TECHNICAL REPORT

ISO/TR 29901

First edition
2007-11-01

Selected illustrations of full factorial experiments with four factors

Illustrations choisies de plans d'expérience factoriels complets à quatre facteurs

Reference number
ISO/TR 29901:2007(E)

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Contents

Page

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

In exceptional circumstances, when a technical committee has collected data of a different kind from that which is normally published as an International Standard ("state of the art", for example), it may decide by a simple majority vote of its participating members to publish a Technical Report. A Technical Report is entirely informative in nature and does not have to be reviewed until the data it provides are considered to be no longer valid or useful.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

ISO/TR 29901 was prepared by Technical Committee ISO/TC 69, *Applications of statistical methods*.

Introduction

The Six Sigma and international statistical standards communities share a philosophy of continuous improvement and many analytical tools. The Six Sigma community tends to adopt a pragmatic approach driven by time and resource constraints. The statistical standards community arrives at rigorous documents through long-term international consensus. The disparities in time pressures, mathematical rigor and statistical software usage have inhibited exchanges, synergy and mutual appreciation between the two groups.

The present document takes one specific statistical tool (full factorial designs with four factors, 2^4 designs) and develops the topic somewhat generically (in the spirit of International Standards) but then illustrates it through the use of five detailed and distinct applications. The generic description focuses on the commonalities across 2^4 designs. These commonalities hold more generally for arbitrary numbers of factors, but a value of four was chosen for this Technical Report. The annexes containing the five illustrations follow the basic framework but also identify the nuances and peculiarities in the specific applications. Each example offers at least one “wrinkle” to the problem, which is generally the case for real Six Sigma applications. It is thus hoped that practitioners can identify with at least one of the five examples, if only to remind them of the basic material on factorial designs that was encountered during their Six Sigma training. Each of the five examples is developed and analysed using statistical software of current vintage. The explanations throughout are devoid of mathematical detail — such material can be readily obtained from the many design and analysis of experiments textbooks available (such as those given in the Bibliography).

Selected illustrations of full factorial experiments with four factors

1 Scope

This Technical Report describes the steps necessary to specify, to use and to analyse 2^4 full factorial designs through illustration, with five distinct applications of this methodology.

Depending on the application, a number of factors other than four may be considered in the experiment.

NOTE 1 Each of these five illustrations is similar in that sufficient resources were available to implement the design. Other commonalities among the five examples are noted (e.g. study objective, two levels for factors, response variable(s), factors effecting the response). The individual illustrations have some salient features that are distinct such as presence/absence of repetitions, centre points, interactions, or different types of response variables. Each illustration takes place in a different environment such as marketing, software, manufacturing, telecommunications and chemical processing.

NOTE 2 For the purposes of this Technical Report, the selection of four factors with two levels (aside from centre points) was made in advance. Furthermore, the detailed use of response surface designs as a follow-up or augmentation of the existing designs was excluded from this Technical Report, although their use is noted in some of the illustrations. Likewise, Taguchi designs and blocking designs were not included.

NOTE 3 Full factorial experiments are often employed by individuals (so-called “black belts” or “green belts”) associated with Six Sigma methods. Six Sigma methods are concerned with problem solving and continuous improvement. A full factorial experiment with four factors is one of many tools available to Six Sigma practitioners, but hitherto has not been addressed in detail in ISO International Standards.

2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 3534-1:2006, *Statistics — Vocabulary and symbols — Part 1: General statistical terms and terms used in probability*

ISO 3534-2, *Statistics — Vocabulary and symbols — Part 2: Applied statistics*

ISO 3534-3:1999, *Statistics —Vocabulary and symbols — Part 3: Design of experiments*

3 Terms and definitions

For the purposes of this document, the terms and definitions in ISO 3534-1, ISO 3534-2, ISO 3534-3 and the following apply.

3.1
analysis of variance
ANOVA
technique which subdivides the total variation of a response variable into meaningful components associated with specific sources of variation

NOTE Adapted from ISO 3534-3:1999, definition 3.4.

3.2
binomial distribution
discrete distribution having the probability mass function

$$P(X = x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

where $x = 0, 1, \ldots, n$ and with indexing parameters $n = 1, 2, \ldots$, and $0 < p < 1$.

NOTE Adapted from ISO 3534-1:2006, definition 2.46.

3.3
block
collection of experimental units more homogeneous than the full set of experimental units

NOTE Adapted from ISO 3534-3:1999, definition 1.11.

3.4
centre point
vector of factor level settings of the form (a_1, a_2, ..., a_k), where all a_i equal 0, as notation for the coded levels of the factors

NOTE Adapted from ISO 3534-3:1999, definition 1.36.

3.5
design matrix
matrix with rows representing individual treatments (possibly transformed according to the assumed model) which can be extended by deduced levels of other functions of factor levels (interactions, quadratic terms, etc.) but are dependent upon the assumed model

NOTE Adapted from ISO 3534-3:1999, definition 2.7.1.

3.6
factor
predictor variable that is varied with the intent of assessing its effect on the response variable

NOTE Adapted from ISO 3534-3:1999, definition 1.5.

3.7
full factorial experiment
experiment consisting of all possible treatments formed from two or more factors, each being studied at two or more levels

NOTE Adapted from ISO 3534-3:1999, definition 2.1.

3.8
interaction
effect for which the apparent influence of one factor on the response variable depends upon one or more other factors

NOTE Adapted from ISO 3534-3:1999, definition 1.17.

3.9
level
potential setting, value or assignment of a factor

NOTE Adapted from ISO 3534-3:1999, definition 1.6.

3.10
normal distribution
continuous distribution having the probability density function

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

where $-\infty < x < \infty$ and with parameters $-\infty < \mu < \infty$ and $\sigma > 0$

NOTE Adapted from ISO 3534-1:2006, definition 2.50.

3.11
predictor variable
variable that can contribute to the explanation of the outcome of an experiment

NOTE Adapted from ISO 3534-3:1999, definition 1.3.

3.12
randomization
process used to assign treatments to experimental units so that each experimental unit has an equal chance of being assigned a particular treatment

NOTE Adapted from ISO 3534-3:1999, definition 1.29.

3.13
replication
performance of an experiment more than once for a given set of predictor variables

NOTE Adapted from ISO 3534-3:1999, definition 1.27.

3.14
split-plot design
design in which a group of experimental units (plot) to which the same level assigned to the principal factor is subdivided (split) so as to study one or more additional principal factors within each level of that factor

NOTE Adapted from ISO 3534-3:1999, definition 2.3.6.

4 Symbols and abbreviated terms

The symbols and abbreviated terms used in this Technical Report are as follows:

y	Response variable
A, B, C, D	Factors
AB, AC, AD, BC, BD, CD	2-way interactions
ABC, ACD, BCD	3-way interactions
ABCD	4-way interactions
+1/–1	High and low settings
2^4	Four factors each with two levels
σ	Standard deviation

5 Generic description of full factorial designs

5.1 Overview of the structure of the four factor examples in Annexes A through E

This Technical Report provides general guidelines on the design, conduct and analyses of two-level full factorial designs and illustrates the steps with five distinct applications given in Annexes A through E. Each of these five examples follows the basic structure given in Table 1.

The steps given in Table 1 apply to design and analysis of experiments in general, although this Technical Report focuses on 2^4 full factorial designs. Each of the seven steps is explained in general below. Specific explanations of the substance of these steps is provided in the examples in Annexes A through E.

Table 1 — Basic steps in experimental design

1	State the overall objective(s) of the experiment
2	Describe the response variable(s)
3	List the factors that might affect the response(s)
4	Select a “full” factorial design
5	Analyse the results – Numerical summaries and graphical displays
6	Present the results
7	Perform a confirmation run

5.2 Overall objective(s) of the experiment

Experiments are conducted for a variety of reasons. The primary motivation for the experiment should be clearly stated and agreed to by all parties involved in the design, conduct, analysis and implications of the experimental effort. There may be secondary objectives which could be addressed with the full factorial experiment.

The ultimate outcome of the experiment could be to take immediate action on factor levels or to obtain a predictive model, both of which dictate some elements of the analyses.

5.3 Response variable(s)

Associated with the objective of an experiment is a measurable outcome or performance measure. A response of interest could involve maximization (larger is better), minimization (smaller is better) or meet a target value (be close to a specified value). The response variable (denoted here by the variable y) should be intimately (if not directly) related to the objective of the experiment. For some situations, there may be multiple characteristics of interest to be considered, although there typically is a primary response variable associated with the experiment. In other cases, multiple responses must be considered; however, for purposes of this document, a single response is considered in each example.

5.4 Factors affecting the response(s)

The response variable likely depends in some unknown way on a variety of conditions that occur or could be set in the course of generating a response variable outcome. These conditions are presumed to relate to controllable factors that may be continuous (temperature, concentration) or discrete (two assembly lines A or B, two vendors, two packaging styles, and so forth). For 2^4 experiments, we simplify the experimental design process by selecting two levels for each factor to be varied in the experiment. For discrete factors with only two possible settings, the levels of this factor are just these two settings. For continuous factors, there is discretion in choosing the two specific values. In some cases, the two settings could be the historical value and a proposed value. In other cases, the two settings could be a nominal adjustment from the historical setting. In any event, the settings should be sufficiently far removed to have an opportunity to reveal an impact subject to the inherent uncertainty, while not being so disparate that the settings are unreasonable from a practical, safety or sensibility standpoint. The setting of levels of continuous factors benefits from the domain expert collaborating on the experiment.

There may be additional factors that could impact that response variable but may be deemed less important than the chosen four factors or are too difficult or expensive to control. Finally, it is the case that the factors are to be set independently from each other. It could, however, be discovered that the factors interact (the setting of one factor impacts how a second factor affects the response).

5.5 “Full” factorial design

This classical design consists of 16 runs obtained by considering all combinations of four factors with two possible levels. Table 2 provides the basic layout in a standard order for ease of understanding. Each row of the table represents one set of experimental conditions that when run will produce a value of the response variable y. The four factors are designated as A, B, C and D. For an individual factor, the level “−1” is the “low” setting, or one of the two levels if the factor is categorical. The level “+1” is the “high” setting, or the other level of the categorical factor. The column “y” is a placeholder for the response value once a run has occurred.

In addition to providing the explicit 16 experimental runs to be conducted, Table 2 provides an example of a randomized run order. The experiment should not be run in the standard order (for example, an increasing trend in the response could be confused with the effect of factor D on the response). The final two columns provide abbreviated names for the 16 runs. The first name coincides with the “+” or “–” levels of the factors given in the order for ABCD. The second naming convention includes a lower case letter for each factor at its high level. By convention, the four settings at the low level are designated as “(1)” since otherwise the abbreviation would be null.

