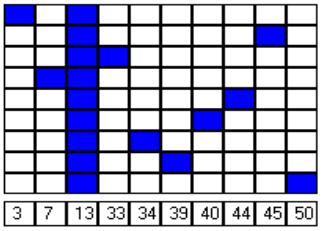
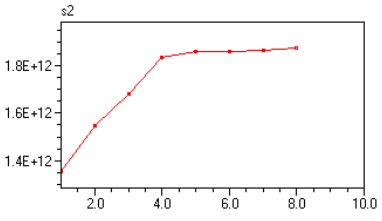


Figure 4. a) Stepwise regression up to nine variables and b) graphs of s^2, AIC and BIC

Regression with 2 variables

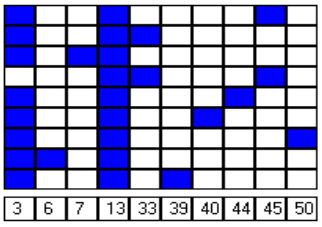
Var	Var	$s^2.10^{-12}$	R^2
b3	b13	1.35	0.65
b13	b45	1.54	0.60
b13	b33	1.67	0.57
b7	b13	1.83	0.53
b13	b44	1.85	0.52
b13	b40	1.85	0.52
b13	b34	1.86	0.52
b13	b39	1.87	0.52

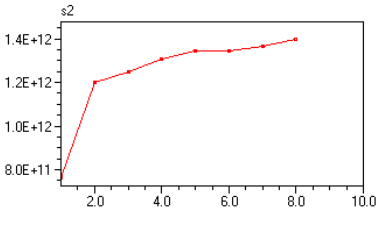




Regression with 3 variables

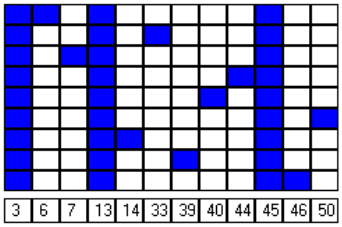
Var	Var	Var	$s^2.10^{-12}$	R^2
b3	b13	b45	0.76	0.814
b3	b13	b33	1.20	0.707
b3	b7	b13	1.25	0.696
b13	b33	b45	1.30	0.682
b3	b13	b44	1.34	0.673
b3	b13	b40	1.34	0.673
b3	b13	b50	1.36	0.667
b3	b6	b13	1.39	0.661

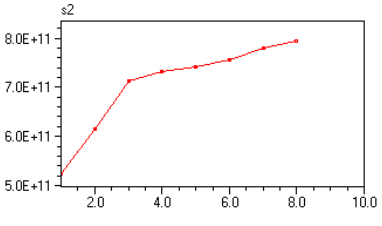




Regression with 4 variables

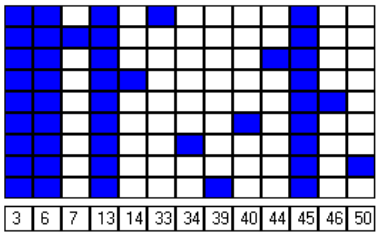
Var	Var	Var	Var	$s^2.10^{-12}$	R^2
b3	b6	b13	b45	0.52	0.878
b3	b13	b33	b45	0.61	0.857
b3	b7	b13	b45	0.71	0.834
b3	b13	b44	b45	0.73	0.829
b3	b13	b40	b45	0.74	0.827
b3	b13	b45	b50	0.75	0.824
b3	b13	b14	b45	0.78	0.818
b3	b13	b39	b45	0.79	0.815

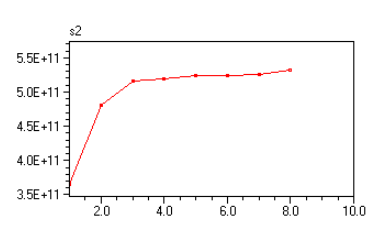




Regression with 5 variables

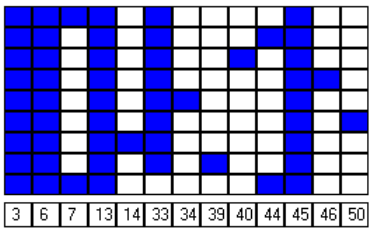
Var	Var	Var	Var	Var	$s^2.10^{-12}$	R^2
b3	b6	b13	b33	b45	0.36	0.918
b3	b6	b7	b13	b45	0.48	0.893
b3	b6	b13	b44	b45	0.51	0.885
b3	b6	b13	b14	b45	0.51	0.884
b3	b6	b13	b45	b46	0.52	0.883
b3	b6	b13	b40	b45	0.52	0.883
b3	b6	b13	b34	b45	0.52	0.883
b3	b6	b13	b45	b50	0.53	0.881

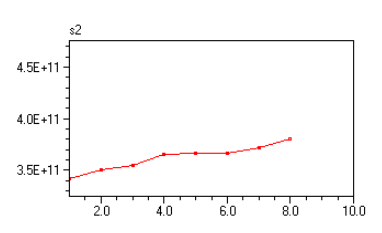




Regression with 6 variables

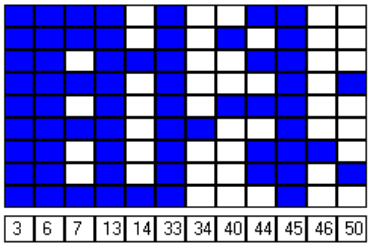
Var	Var	Var	Var	Var	Var	$s^2.10^{-12}$	R^2
b3	b6	b7	b13	b33	b45	0.34	0.927
b3	b6	b13	b33	b44	b45	0.35	0.926
b3	b6	b13	b33	b40	b45	0.35	0.924
b3	b6	b13	b33	b45	b46	0.36	0.922
b3	b6	b13	b33	b34	b45	0.36	0.922
b3	b6	b13	b33	b45	b50	0.36	0.922
b3	b6	b13	b14	b33	b45	0.37	0.921
b3	b6	b13	b33	b39	b45	0.38	0.919





Regression with 7 variables

Var	Var	Var	Var	Var	Var	Var	$s^2.10^{-12}$	R^2
b3	b6	b7	b13	b33	b44	b45	0.30	0.937
b3	b6	b7	b13	b33	b40	b45	0.31	0.936
b3	b6	b13	b14	b33	b44	b45	0.32	0.935
b3	b6	b7	b13	b33	b45	b50	0.33	0.933
b3	b6	b13	b33	b40	b44	b45	0.33	0.932
b3	b6	b7	b13	b33	b34	b45	0.34	0.931
b3	b6	b13	b33	b44	b45	b46	0.34	0.930
b3	b6	b13	b33	b44	b45	b50	0.34	0.930



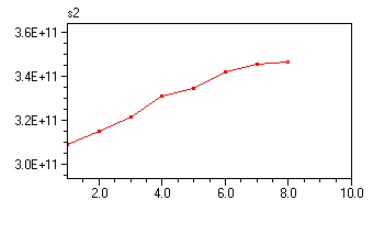


Figure 5. All subset regressions for $f=2$ to $f=7$ variables. The first column of the table gives the name of the selected variables in each model and the corresponding values of R^2 and s^2 for the eight best solutions. The second column shows the mapping of the selected variables where the presence of the variables is marked with a blue square. The third column shows the s^2 evolution for a fixed number of variables, corresponding to different models, numbered 1 to 8, with s^2 increasing.

The study of BIC (Bayesian Information Criterion) (Schwarz, 1978) and AIC (Akaike's Information Criterion) (Akaike, 1974) values can also be used to interpret the results of the supersaturated design in order to determine the number of necessary variables in the model.

Figure 4b shows the evolution of s^2 (residual variance), AIC and BIC in terms of the number of variables in the model. These graphs confirm that the number of active factors is probably not higher than five.

Therefore, the second stage involves all subsets regressions for a number of factors (f) from two to five. Figure 5 summarizes the values of R^2 and s^2 of the different models for each value of f.