Table 2 — Layout of a generic 2^4 full factorial design

Row number	A	B	C	D	y	Run order	Name of run	Alternate name of run
1	–1	–1	–1	–1	y_1	6	− − − −	(1)
2	+1	–1	–1	–1	y_2	14	+ − − −	a
3	–1	+1	–1	–1	y_3	4	− + − −	b
4	+1	+1	–1	–1	y_4	11	+ + − −	ab
5	–1	–1	+1	–1	y_5	9	− − + −	c
6	+1	–1	+1	–1	y_6	2	+ − + −	ac
7	–1	+1	+1	–1	y_7	3	− + + −	bc
8	+1	+1	+1	–1	y_8	1	+ + + −	abc
9	–1	–1	–1	+1	y_9	8	− − − +	d
10	+1	–1	–1	+1	y_{10}	13	+ − − +	ad
11	–1	+1	–1	+1	y_{11}	7	− + − +	bd
12	+1	+1	–1	+1	y_{12}	10	+ + − +	abd
13	–1	–1	+1	+1	y_{13}	15	− − + +	cd
14	+1	–1	+1	+1	y_{14}	16	+ − + +	acd
15	–1	+1	+1	+1	y_{15}	5	− + + +	bcd
16	+1	+1	+1	+1	y_{16}	12	+ + + +	abcd

In the situation with all four factors being continuous (quantitative), some experimenters elect to run additional “centre” points. The levels of the centre point are at the midpoint of the two levels of each factor. The inclusion of centre points facilitates statistical testing of effects, using the variability of the responses at the centre points as a guideline. It also provides the means to test for non-linearity (curvature) of the response.

In the situation where a particular factor is difficult or expensive to change from one level to the other, the experiment could be conducted in “blocks”, one block for each level of the difficult factor.

One final situation concerns replication of the runs in the design. Multiple replications at each experimental setting accommodate estimates of the overall inherent variability which can then be pooled for statistical testing purposes.

5.6 Analyse the results — Numerical summaries and graphical displays

At the completion of the conduct of the experiment, the y_i values would be replaced by the actual observed responses. Many existing statistical software packages exist to facilitate the generation of output to aid in the understanding of the results of the experiment. Of immediate concern is the determination of the impact of the four factors individually on the response variable. Thus, the main effects for A, B, C and D are to be estimated as well as the following interaction terms:

- two-way: AB, AC, AD, BC, BD, CD;
- three-way: ABC, ABD, ACD, BCD;
- four-way: ABCD.

In addition to the actual estimates, it is useful to arrange the estimated effects from largest to smallest in the form of a Pareto Chart. The effects can also be presented on a normal or half-normal plot to identify the stronger effects. Depending on the inclusion of centre points or replication, the experimental error can be estimated directly and in turn the standard error of the effects can be determined. In the absence of these experimental runs, the experimental error can be estimated indirectly by presuming that three-way and four-way interactions are negligible. Effects plots, interaction plots and residual plots are also generally considered in assessing the results of an experiment. These are illustrated in Annexes A through E.

5.7 Present the results

Frequently, the purpose of the experiment is to develop a predictive model so as to explore alternate settings of the factors. Contour plots are available from most software packages. Moreover, with a predictive model, tentative consideration of optimal settings can be identified. The predictive model will consist of a function (typically linear) of the estimated effects and possibly interactions. Alternate mathematical models may be envisaged depending on examination of the residuals (collection of predictive values minus observed values).

5.8 Perform confirmation runs

Follow-up experiments may prove useful to demonstrate that the lessons learned from the 2^4 full factorial experiments are verified in subsequent runs. A natural follow-on experiment could be to identify a promising direction in which to adjust the factors for improved performance in the response variable.

6 Description of Annexes A through E

6.1 Comparing and contrasting the examples

Five distinct examples of 2^4 full factorial designs are illustrated in Annexes A to E. Each of these examples follows the same general template as given in Table 1 and follows a version of the standard design given in Table 2.

6.2 Experiment summaries

Table 3 summarizes the five examples detailed in the annexes and indicates aspects of the analyses which were unique to each experiment.

Table 3 — Experiment summaries found by annex

Annex	Experiment	Problem-specific aspects
A	Solder bars	Mixture of discrete and continuous factors; repetitions; important 2-way interaction
B	Direct mail campaign	Proportion response variable, standard errors based on binomial distribution
C	Button tactility	Centre points; curvature in response
D	PVC formulation optimization	Split-plot; centre points; contour plots
E	Genetic algorithms	Randomized order not necessary; set up for interim analyses; replication; dispersion effect

Annex A
(informative)

Solder bar experiment[1)]

A.1 General

The material solder can be produced in a number of forms. One of the most common is a bar of solid solder. The bars are sold by mass, namely 0,5 kg, 1,0 kg and 2,0 kg.

A.2 Overall objective for the experiment

The process had recently been relocated to a new site and subsequently, on the underside of the bars there were a large number of “rosettes” (small pits on the surface of the solder bar). This led to numerous customer complaints, customers assuming that the solder was of an inferior standard. The current level of rosettes on bars was assessed to be about 30 per bar.

The operational staff had tried over several months various “one-factor-at-a-time” experiments, none of which produced a large enough improvement for the producer to say to their customers that they had solved the problem. It was decided to design and then run an experiment to understand the factors responsible for the defect and then how to manage the process so as to minimize the number of rosettes.

A.3 Description of the process

Solder bars are produced by pouring molten solder into moulds pre-selected according to the target mass of the bars to be produced.

Ingots of solder are firstly melted in an electric furnace at about 290 °C. An operator then produces the bars by filling a “kettle” with the molten solder and then by pouring the solder from the kettle into moulds. A kettle is similar to the kitchen utensil of the same name but with extra spouts, thus enabling several bars to be poured simultaneously.

The bars are cast upside down so the bottom of each bar is uppermost and open to the atmosphere. The moulds are arranged in banks on a table with water-cooling integrated into it that can be turned on or off. Before casting, the moulds can be “smoked”. This is a layer of carbon applied with an oxyacetylene torch.

After a short period of time during which the bars solidify, they are de-moulded, stacked and then packed.

A.4 Response variable

A.4.1 Choice of variable

The response chosen for the experiment was the average number of rosettes counted on the surface of the bars. The objective was to minimize the average number of rosettes.

1) This example has been kindly donated by Cookson Electronics (Fry's Metals Glasgow).

A.4.2 Measurement of the response variable

Approximately 160 bars were produced in each experimental “run”. The number of rosettes was counted across all of the solder bars and the arithmetic mean of the number of rosettes per bar calculated for each experimental run. This became the output for that experimental run.

A.4.3 Relationship of the response variable to the objective of the experiment

The mean number of rosettes per bar is directly linked to the objective of the experiment.

A.5 Factors affecting the response

A.5.1 Description of each factor (continuous/discrete) to be varied

The factors chosen to be varied within the experiment were determined using the knowledge of the production operators and the technical staff. These were:

A: the casting temperature of the molten solder;

B: water-cooling, applied or not;

C: the pouring rate of the molten solder into the moulds; and

D: whether the moulds were smoked or not.

The experiment was to be run using these four factors each at two levels.

Factor A was clearly a factor that could be adjusted to any given temperature. Factor C was, in theory, one that could be varied as a continuous variable. However, in practice, it was handled as a categorical factor and the rate was either “Normal” or “Maximum”.

The other two factors were categorical. Water cooling, factor B, was either “Off” or “On” and the moulds, factor D, were either not smoked “No” or smoked “Yes”.

A.5.2 Selection of levels (related to size of effect to be determined)

The factors used in the experiment and their associated levels are given in Table A.1.

Table A.1 — Factors and their associated levels

Factor	Level 1	Level 2
A: Casting temperature (°C)	260	320
B: Water cooling	Off	On
C : Pouring rate	Normal	Maximum
D: Mould conditioning	No	Yes

The casting temperatures selected were determined to be the extremes that would be considered during production. This was a result of discussions with the technical and metallurgical staff.

The levels for the remaining factors were deemed to be discrete, e.g. “On” or “Off”.

Prior to the experiment, it was really not known if the selected factors would have any effect on the response, and so this might be regarded as a “screening” experiment. The considered opinion of the technical department was that these factors were very likely to be linked to the response and that the levels chosen for the casting temperature were wide enough to generate an observable change in the response.

A.5.3 Other factors noted but not incorporated due to issues with controllability or relevance

Other factors that might have been considered, but were not included in the experiment, are given in Table A.2.

Table A.2 — Factors not included in the experiment

Factor	Reason for exclusion
Air temperature	Air temperature was not regarded as important to the response and the factory was not temperature controlled and this would have been very difficult to control.
Humidity	Considered irrelevant and the factory did not posses air conditioning so therefore it would have been impossible to control.
Air cleanliness	The atmosphere within the factory was naturally smoky and it would have been impossible to control the level of cleanliness.
Operator	There was only one operator involved with this process and so the inclusion of different operators for this experiment was seen as irrelevant.
Day	The day of the week was regarded as irrelevant for this experiment.

A.6 Full factorial design

A.6.1 Design matrix (standard order, run order)

The design selected was a full factorial that allowed for all interactions to be examined. Table A.3 shows the design.

Table A.3 — Design matrix

A: Casting temperature °C	B: Water cooling	C: Pouring rate	D: Mould conditioning	Standard order	Run order
320	On	Maximum	No	8	1
320	Off	Maximum	No	6	2
260	On	Maximum	No	7	3
260	On	Normal	No	3	4
260	On	Maximum	Yes	15	5
260	Off	Normal	No	1	6
260	On	Normal	Yes	11	7
260	Off	Normal	Yes	9	8
260	Off	Maximum	No	5	9
320	On	Normal	Yes	12	10
320	On	Normal	No	4	11
320	On	Maximum	Yes	16	12
320	Off	Normal	Yes	10	13
320	Off	Normal	No	2	14
260	Off	Maximum	Yes	13	15
320	Off	Maximum	Yes	14	16

The experiment was randomized and the column headed "Run order" was the order in which the runs were performed during the experiment. The column "Standard order" specifies where the particular factor combination would be found in the "design matrix" for the 4-factor 2-level design.

A.6.2 Centre points

Centre points were not selected for this experiment, partly because three of the four factors were taken as categorical (binary) factors, making centre points inconvenient to arrange.

A.6.3 Replication and repetition

The factory could not afford any more experimental time than that needed to conduct the 16 run experiment shown above. Therefore, there was no replication of any of the experimental runs. However, once an experimental run had been set up, about 160 bars were produced, providing abundant repetition.

Repetition is important to overcome effects of measurement uncertainty and to improve the reliability of the analysis. The decision taken to average the number of rosettes was seen as appropriate in this case. At other times different decisions might be taken, e.g. to calculate a standard deviation from the repeated measurements for each replicated run. These repetitions provide the analyst with information about factors that might influence the process variability.

A.7 Analysis of results

A.7.1 Data acquired through the experiment and analysis considerations

The results from the experiment are given in Table A.4. The response analysed is the mean number of rosettes per bar per run. They were keyed into a software program called MINITAB™[2] .