These values yield unequivocal results: five factors (U_{13}, U_{45}, U_3, U_6 and U_{33}) are detected as active with an R^2 value equal to 0.92 and s^2 value equal to 3.6×10^{11}. The addition of a supplementary variable in the model ($f=6$) does not change the values of the R^2 and s^2significantly, which allows us the selection of the model with 5 variables.

3.2.3. Group screening

First, the experimenter must use knowledge of the problem to organize the factors into logical groups. The experimenter attributes a probability upon the possible influence: 1 corresponds to a small probability and 2, to a large probability. Factors where no knowledge is available will be studied alone (a group with one factor). Table 1 summarizes the attribution of the levels (the level + chosen as increasing the response), the probability *a priori* proposed by the experimenters and the construction of the groups.

Table 1. Factors, probability a priori, groups

	(-)	(+)	Proba	Group		(-)	(+)	Proba	Group
U_1	0.95	1.0	1	G_1	U_{27}	76	78	1	G_3
U_2	0.95	1.0	2	G_7	U_{28}	11	9	1	G_3
U_3	0.5	1.5	2	G_7	U_{29}	62	60	1	G_3
U_4	0.011	0.0090	2	G_7	U_{30}	34	36	1	G_4
U_5	1.02	0.98	1	G_1	U_{31}	11	9	1	G_4
U_6	1.2	0.8	2	G_7	U_{32}	59	71	1	G_4
U_7	48	42	?	G_9	U_{33}	42	61	1	G_4
U_8	0.065	0.075	?	G_{10}	U_{34}	2000	6000	?	G_{14}
U_9	61	65	?	G_{11}	U_{35}	2550	2450	1	G_4
U_{10}	0.35	0.25	2	G_7	U_{36}	2500	1500	1	G_5
U_{11}	0.29	0.15	2	G_8	U_{37}	1000	2000	1	G_5
U_{12}	1.05	0.95	1	G_1	U_{38}	700	1300	1	G_5
U_{13}	2	0.6	2	G_8	U_{39}	3000	2000	1	G_5
U_{14}	0.23	0.17	?	G_{12}	U_{40}	700	1300	1	G_5
U_{15}	3500	3700	?	G_{13}	U_{41}	2000	3000	1	G_6
U_{16}	11	9	1	G_1	U_{42}	2500	3500	1	G_6
U_{17}	71	69	1	G_1	U_{43}	3500	4500	1	G_6
U_{18}	48	66	2	G_8	U_{44}	4500	5500	1	G_6
U_{19}	78	70	2	G_8	U_{45}	3500	4500	2	G_8
U_{20}	16	14	1	G_2	U_{46}	3500	4500	?	G_{15}
U_{21}	38	36	1	G_2	U_{47}	4500	5500	?	G_{16}
U_{22}	26	28	1	G_2	U_{48}	700	1300	?	G_{17}
U_{23}	15	13	1	G_2	U_{49}	4500	5500	?	G_{18}
U_{24}	63	61	1	G_2	U_{50}	700	1300	?	G_{19}
U_{25}	21	19	1	G_3	U_{51}	3000	2000	1	G_6
U_{26}	46	44	1	G_3					

For the study of these 19 grouped factors, a non-regular Hadamard design with 20 experiments was constructed, as shown in Table 2.