Table A.4 — Experiment results

A: Casting temperature °C	B: Water cooling	C: Pouring rate	D: Mould conditioning	Standard order	Run order	Mean rosettes per bar
320	On	Maximum	No	8	1	85,1
320	Off	Maximum	No	6	2	77,9
260	On	Maximum	No	7	3	89,1
260	On	Normal	No	3	4	87,7
260	On	Maximum	Yes	15	5	92,2
260	Off	Normal	No	1	6	100,4
260	On	Normal	Yes	11	7	84,7
260	Off	Normal	Yes	9	8	75,4
260	Off	Maximum	No	5	9	84,5
320	On	Normal	Yes	12	10	15,3
320	On	Normal	No	4	11	80,6
320	On	Maximum	Yes	16	12	8,1
320	Off	Normal	Yes	10	13	1,0
320	Off	Normal	No	2	14	84,8
260	Off	Maximum	Yes	13	15	87,6
320	Off	Maximum	Yes	14	16	2,3

A decision was taken that only the first-order interactions, e.g. AB, would be analysed, together with the main effects. Therefore, the higher order interactions were initially considered to be unimportant and insignificant (subsequently confirmed).

2) MINITAB™ is the trade name of a product supplied by Minitab Inc. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of the product named. Equivalent products may be used if they can be shown to lead to the same results.

A.7.2 Estimation of effects

```
Estimated Effects and Coefficients for Mean rosettes

Term                               Effect    Coef  SE Coef       T      P
Constant                                    66.04    1.606   41.14  0.000
Casting temperature (deg C)        -43.31  -21.66    1.606  -13.49  0.000
Water cooling                        3.61    1.81    1.606    1.13  0.312
Pouring rate                        -0.39   -0.19    1.606   -0.12  0.909
Mould conditioning                 -40.44  -20.22    1.606  -12.59  0.000
Casting temperature (deg C)*         2.16    1.08    1.606    0.67  0.531
  Water cooling
Casting temperature (deg C)*        -1.69   -0.84    1.606   -0.53  0.622
  Pouring rate
Casting temperature (deg C)*       -34.99  -17.49    1.606  -10.90  0.000
  Mould conditioning
Water cooling*Pouring rate           1.94    0.97    1.606    0.60  0.573
Water cooling*Mould conditioning     4.89    2.44    1.606    1.52  0.188
Pouring rate*Mould conditioning      3.84    1.92    1.606    1.20  0.286

S = 6.42204   R-Sq = 98.94%   R-Sq(adj) = 96.81%

Analysis of Variance for Mean rosettes

Source              DF   Seq SS   Adj SS   Adj MS      F      P
Main Effects         4  14097.5  14097.5  3524.36  85.45  0.000
2-Way Interactions   6   5096.1   5096.1   849.34  20.59  0.002
Residual Error       5    206.2    206.2    41.24
Total               15  19399.7
```

Figure A.1 — Estimated effects and ANOVA table

The above output (in Figure A.1) indicates that there are significant main effects (“Casting temperature” and “Mould conditioning”) as well as an interaction between these two factors.

In addition, the probability plot in Figure A.2 indicates the significant factors.

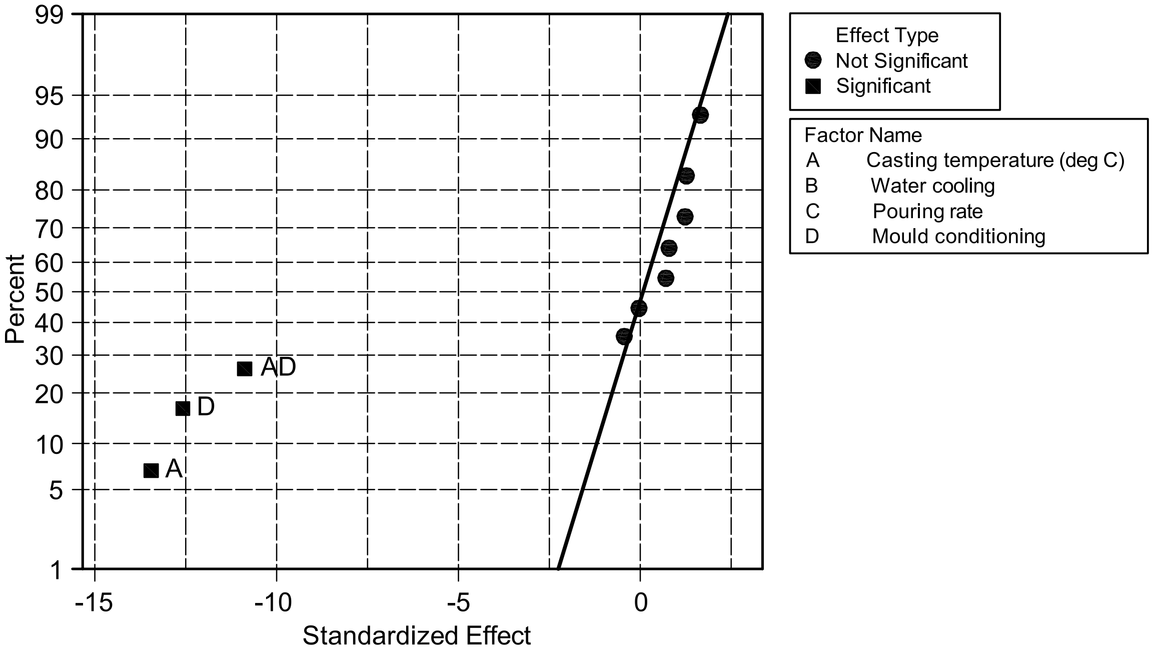


Figure A.2 — Normal probability plot of the standardized effects

The size of the effects and their significance can be seen in the Pareto chart in Figure A.3.

A.7.3 Pareto chart of effects

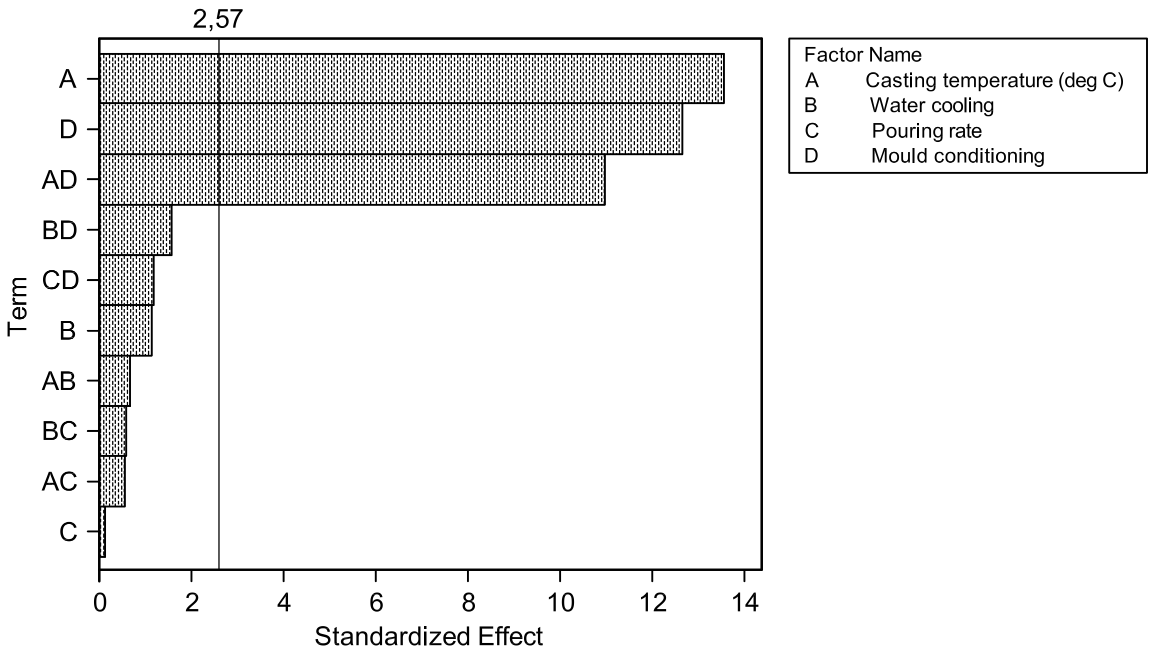


Figure A.3 — Pareto chart of the standardized effects

The vertical line on the Pareto chart indicates the point beyond which the size of the effect would be considered significant (at the 0,05 level). It clearly supports the analysis that factors A, D and their interaction are significant in influencing the occurrence of rosettes.

The plot shows the "standardized" effect. This means the actual effect has been standardized in the statistical sense.

A.7.4 Estimation of experimental error and standard error of effects

It was mentioned above that only the two-factor interactions were considered. The degrees of freedom of those discounted interactions (ABC, ABD, ACD, BCD and ABCD) are all assigned to the estimation of the residual error. Therefore, the degrees of freedom in the ANOVA table in Figure A.1 show 5 for the residual error. If all of the terms had been used, it would have resulted in zero degrees of freedom to estimate the residual error and would have prevented any statistical analysis to have taken place in this experiment.

The standard error of the effects have been calculated using the usual formula, i.e. $\sigma_e/\sqrt{n}$. The estimate of σ_e can be found by taking the square root of the adjusted mean square (Adj MS) value for the residual error, i.e. $\sqrt{41{,}24}$ or 6,422. Since each level of each factor occurs 8 times in the experiment, the standard error can be calculated as 6,422 /$\sqrt{8}$, or 2,271.

A.8 Presentation of results — Effects plots

The main effects plot is shown in Figure A.4 clearly indicating that the "Casting temperature" needs to be set at the 320 °C level and that moulds need to be conditioned to reduce the level of rosettes.

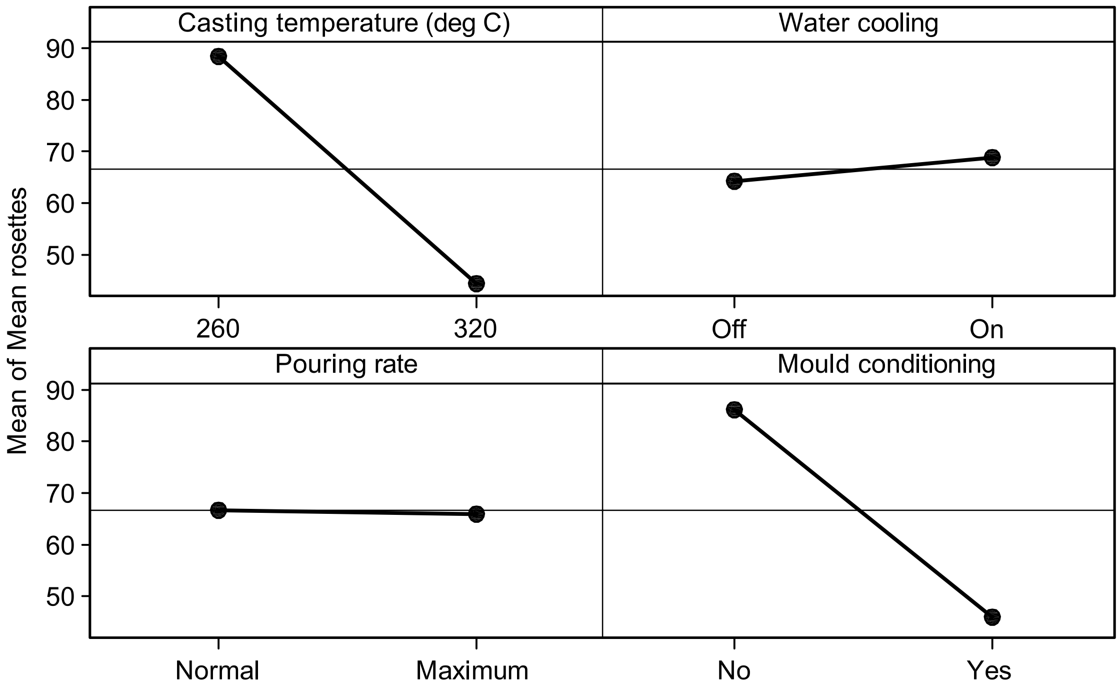


Figure A.4 — Main effects plot

Figure A.5 is an interaction plot between the factors. The interaction of note is that between the "Casting temperature" and the "Mould conditioning", indicating that it is important to have both of these factors set to their "higher" settings to really reduce the level of rosettes.