Table 2. Hadamard design with 20 experiments

N°Exp	G_1	G_2	G_3	G_4	G_5	G_6	G_7	G_8	G_9	G_{10}	G_{11}	G_{12}	G_{13}	G_{14}	G_{15}	G_{16}	G_{17}	G_{18}	G_{19}
1	+	+	−	−	+	+	+	+	−	+	−	+	−	−	−	−	+	+	−
2	−	+	+	−	−	+	+	+	+	−	+	−	+	−	−	−	−	+	+
3	+	−	+	+	−	−	+	+	+	+	−	+	−	+	−	−	−	−	+
4	+	+	−	+	+	−	−	+	+	+	+	−	+	−	+	−	−	−	−
5	−	+	+	−	+	+	−	−	+	+	+	+	−	+	−	+	−	−	−
6	−	−	+	+	−	+	+	−	−	+	+	+	+	−	+	−	+	−	−
7	−	−	−	+	+	−	+	+	−	−	+	+	+	+	−	+	−	+	−
8	−	−	−	−	+	+	−	+	+	−	−	+	+	+	+	−	+	−	+
9	+	−	−	−	−	+	+	−	+	+	−	−	+	+	+	+	−	+	−
10	−	+	−	−	−	−	+	+	−	+	+	−	−	+	+	+	+	−	+
11	+	−	+	−	−	−	−	+	+	−	+	+	−	−	+	+	+	+	−
12	−	+	−	+	−	−	−	−	+	+	−	+	+	−	−	+	+	+	+
13	+	−	+	−	+	−	−	−	−	+	+	−	+	+	−	−	+	+	+
14	+	+	−	+	−	+	−	−	−	−	+	+	−	+	+	−	−	+	+
15	+	+	+	−	+	−	+	−	−	−	−	+	+	−	+	+	−	−	+
16	+	+	+	+	−	+	−	+	−	−	−	−	+	+	−	+	+	−	−
17	−	+	+	+	+	−	+	−	+	−	−	−	−	+	+	−	+	+	−
18	−	−	+	+	+	+	−	+	−	+	−	−	−	−	+	+	−	+	+
19	+	−	−	+	+	+	+	−	+	−	+	−	−	−	−	+	+	−	+
20	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−

A small explanation is necessary in order to understand the reading of this design. For example, run n° 1 is described by conditions for grouped factors which are:

N°Exp	G_1	G_2	G_3	G_4	G_5	G_6	G_7	G_8	G_9	G_{10}	G_{11}	G_{12}	G_{13}	G_{14}	G_{15}	G_{16}	G_{17}	G_{18}	G_{19}
1	+	+	−	−	+	+	+	+	−	+	−	+	−	−	−	−	+	+	−

This run is characterized by conditions for each factor which are:

	G_1 +					G_2 +					G_3 −					G_4 −					
N°Exp	X_1	X_5	X_{12}	X_{16}	X_{17}	X_{20}	X_{21}	X_{22}	X_{23}	X_{24}	X_{25}	X_{26}	X_{27}	X_{28}	X_{29}	X_{30}	X_{31}	X_{32}	X_{33}	X_{35}	…
1	+	+	+	+	+	+	+	+	+	+	−	−	−	−	−	−	−	−	−	−	…

From the results of the 20 simulations, an estimation of the coefficients was obtained; the interpretation of the results reveals 2 active groups (G_7, G_8) and according to the group screening method, the10 factors within these groups (U_2, U_3, U_4, U_6, U_{10}, U_{11}, U_{13}, U_{18}, U_{19}, U_{45}) are individually studied in the second step with an Hadamard design involving 12 experiments.

From this screening design, the estimation of the coefficients (b_1, b_2, ..., b_{10})was obtained and the Bayesian approach (Bernardo and Smith, 1994) (Figure 6) shows a non-influence of the factors U_2, U_4, U_{10} and U_{19}; the 6 others variables can be classified on order of importance:**U_{13}, U_{45}, U_3, U_6, U_{11}** and **U_{18}**.

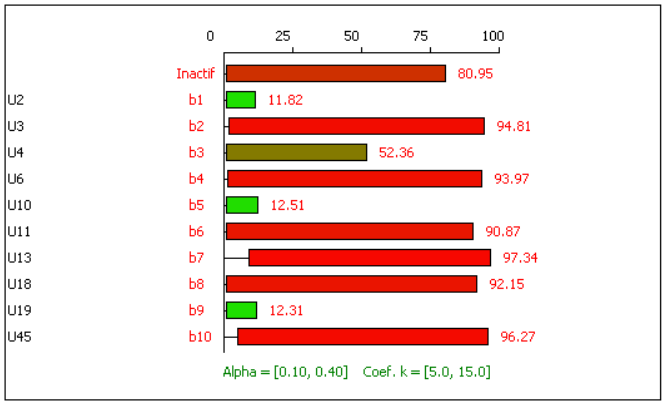


Figure 6. Bayesian plot for the response "production at 2000 days×10^{-3}" for the Hadamard design in 12 experiments in the second step of the group screening method.

3.2.4. Multiple Group screening

In the first stage, the 51 factors were grouped into 7 groups of type 1 (G_1, G_2, ... , G_7) and 7 groups of type 2 (H_1, H_2, ... H_7) such that the intersection of any group of type 1 with any group of type 2 contains one factor, except for the intersection G_6/H_6 and G_7/$H_7$with 2 factors. This arrangement is depicted graphically in Table 3.

To study the type-1 groups (G) and the type-2 groups (H), two Hadamard designs with 8 experiments were performed. The interpretation of the results reveals 3 active type-1 groups (G_1, G_2 and G_3), containing 21 factors and 2 active type-2 groups (H_2 and H_3), containing 14 factors. The factors retained for the second stage are those for which the two types of groups are apparently active: U_3, U_4, U_{12}, U_{13}, U_{19} and U_{45}.

In the second stage, the 6 factors are individually studied with a non-regular Hadamard design in 12 experiments. The estimation of the effects of these factors (b_1, b_2, ... , b_6) are obtained and the effect plot (Figure 7) reveals 3 active factors: U_3, U_{13} and U_{45}.

Table 3. Arrangement of the multiple groups

7 groups of type 1

7 groups of type 2	G_1	G_2	G_3	G_4	G_5	G_6	G_7
H_1	U_2	U_{18}	U_{11}	U_{10}	U_{29}	U_{35}	U_{40}
H_2	U_3	U_{19}	U_{13}	U_{28}	U_{30}	U_{36}	U_{41}
H_3	U_4	U_{45}	U_{12}	U_{27}	U_{31}	U_{37}	U_{42}
H_4	U_6	U_1	U_{16}	U_{26}	U_{32}	U_{38}	U_{43}
H_5	U_{17}	U_5	U_{24}	U_{25}	U_{33}	U_{39}	U_{51}
H_6	U_{49}	U_{20}	U_{23}	U_7	U_9	U_{34} U_{44}	U_{46}
H_7	U_{50}	U_{21}	U_{22}	U_8	U_{14}	U_{15}	$U_{47}U_{48}$