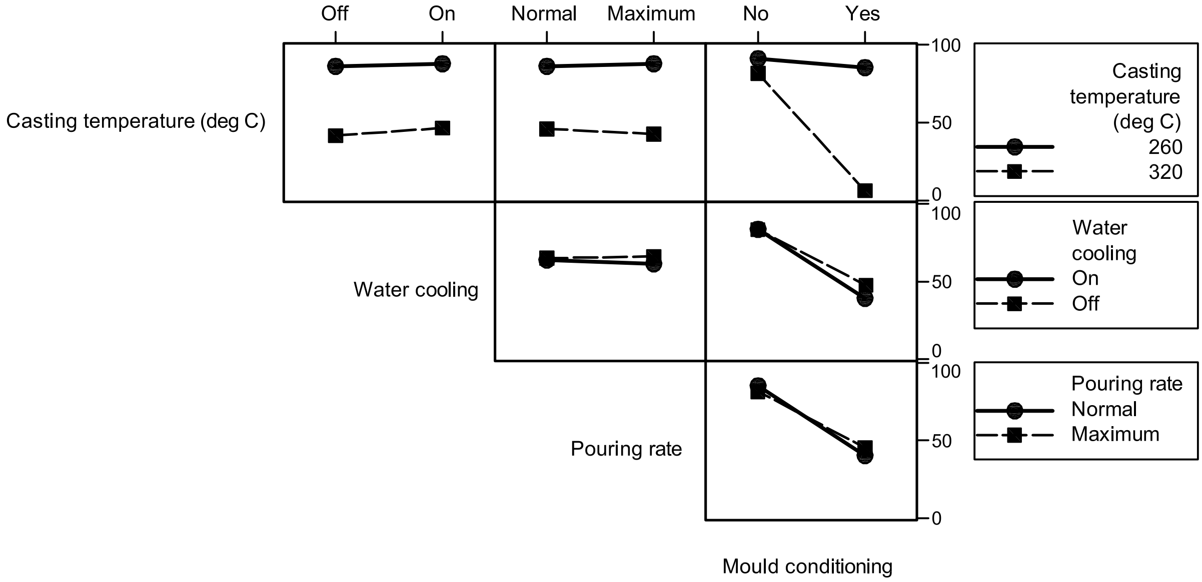


Figure A.5 — Interaction plot

A.9 Presentation of results — Optimization recommendations

The above analysis indicates that it is important to have the casting temperature set at 320 °C and the moulds to be conditioned for the response to be minimized. The other factors and their interactions were unimportant. Consequently, it was decided to run the process with the settings shown in Table A.5.

Table A.5 — Process settings

Factor	Setting
A: Casting temperature (°C)	320
B: Water cooling	Off
C: Pouring rate	Normal
D: Mould conditioning	Yes

A.10 Confirmation results

A run was conducted with the chosen settings described in Table A.5 above. The results were that no rosettes were found.

There was only one draw-back to running the process in this way, which was that the bars showed slight symptoms of oxidation owing to the high casting temperature. This gave a slight "yellowing" of the bars instead of the usual bright shiny appearance. A subsequent experiment was then called for to determine a temperature below 320 °C at which the process could be run and still hold a very low occurrence of rosettes.

Annex B
(informative)

Direct mail marketing campaign[3)]

B.1 General

Direct mail material for a direct marketing campaign can be sent in a variety of packaging styles and with various textual and graphical cues to encourage a high response rate. Historically, the company had experimented with a small subset of their campaigns in which one aspect of the package would be varied. The company was receptive to an approach that was more sophisticated than their previous “one factor at a time” style.

B.2 Overall objective for the experiment

The objective of the experiment was to maximize the magazine subscription response rate based on a direct mailing campaign.

B.3 Description of the process

A typical mailing involves solicitations to 400 000 potential subscribers. The actual package mailed can be varied regarding outer envelope design and other features that had previously enhanced response rate. For this study, the parent company opted to devote 40 000 packages to the experiment. The mailing itself is subcontracted to a third party that specializes in assembling these packages.

B.4 Response variable

B.4.1 Choice of variable

The response is the number of people who subscribed and paid (either by credit card or cheque).

B.4.2 Measurement of the response variable

Exactly 2 500 mailings went out for each combination of package contents. Subscribers were tracked to determine which package of materials they had been sent.

B.4.3 Relationship of the response variable to the objective of the experiment

The publisher expected a response rate of about 2 %. Furthermore, the publisher wanted to be fairly certain to be able to recognize a 0,5 % increase and at least a decent chance of detecting a 0,25 % increase. The response rate is the critical factor in addressing the objective of the experiment as the package variations have roughly the same cost in preparation. Of secondary interest is the proportion of positive responders who pay immediately as well, which avoids a later billing.

---

3) This example has been adapted from the article “Using a Fractional Factorial Design to Increase Direct Mail Response,” by J. Ledolter and A. J. Swersey (2006), *Quality Engineering*, **18**, pp. 469-475. Used with permission from the publisher, the American Society of Quality, Milwaukee, Wisconsin.

B.5 Factors affecting the response

B.5.1 Description of each factor (continuous/discrete) to be varied

The factors chosen to be varied within the experiment were determined using the publisher’s assessment of possible improvements and an appreciation of cost-effective adjustments. These were:

A: Presence/absence of an “Act Now to Respond/Pay today” insert;

B: Additional option of paying by credit card (rather than only by personal cheque);

C: Offer wording strength;

D: Presence/absence of phrase on outer envelope with mild profanity.

Each of these factors are categorical with two levels.

B.5.2 Selection of levels (related to size of effect to be determined)

The factors used in the experiment and their associated levels are given in Table B.1.

Table B.1 — Factors and their associated levels

Factor	Level 1	Level 2
A: Act now	No insert	Insert
B: Payment	Personal cheque	Cheque or credit card
C: Offer phrase	No mention	Mention “hot off press”
D: Profanity	No profanity	Mild profanity

B.5.3 Other factors noted but not incorporated due to issues with controllability or relevance

Other factors that might have been considered, but were not included in the experiment, are shown in Table B.2.

Table B.2 — Other factors considered

Factor	Reason for exclusion
Enclosure of bumper sticker	An attractive but costly bumper sticker mentioning the magazine could induce some to respond favourably to the campaign. The size of the bumper sticker impacts the mailing package and postal costs. No significant improvement with bumper sticker was expected.
Guarantee	Money-back guarantee for part or all of the subscription was considered. This factor suggests a defeatist attitude and was rejected for consideration. Moreover, the implementation of a money-back policy incurs costs that directly impact profits.
Testimonials	Celebrities could be hired to extol the virtues of the magazine. In light of the subject matter content, it was not determined who would be most appropriate and if in fact this celebrity could discourage some respondents.
Personalized salutations	For considerably more expense in preparation of the package, the customer’s name could be used. To recoup this cost, an extraordinary response rate would be required and this was deemed risky.

B.6 Full factorial design

B.6.1 Design matrix (standard order, run order)

The design selected was a “full” factorial that allowed for all interactions to be examined.

Table B.3 shows the design.

Table B.3 — Design matrix

A: Act now	B: Payment options	C: Offer phrase	D: Profanity	Standard order
No insert	No credit card	None	None	1
Insert	No credit card	None	None	2
No insert	Credit card	None	None	3
Insert	Credit card	None	None	4
No insert	No credit card	Hot off press	None	5
Insert	No credit card	Hot off press	None	6
No insert	Credit card	Hot off press	None	7
Insert	Credit card	Hot off press	None	8
No insert	No credit card	None	Mild profanity	9
Insert	No credit card	None	Mild profanity	10
No insert	Credit card	None	Mild profanity	11
Insert	Credit card	None	Mild profanity	12
No insert	No credit card	Hot off press	Mild profanity	13
Insert	No credit card	Hot off press	Mild profanity	14
No insert	Credit card	Hot off press	Mild profanity	15
Insert	Credit card	Hot off press	Mild profanity	16

Randomization was used in selecting the 2 500 addresses out of the 40 000 addresses constituting the sampling frame. Considerable effort was expended to ensure that the third party provider correctly assembled the mailed packages.

B.6.2 Centre points

Centre points were not selected for this experiment because the variables are categorical. Moreover, the response rate can be assessed as a binomial random variable with a nominal response rate of 2 % so that an independent estimate of the uncertainty can be calculated.

B.6.3 Replication, repetition

As mentioned in B.4.2, the mailings were sent to 2 500 potential subscribers. This experiment had neither replication nor repetition.

B.7 Analysis of results

B.7.1 Data acquired through the experiment and analysis considerations

The results from the experiment are given in the Table B.4. They were keyed into a software program called JMP[4].

Table B.4 — Experiment results

A: Act now	B: Payment options	C: Offer phrase	D: Profanity	Standard order	Response rate %
No insert	No credit card	None	None	1	2,08
Insert	No credit card	None	None	2	2,52
No insert	Credit card	None	None	3	2,36
Insert	Credit card	None	None	4	2,12
No insert	No credit card	Hot off press	None	5	2,12
Insert	No credit card	Hot off press	None	6	2,52
No insert	Credit card	Hot off press	None	7	1,96
Insert	Credit card	Hot off press	None	8	2,64
No insert	No credit card	None	Mild profanity	9	2,40
Insert	No credit card	None	Mild profanity	10	2,76
No insert	Credit card	None	Mild profanity	11	3,24
Insert	Credit card	None	Mild profanity	12	3,04
No insert	No credit card	Hot off press	Mild profanity	13	2,36
Insert	No credit card	Hot off press	Mild profanity	14	3,12
No insert	Credit card	Hot off press	Mild profanity	15	2,64
Insert	Credit card	Hot off press	Mild profanity	16	3,20

All effects can be estimated since we have an independent estimate of error based on the binomial response variable. Each effect is estimated by the average response rate for the effect at the high level minus the response rate for the effect at its low level. Since the overall response rate is 2,57 %, the standard error of each estimated effect is $\sqrt{[4(2{,}57)(100 - 2{,}57) / 40\,000]} = 0{,}158$ %. The Z ratio is effect divided by standard error for each term. The expression "Prob>|z|" refers to the probability that a standard normal variable is greater than z or less than $-z$.

4) JMP is the trade name of a product supplied by SAS Inc. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of the product named. Equivalent products may be used if they can be shown to lead to the same results.

Table B.5 — Estimation of effects

Term	Coefficient	Effect	Standard error	Z ratio	Prob>|z|
A	0,172 5	0,345	0,158	2,18	0,029
B	0,082 5	0,165	0,158	1,04	0,297
A*B	−0,072 5	−0,145	0,158	−0,92	0,359
C	0,002 5	0,005	0,158	0,03	0,975
A*C	0,127 5	0,255	0,158	1,61	0,107
B*C	−0,042 5	−0,085	0,158	−0,54	0,591
A*B*C	0,082 5	0,165	0,158	1,04	0,297
D	0,277 5	0,555	0,158	3,51	0,000
A*D	0,012 5	0,025	0,158	0,16	0,874
B*D	0,102 5	0,205	0,158	1,30	0,195
A*B*D	−0,022 5	−0,045	0,158	−0,28	0,776
C*D	−0,017 5	−0,035	0,158	−0,22	0,825
A*C*D	0,017 5	0,035	0,158	0,22	0,825
B*C*D	−0,052 5	−0,105	0,158	−0,66	0,507
A*B*C*D	−0,037 5	−0,075	0,158	−0,47	0,636

The above summary (Table B.5) indicates the significance of terms A and D and a marginal AC interaction. The coefficients were generated in JMP and the results stored in a data sheet. The Z ratio and Prob>|z| were calculated using the formula feature.

In addition, the probability plot in Figure B.1 indicates potentially significant factors. However, the significance is based on a separate calculation that is not pertinent to the binomial response.

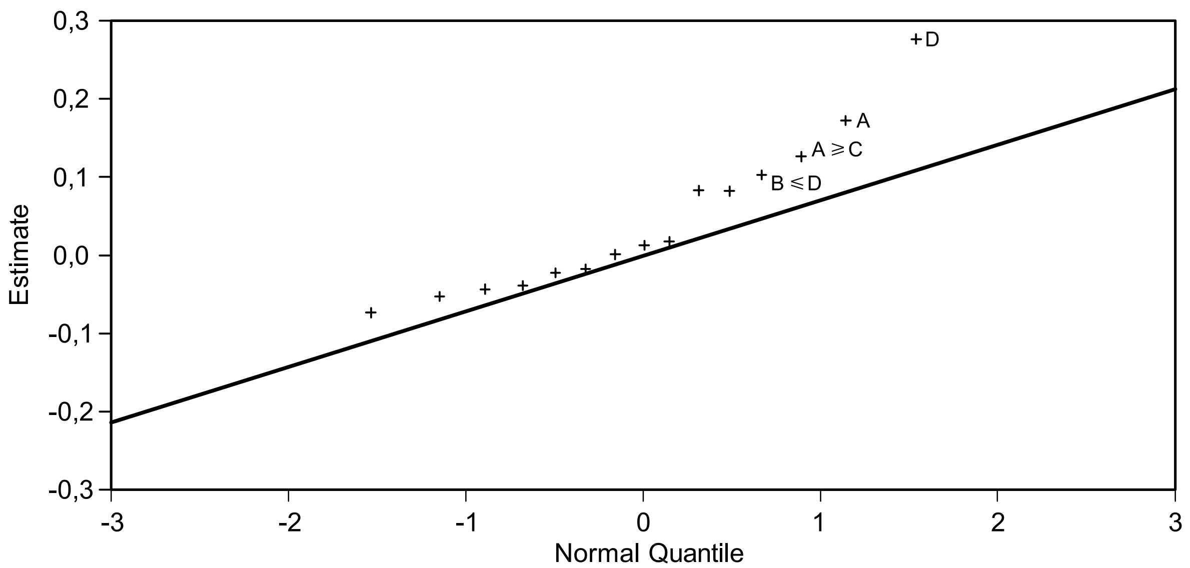


Figure B.1 — Probability plot

B.7.2 Pareto chart of effects

The relative size of the effects can be seen in the Pareto chart (see Figure B.2).

Term	Estimate
D	0,2775000
A	0,1725000
A ± C	0,1275000
B ± D	0,1025000
A ± B ± C	0,0825000
B	0,0825000
A ± B	-0,0725000
B ± C ± D	-0,0525000
B ± C	-0,0425000
A±B±C±D	-0,0375000
A ± B ± D	0,0225000
A ± C ± D	0,0175000
C ± D	-0,0175000
A ± D	0,0125000
C	0,0025000

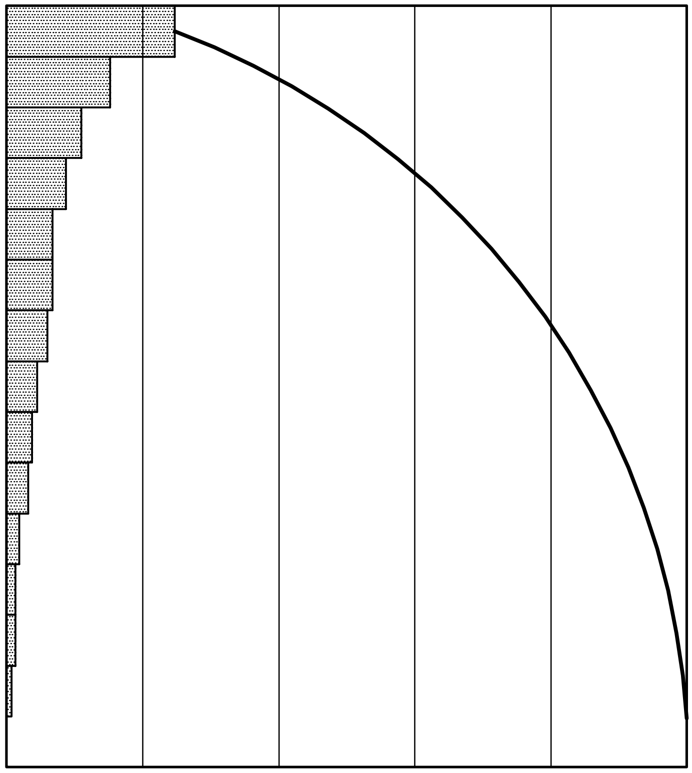


Figure B.2 — Pareto chart of estimates

B.7.3 Estimation of experimental error and standard error of effects

Each effect is estimated by the average response rate for the effect at the high level minus the response rate for the effect at its low level. Since the overall response rate is 2,57 %, the standard error of each estimated effect is $\sqrt{[4(2,57)(100 - 2,57) / 40\,000]} = 0,158$ %.

B.8 Presentation of results — Effects plots

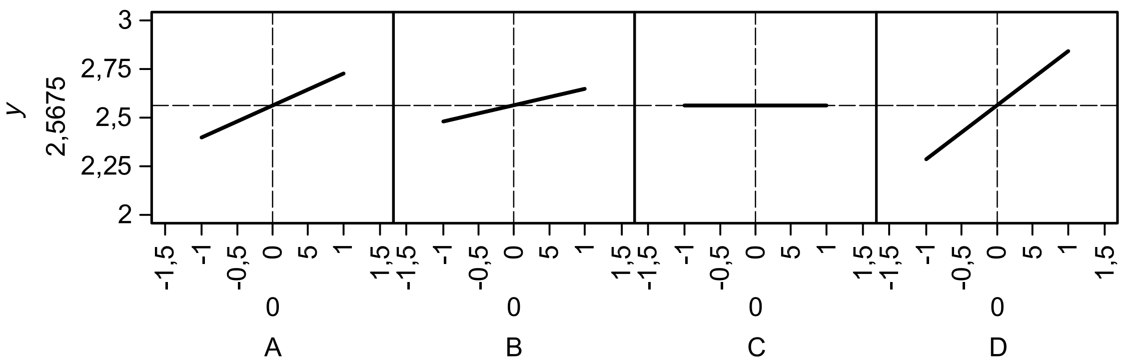


Figure B.3 — Effects plots

The relative slopes of the lines in Figure B.3 are consistent with the significance of the factors (D, A, B and C).

Figure B.4 is an interaction plot between the factors. The interaction of note is that between factors A (Act now) and C (Offer phrase).

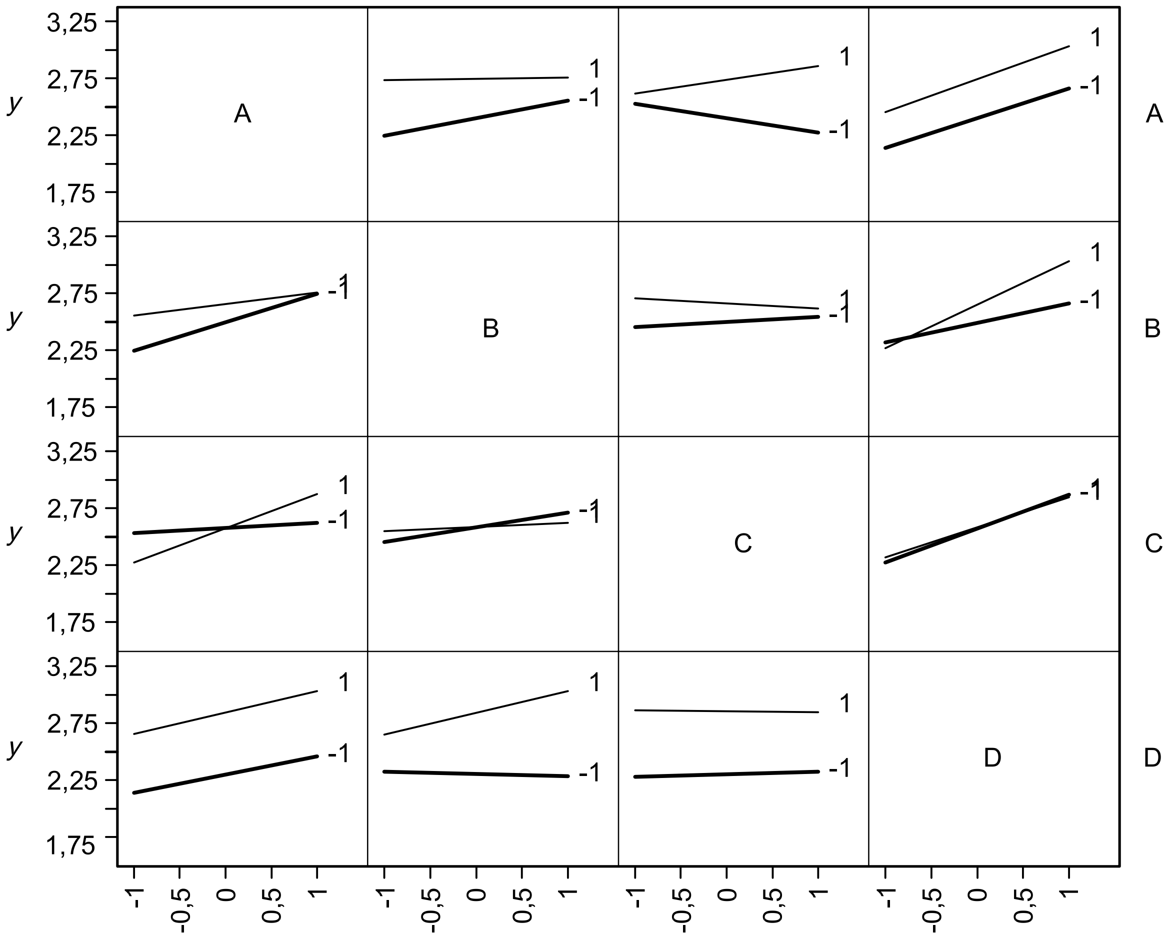


Figure B.4 — Interaction profiles

B.9 Presentation of results — Optimization recommendations

The preferred settings for future mailings are given in Table B.6.