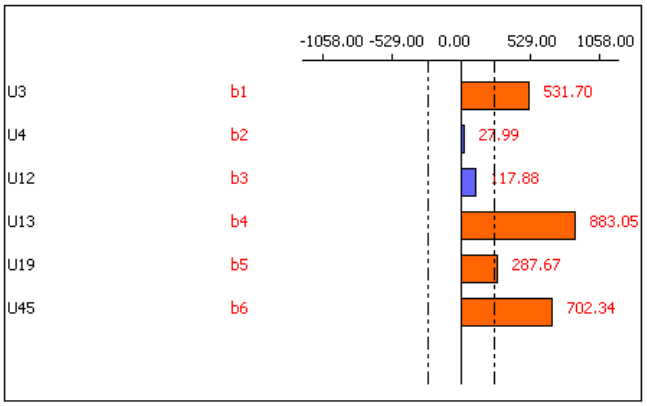
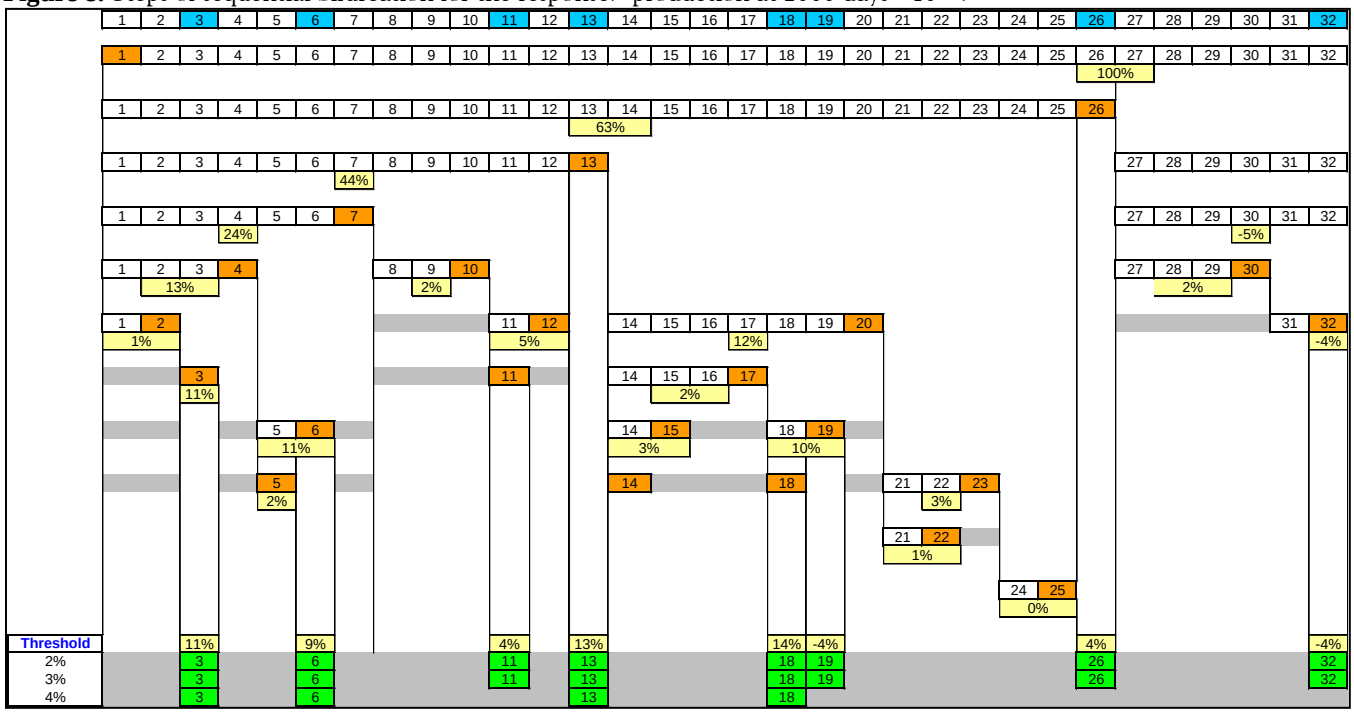


Figure 7. Effects plot for the response "production at 2000 days × 10^{-3}" for the Hadamard design in 12 experiments in the second step of the multiple group screening method.

3.3. Sequential bifurcation

The method of sequential bifurcation was applied in order to study the 51 factors. We begin sequential bifurcation by calculating the response when all the factors are their low levels, $y_{(0)} = 1{,}516.920$, and when all the factors are at their high level, $y_{(51)} = 7{,}088.965$. So the estimated effect of all 51 factors aggregated together is obtained from $y_{(51)} - y_{(0)} = 5{,}571.445$, considered as a variation of 100 %. The next step is to divide the current group of 51 factors into two equal subgroups, hence the term symmetric bifurcation, and to compute $y_{(26)}$ by fixing the first 26 factors at their low level, while the 25 others factors are set at their high level. At this stage, $y_{(26)}$ is compared to $y_{(0)}$ and $y_{(51)}$ and the corresponding variation percentage of the response is calculated (Figure 8). The bifurcation continues by splitting active subgroups into smaller subgroups and eliminating non-active subgroups until the percentage of variation becomes non-significant. Figure 8 shows the different steps of the sequential bifurcation leading to the identification of 5 active factors $\mathbf{U_3}$, $\mathbf{U_6}$, $\mathbf{U_{13}}$, $\mathbf{U_{18}}$, $\mathbf{U_{45}}$ and 5 factors to a lesser extent : $\mathbf{U_{11}}$, $\mathbf{U_{19}}$, $\mathbf{U_{26}}$, $\mathbf{U_{32}}$, $\mathbf{U_{33}}$, in 29 simulations.

Figure 8. Steps of sequential bifurcation for the response: "production at 2000 days$\times 10^{-3}$".