Table B.6 — Preferred settings

Factor	Setting
A: Act now	Insert
B: Payment options	Include credit card option
C: Offer phrase	Include “hot off press”
D: Profanity	Include “off-colour” word

Note that although factor B was not deemed statistically significant (the subscription rate was increased by an estimated 0,165 %), it was noticed that the number of people who returned the reply card without payment declined from 10 % to 5 %.

B.10 Confirmation results

Subsequent mailings confirmed the benefits of offering the credit card payment option in addition to the personal cheque option.

Annex C
(informative)

Button tactility experiment[5)]

C.1 General

Button tactility (perceptibility by touch) is a critical requirement for wireless products such as two-way radios and cellular phones. Poor button tactility will lead to customer complaints and dissatisfaction. The main purpose of this experiment is to provide good button tactility for two-way radios by optimizing specification settings of the existing component using statistical tools.

C.2 Overall objective for the experiment

The objective of the experiment is to achieve good button tactility on the emergency button. The study is focused on the emergency button due to more frequent customer complaints associated with this button on previous products.

C.3 Description of the process

When a keypad button is depressed, the keypad button plunger pushes on a metal dome. Once this force exceeds the actuation force of the dome, the dome collapses and the force subsequently reduces to a new force called the return force. The ratio of these two forces, called the snap ratio, is a measure of tactility. Electrical connectivity is activated in the interval between the actuation and the return forces.

A customer experiences tactility when he/she feels a “click” as a button is depressed. Tactility is important as it gives feedback to the user that the button has been activated. When a button has poor tactility, the user does not receive physical feedback that the key has been actuated and has to check by some other means, which is not desirable in a mission critical situation.

C.4 Response variable

C.4.1 Choice of variable

The button tactility is defined as the snap ratio (%) ➔ $[(F_1 - F_2) / F_1] \times 100~\%$

where

F_1 is the actuation force, expressed in gram force (gf) [6)]; and

F_2 is the return force, expressed in gram force (gf).

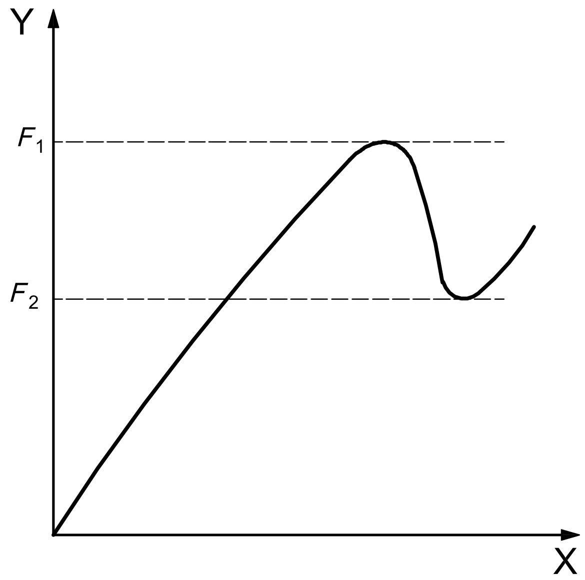
---

5) This example has been kindly donated by Motorola (Penang Engineering Group).

6) 1 gf = $9{,}806\ 65 \times 10^{-3}$ N.

C.4.2 Measurement of the response variable

To measure F_1 and F_2, a mechanical plunger is used to depress the button and the force curve graph is plotted. The first peak on the graph (Figure C.1) is the actuation force and the lowest point after the first peak is the return force.



Key

X travel distance, in millimetres (mm)

Y force, in gram force (gf)

Figure C.1 — Force curve

C.5 Factors affecting the response

C.5.1 Description of each factor to be varied

The factors chosen to be varied within the experiment were determined using the knowledge of the design engineering staff. These were:

A: Duro hardness button;

B: Air vent width;

C: Actuation force of dome; and

D: Plunger length.

The experiment was to be run using these four factors each at two levels:

- factor A: measure of the hardness of the button, higher is harder;
- factor B: width of the vent required under the metal dome which allows the metal dome to collapse and return;
- factor C: force required to collapse the metal dome;
- factor D: length of the plunger located beneath each key of the keypad.

C.5.2 Selection of levels (related to size of effect to be determined)

The continuous factors used in the experiment and their associated levels are given in Table C.1

Table C.1 — Continuous factors and their associated levels

Factor	Level 1	Level 2
A: Duro hardness button	40	80
B: Air vent width	0,6	1,8
C: Actuation force of dome	120	200
D: Plunger length	0,7	1

C.5.3 Other factors noted but not incorporated due to issues with controllability or relevance

Other factors that might have been considered, but were not included in the experiment, are shown in Table C.2.

Table C.2 — Other factors

Factor	Reason for exclusion
Surface area of the button	The area for the button on this product was finalized upfront by the industrial design team based on user feedback

.

C.6 Full factorial design

C.6.1 Design matrix (standard order, run order)

The design selected was a full factorial design that allowed for all interactions to be examined. Three centre points were added to detect curvature and estimate the pure error.

Table C.3 shows the design.

The experiment was randomized and the column headed “Run order” was the order in which the runs were performed during the experiment. The column “Standard order” specifies where the particular factor combination would be found in the “design matrix” for the 4-factor 2-level design. The three centre points were deliberately spread throughout the runs in order to capture variation in the course of the experiment.

Table C.3 — Design matrix

Centre point	Blocks	A: Duro hardness button	B: Air vent width	C: Actuation force of dome	D: Plunger length	Standard order	Run order
1	1	40	0,6	120	0,7	1	1
1	1	80	1,8	120	1	12	2
0	1	60	1,2	160	0,85	19	3
1	1	80	1,8	200	0,7	8	4
0	1	60	1,2	160	0,85	17	5
1	1	40	1,8	120	1	11	6
1	1	40	0,6	200	0,7	5	7
1	1	40	1,8	120	0,7	3	8
1	1	40	1,8	200	1	15	9
1	1	80	1,8	200	1	16	10
1	1	40	0,6	120	1	9	11
1	1	80	0,6	200	0,7	6	12
1	1	80	1,8	120	0,7	4	13
1	1	40	1,8	200	0,7	7	14
1	1	80	0,6	120	1	10	15
1	1	80	0,6	120	0,7	2	16
1	1	40	0,6	200	1	13	17
1	1	80	0,6	200	1	14	18
0	1	60	1,2	160	0,85	18	19

C.6.2 Centre points

Three centre points were added to the original design.

C.6.3 Replication, repetition

Since three centre points were included in the design, it was not necessary to replicate other design points. The centre points provided the estimate of pure error.

C.7 Analysis of results

C.7.1 Data acquired through the experiment and analysis considerations

The results from the experiment are given in Table C.4. They were keyed into the software program MINITAB™ [7].

7) MINITAB™ is the trade name of a product supplied by Minitab Inc. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of the product named. Equivalent products may be used if they can be shown to lead to the same results.

Table C.4 — Experiment results

Centre point	Blocks	A: Duro hardness button	B: Air vent width	C: Actuation force of dome	D: Plunger length	Standard order	Run order	Respond
1	1	40	0,6	120	0,7	1	1	24,76
1	1	80	1,8	120	1	12	2	27,17
0	1	60	1,2	160	0,85	19	3	33,24
1	1	80	1,8	200	0,7	8	4	33,55
0	1	60	1,2	160	0,85	17	5	33,27
1	1	40	1,8	120	1	11	6	18,01
1	1	40	0,6	200	0,7	5	7	19,55
1	1	40	1,8	120	0,7	3	8	35,07
1	1	40	1,8	200	1	15	9	17,16
1	1	80	1,8	200	1	16	10	32,78
1	1	40	0,6	120	1	9	11	22,38
1	1	80	0,6	200	0,7	6	12	39,77
1	1	80	1,8	120	0,7	4	13	27,95
1	1	40	1,8	200	0,7	7	14	20,37
1	1	80	0,6	120	1	10	15	27,23
1	1	80	0,6	120	0,7	2	16	21,56
1	1	40	0,6	200	1	13	17	19,22
1	1	80	0,6	200	1	14	18	36,79
0	1	60	1,2	160	0,85	18	19	32,58

An initial assumption was made that only the first-order interactions, e.g. AB, would be analysed, together with the main effects. Therefore, the higher-order interactions were considered to be unimportant and insignificant. This assumption was later confirmed by fitting a saturated model and recognizing that the three-way and four-way interactions were in fact not statistically significant, using the centre points to provide an estimate of pure error.

C.7.2 Estimation of effects

```
Estimated Effects and Coefficients for Respond (coded units)

Term                                 Effect    Coef  SE Coef      T      P
Constant                                     27.495    1.119  24.56  0.000
Duro hardness button                  8.785   4.392    1.220   3.60  0.007
Air vent width                        0.100   0.050    1.220   0.04  0.968
Actuation force of dome               1.883   0.941    1.220   0.77  0.462
Plunger length                       -2.730  -1.365    1.220  -1.12  0.296
Duro hardness button*Air vent width  -1.075  -0.537    1.220  -0.44  0.671
Duro hardness button*                 7.863   3.931    1.220   3.22  0.012
  Actuation force of dome
Duro hardness button*Plunger length   3.015   1.508    1.220   1.24  0.252
Air vent width*                      -2.968  -1.484    1.220  -1.22  0.258
  Actuation force of dome
Air vent width*Plunger length        -2.725  -1.362    1.220  -1.12  0.296
Actuation force of dome*              0.907   0.454    1.220   0.37  0.720
  Plunger length


S = 4.87905   R-Sq = 78.83%   R-Sq(adj) = 52.37%


Analysis of Variance for Respond (coded units)

Source              DF   Seq SS   Adj SS   Adj MS       F      P
Main Effects         4  352.732  352.732   88.183    3.70  0.054
2-Way Interactions   6  356.480  356.480   59.413    2.50  0.115
Residual Error       8  190.441  190.441   23.805
  Curvature          1  109.131  109.131  109.131    9.40  0.018
  Lack of Fit        5   81.006   81.006   16.201  106.52  0.009
  Pure Error         2    0.304    0.304    0.152
Total               18  899.653
```

Figure C.2 — Factorial fit: Respond versus Duro, air vent, force, plunger length

The above output (see Figure C.2) indicates that there are significant differences for some of the main factors (Duro hardness button) as well as an interaction between the Duro hardness button (Duro) and the actuation force of dome (force). The output also indicates curvature and lack of fit is significant (see C.9).

In addition, the probability plot (see Figure C.3) indicates the significant factors.

A reduced model including Duro hardness button and actuation force of the dome and their interaction could be refitted to confirm the significance of these factors.

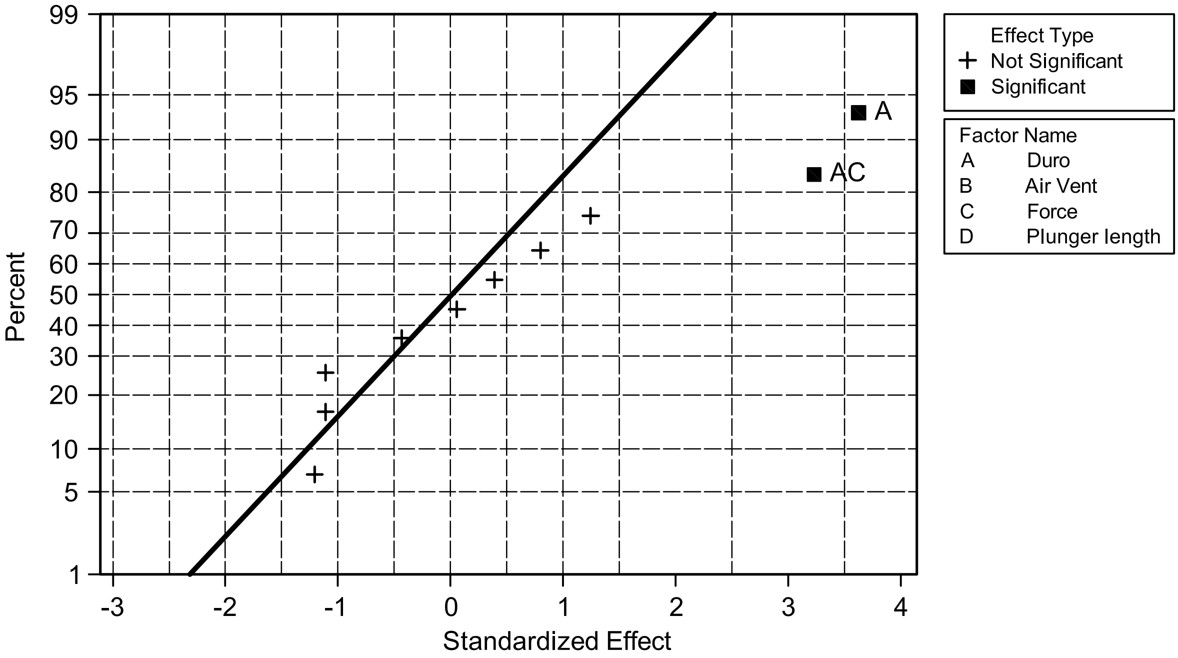


Figure C.3 — Normal probability plot of the standardized effects

C.7.3 Pareto chart of effects

The size of the effects and their significance can be seen in the Pareto chart in Figure C.4.

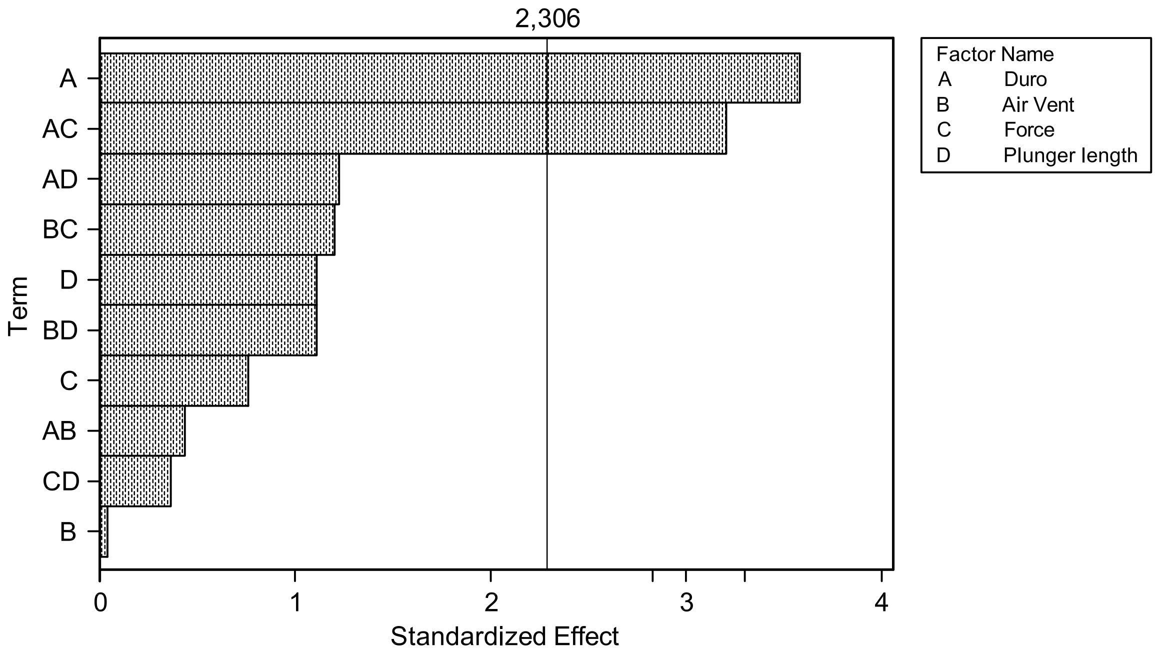


Figure C.4 — Pareto chart of the standardized effects

The vertical line on the Pareto chart indicates the point beyond which the size of the effect would be considered significant (at the 0,05 level). The chart clearly supports the analysis that factor A and the interactions between A and C are significant in influencing the button tactility.

The plot shows the "standardized" effect. This means the actual effect has been standardized in the statistical sense.

C.7.4 Estimation of experimental error and standard error of effects

It was mentioned above that only the two-factor interactions were considered. Therefore, the degrees of freedom of those discounted interactions (ABC, ABD, ACD, BCD and ABCD) are all assigned to the estimation of the residual error. Two degrees of freedom for centre points are also included in the residual error. Therefore, the number of degrees of freedom in the ANOVA table is eight for the residual error.

The standard error of the effects has been calculated using the usual formula, i.e. $\sigma_e/\sqrt{n}$. The estimate of σ_e can be found by taking the square root of the adjusted mean square (Adj MS) value for the residual error, i.e. $\sqrt{23,805}$ or 4,879. Since each level of each factor occurs eight times in the experiment, the standard error can be calculated as 4,879/$\sqrt{8}$, or 1,725.

C.8 Presentation of results — Effects plots

See Figures C.5 and C.6.

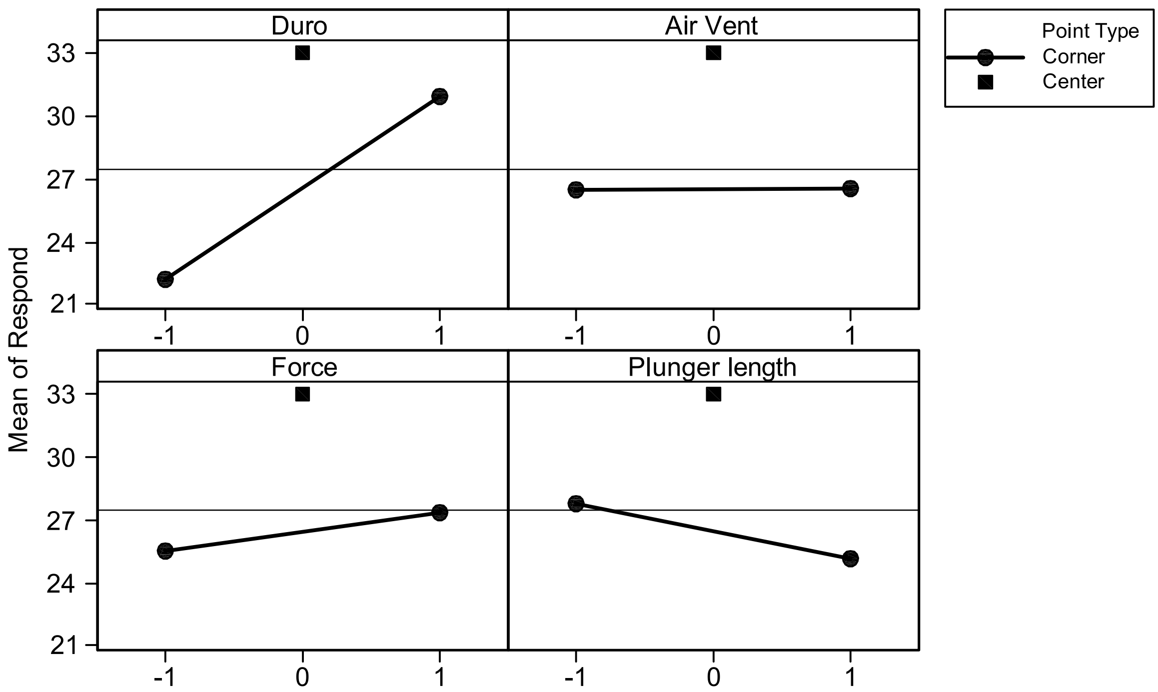


Figure C.5 — Main effects plot

C.9 Presentation of results — Optimization recommendations

The above analysis indicates a presence of curvature, i.e. quadratic terms could be included in the model. It is recommended to follow up with a Response Surface Experiment using two factors: 1) the Duro hardness button (Duro); and 2) the actuation force of dome (force).