METHODS	Number of experiments
OFAT	102
HADAMARD R III	56
HADAMARD R IV	112
SUPERSATURATED	28
GROUP SCREENING	32
MULTIPLE GROUP SCREENING	28
SEQUENTIAL BIFURCATION	29

Table 4. Synthesis of results obtained with the different screening methods

Conclusion

This first complete application of numerical simulation allowed us to compare a large range of screening methods, which can be successful if the assumptions associated to each of them are respected, that is:

- Additivity assumption for the Hadamard R_{III} designs,
- Very few active factors for the supersaturated designs ("sparsity effect"),
- Knowledge on the direction of possible effects in order to suitably attribute the levels, for the group screening and the sequential bifurcation methods.

To conclude, a summary establishes a synthesis of the results (Table 4) with the different methods and from this table, it can be noted that:

- All methods, even with very few experiments (N=28, N=29), detect the most important factors : U_3, U_{13}, $U_{45,}$
- The Plackett and Burmann R_{IV} designs, with more experiments (N=112) identify more precisely the active factors, by quantification of the effects, unbiased by two-factor interactions,
- Sequential bifurcation inevitably leads to the identification of the active factors, but this approach requires knowing, *a priori*, the direction of the effects.

These conclusions on the comparison of different screening methods underline the importance of the underlying assumptions. In practice, the choice of the method depends on the objective of the study, on the possible number of simulations and on the risk (false positive or false negative) the experimenter will accept. It is obvious that the classical Plackett and Burmann R_{IV} designs lead to most reliable conclusions and if the experimenters do not want to take any risk, this is the best strategy. Nevertheless, the Plackett and Burmann R_{III} designs seem to be a good compromise in order to decrease the number of simulations without necessarily going to supersaturated designs and without any assumptions imposed on the direction of the effects.

Acknowledgements: This study has been supported by DICE Consortium (Deep Inside Computer Experiments).

References

Abraham, B., Chipman, H., 1999. Some risks in the construction and analysis of supersatured designs. *Technometrics* 41,135-141.

Akaike, H., 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 716-723.

Azarian, A., Vasseur, O., Bennaï, B., Jolivet, V., Claeys-Bruno, M., Sergent, M., Bourdon, P., 2011.Sensitivity analysis of interference optical systems by numerical designs. *Journal de la SFdS* 152, 1, 118-130.

Barton, R. 2001. Designing simulation experiments, 2001 Winter Simulation Conference. in: Peters, B.A., Medeiros, D.J., Rohrer, M.W., (Eds), pp 47-52.

Bernardo, J., Smith, A. F. M., 1994. Bayesian Theory. John Wile, New York.

Beth, T., Jungnickel, D., Lenz, H.,1999. *Design Theory*. Cambridge.

Bettonvil, B., 1990. Detection of important factors by sequential bifurcation. University Press ed., Tilburg.

Bettonvil, B., Kleijnen, J. P. C., 1997. Searching for important factors in simulation models with many factors: Sequential bifurcation. *European Journal of Operational Research* 96, 180-194.

Booth, K. H. V., Cox, D. R., 1962. Some systematic supersaturated designs. *Technometrics* 4, 489-495.

Butler, N. A., Mead, R., Eskridge, K. M., Gilmour, S. G., 2001. A general method of constructing $E(s^2)$-optimal supersaturated designs. *Journal of the Royal Statistical Society* B 63, 621-632.

Catterjee, K., Gupta, S., 2003. Construction of supersaturated designs involvings-level factors. *Journal of statistical planning and inference* 113, 589-595.

Cheng, C. S., 1997. E(s2)-optimal Supersaturated designs. Statistica Sinica 7, 929-939.

Cela, R., Phan-Tan-Luu, R., Claeys-Bruno, M., 2009.*Screening Strategies.* In Comprehensive Chemometrics, Elsevier, Oxford.

Claeys-Bruno, M., Dobrijevic, M., Phan-Tan-Luu, R., Sergent, M., 2009. A new class of supersaturated design: application to a sensitivity study of a photochemical model. *Chemometrics and Intelligent laboratory systems*95, 86-93.

Claeys-Bruno, M., Cela, R., Phan-Tan-Luu, R., Sergent, 2011. Supersaturated designs for computer experiments: Comparison of construction methods and new methods of treatment adapted at the high dimension problem. *Chemometrics and Intelligent Laboratory Systems* 105, 137-146.

Deng, L.-Y., Lin, D. K. J., Wang, J., 1999. A resolution rank criterion for supersaturated designs. *Statistica Sinica* 9, 605-610.

Dorfman, R., 1943. The detection of defective members of large populations. Annals of *Mathematical Statistics* 14, 436-440.

Finucan, H. M., 1964. The blood testing problem. *Applied Statistics* 13, (43-50).

Hunter, W. G., Mezaki, R., 1964. Catalyst selection by group screening. *Industrial et Engineering Chemistry* 56, (3), 38-40.

Kleijnen, J. P. C., 2004. An overview of the design and analysis of simulation experiments for sensitivity analysis. *European Journal of Operational Research*.

Lenth, R. V., 1989. Quick and easy analysis of unreplicated factorials. *Technometrics* 31, (4), 469-473.

Lin, D. K. J., 1993. A New class of Supersaturated designs. *Technometrics* 35, 28-31.

Lin, D. K. J., 1995. Generating systematic supersaturated designs. *Technometrics* 37, 213-225.

Liu, M.-Q., Hickernell, F. J., 2002. E(s^2)-optimality and minimum discrepancy in 2-level supersaturated designs. *Statistica Sinica* 12, 931-939.

Liu, Y., Dean, A., 2004. k-Circulant Supersaturated Designs. *Technometrics* 43, 32-42.

Lu, X., Meng, Y., 2000. A new method in the construction of two-level supersaturated design. *Journal of Statistical Planning and Inference* 86, 229-238.

Lu, X., Wu, X., 2004. A Strategy of Searching Active Factors in Supersaturated Screening Experiments. *Journal of Quality Technology* 36, (4), 392-399.

Mathieu, D., Nony, J.,Phan-Tan-Luu, R. 2007. NEMROD-W software, LPRAI,www.NemrodW.com: Marseille.

Morris, M. D., 1987. Two-stage factor screening procedures using multiple grouping assignments. Communications in statistics, *Theory and methods* 16, (10), 3051-3067.

Nguyen, N.-K., 1996. An Algorithmic Approach to Constructing Supersaturated Designs. *Technometrics* 38, 69-73.

Patel, M. S., 1962. Group-screening with more than two stages. *Technometrics* 4, 209-217.

Plackett, R. L., Burman, J. P., 1946. Design of Optimum Multifactorial Experiments. *Biometrika* 33, 305-325.

Sanchez, S. M., Wang, H., Lucas, T. W. 2005. In A two-phase screening procedure for simulation experiments, 2005 Winter Simulation Conference, Kulh M.K., Armstrong F.B., Joines J.A. (Eds).

Satterwaithe, F., 1959. Random balance experimentation (with discussion). *Technometrics* 1, 111-137.

Schwarz, G., 1978. Estimating the dimension of a model. *Annals of Statistics* 6, 461-464.

Tang, B., Wu, C. F. J., 1977. A method for constructing supersatured designs and its E(s2) optimality. *Canadian Journal of Statistics* 25, 191-201.

Watson, G. S., 1961. A study of the group screening method. *Technometrics* 3, (3), 371-388.

Welch, W. J., Buck, R. J, Sacks, J., Wynn, H. P., Mitchell, T. J., Morris, M. D., 1992. Screening, predicting and computer experiments. *Technometrics* 34, 15-25.

Wu, C. F. J., 1993. Construction of supersaturated designs through partially aliased interactions. *Biometrika* 80, 661-669.

Yamada, S., Lin, D. K. J., 1997. Supersaturated designs including an orthogonal base. *Canadian Journal of Statistics* 25, 203-213.

Correspondence: m.claeys-bruno@univ-amu.fr