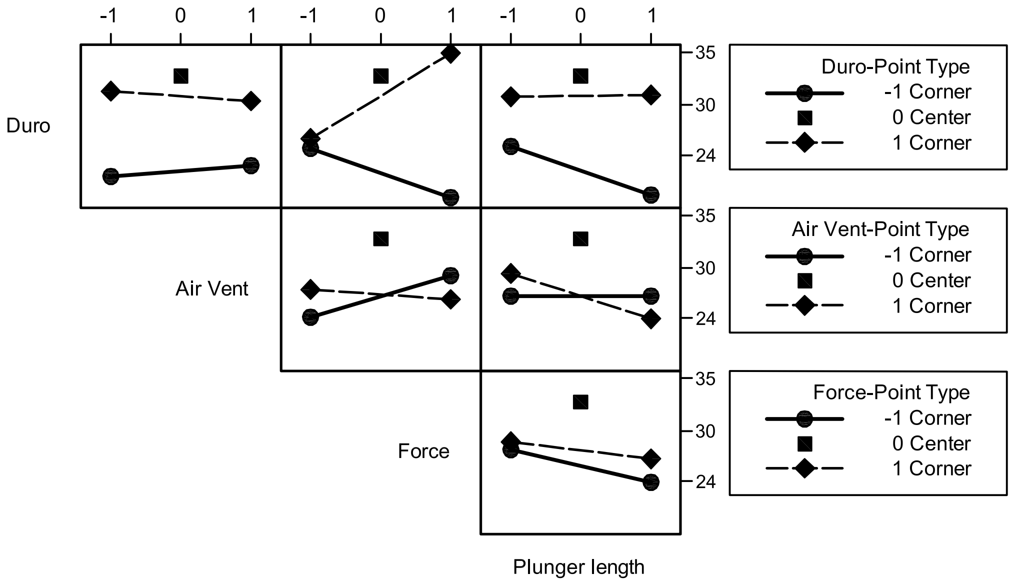


Figure C.6 — Interaction plot

The residuals versus Duro plot (see Figure C.7) shows that the residual values at the centre points are on one side of the zero line. This indicates the presence of curvature.

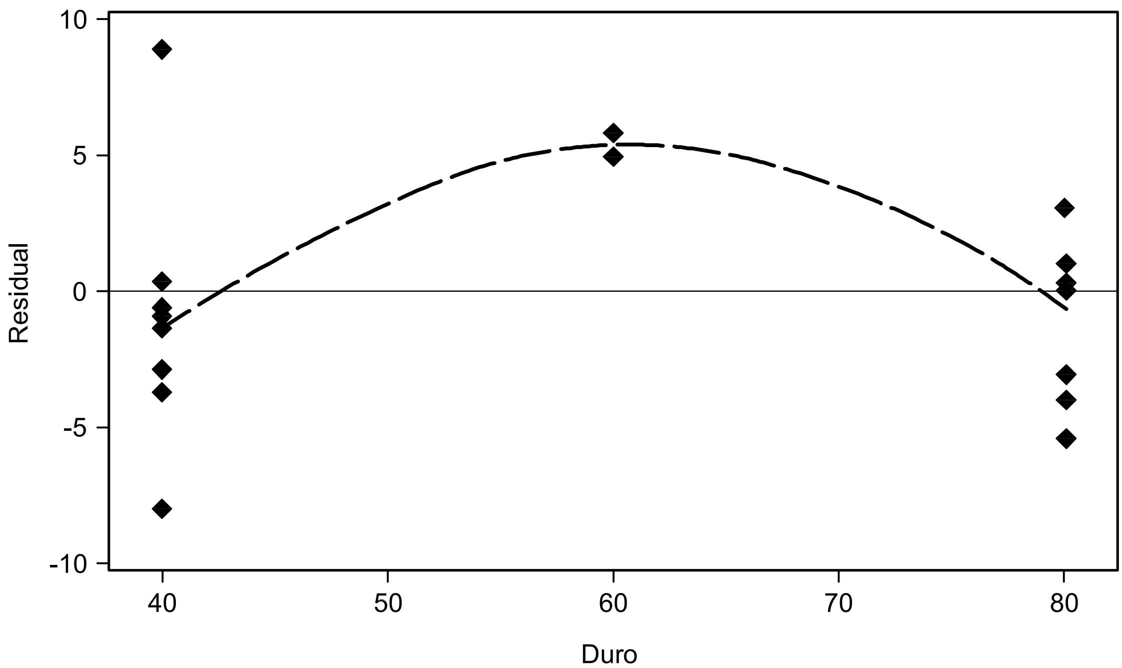


Figure C.7 — Residuals versus Duro plot

C.10 Confirmation results

A confirmation run should be performed after the Response Surface Experiment is conducted.

Annex D
(informative)

Optimizing a customer PVC formulation[8)]

D.1 General

Processing Aids (PA) facilitate the manufacturing of PVC parts by changing the physical properties of the plastic. Formulation of PVC varies depending on application, equipment and component types and loading used. The PA manufacturing company is now selling a new, more efficient product that requires some modification to the PVC formulation, based on customer specifics. One of the PA manufacturing company customers, a PVC manufacturer, has just switched to the new product with a reduced PA loading. He reports that the new product is responsible for degraded processing conditions.

D.2 Overall objective for the experiment

At the manufacturing company laboratory, it was decided to design and run an experiment to reproduce the customer experience. As a starting point, the same formulation and processing conditions were used with the new product that was used for the old one. Key factors were varied within an experimental window around this point, so that the analysis of the experiment would determine the optimum component loadings and process conditions.

D.3 Description of the process

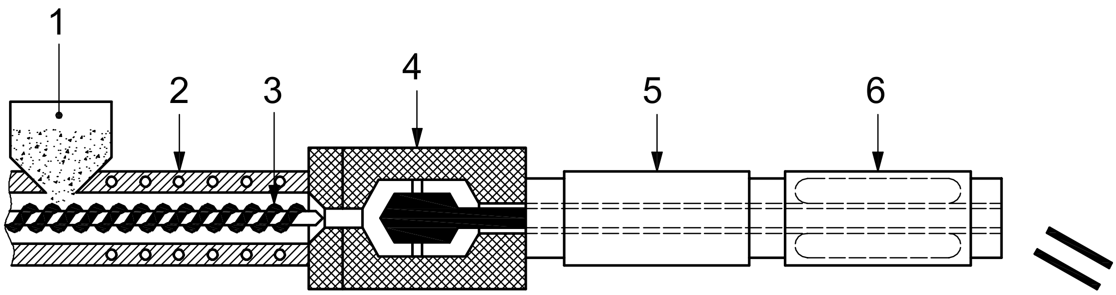
A standard foam formulation, like that used by the customer, has the components (in pellet form) shown in Table D.1.

Table D.1 — Components (in pellet form)

Components	Quantity
PVC suspension resin	100
Stabilizer	2 to 4
Lubricant	2
Processing Aid (PA)	4 to 8
Blowing Agent (BA)	0,5 to 1,5
Filler	5 to 20

Components are thoroughly blended and the resulting material placed in the extruder hopper as shown in Figure D.1. Material is conveyed continuously forward by a rotating screw inside a heated barrel. Softened plastic is forced through a die and directly into cold water where the product solidifies. See Figure D.1.

8) This example has been kindly donated by Rohm and Haas, European Laboratories.



Key

1 hopper
2 barrel
3 screw
4 die
5 cold water
6 caterpillar

Figure D.1 — PVC blend

Temperature can be varied in four different zones of the extruder barrel. It is not known precisely at which temperature profile problems may have been experienced by customers. It was decided to study two different profiles, one flatter than the other, with ranges representing most of the customer's operating conditions.

D.4 Response variable

D.4.1 Choice of variable

Many responses were studied in this experiment. Only one of them, the expansion ratio, amongst the most critical, is presented here.

D.4.2 Measurement of the response variable

The expansion ratio is a measure of part swell. It is calculated as the ratio of the extruded rod diameter to the die diameter. A higher ratio is better, indicating that the melted PVC will mould more easily around difficult shapes.

D.4.3 Relationship of the response variable to the objective of the experiment

The expansion ratio is directly linked to the objective of the experiment — to develop a formulation that will be robust against processing problems.

D.5 Factors affecting the response

D.5.1 Description of each factor (continuous/discrete) to be varied

Four factors were investigated in this experiment:

A: Type of PA used (old or new PA) — The old PA was needed as a baseline against which the new PA would be matched;

B: PA loading;

C: Blowing Agent (BA) loading;

D: Temperature profile in the extruder barrel.

The experiment was run using these four factors each at two levels.

D.5.2 Selection of levels (related to size of effect to be determined)

The factors used in the experiment and their associated levels are shown in Table D.2.

Table D.2 — Factors and their associated levels

Factor	Factor type	Level 1	Level 2
A: BA loading	Continuous	1,6	2,0
B: PA type	Discrete	Old	New
C: PA loading	Continuous	4	6
D: Temperature profile	Discrete	Rising (135/165/190/160 °C)	Flatter (150/170/180/160 °C)

The levels for the Blowing Agent (BA) were fixed at 1,6 phr [9] and 2,0 phr to frame the current level of 1,8 phr used with the old product. These were determined to be the extremes that would be considered during production.

The levels for the Processing Aid (PA) were fixed at 4 phr and 6 phr with 6 phr the recommended level of the old PA. It was known that the new PA was more efficient, thus requiring a smaller quantity to be used.

Each combination of factors A, B and C represented a unique blend. Thus, there were eight blends in the experiment. Two samples of each blend were tested each at a different temperature profile. This constitutes a split-plot design.

It was known prior to this experiment that the three formulation factors (A, B and C) would have an effect on the response. There was no certainty about the effect of factor D.

D.5.3 Other factors noted but not incorporated due to issues with controllability and relevance

Other factors present in the formulation were the loading of PVC, stabilizer, lubricant and filler (customer specific); these were not varied.

9) phr = parts per hundred resin.

D.6 Full factorial design

D.6.1 Design matrix (standard order, run order)

The design selected was a full factorial that allowed for all interactions to be examined. Table D.3 shows the design. This table contains the 16 expected runs from the full factorial design plus eight additional runs for centre points.

Table D.3 — Design matrix

BA loading	PA type	PA loading	Temp. profile	Standard order	Run order	Blend whole plot	Sample sub-plot	Centre points (CP)
1,6	Old	6	Rising	5	1	5	1	
2	Old	4	Rising	2	2	2	1	
1,8	New	5	Rising	18	3	11	1	CP
2	New	4	Rising	4	4	4	1	
1,6	New	6	Rising	7	5	7	1	
1,8	New	5	Rising	22	6	12	1	CP
1,8	Old	5	Rising	21	7	10	1	CP
1,6	Old	4	Rising	1	8	1	1	
1,6	New	4	Rising	3	9	3	1	
1,8	Old	5	Rising	17	10	9	1	CP
2	New	6	Rising	8	11	8	1	
2	Old	6	Rising	6	12	6	1	
1,6	Old	4	Flatter	9	13	1	2	
1,6	Old	6	Flatter	13	14	5	2	
1,8	New	5	Flatter	24	15	12	2	CP
1,6	New	4	Flatter	11	16	3	2	
2	New	6	Flatter	16	17	8	2	
2	Old	6	Flatter	14	18	6	2	
2	Old	4	Flatter	10	19	2	2	
1,8	Old	5	Flatter	23	20	10	2	CP
2	New	4	Flatter	12	21	4	2	
1,8	Old	5	Flatter	19	22	9	2	CP
1,8	New	5	Flatter	20	23	11	2	CP
1,6	New	6	Flatter	15	24	7	2	

Each of the blends in the experiment was divided into two samples, one sample extruded with a flatter temperature profile, one sample with a rising temperature profile. The column headed “blend” contains the blend number. The column headed “sample” contains the sample number for a given blend number.

This experimental arrangement is that of a split-plot design. There are two sizes of experimental units. One is the “blend”, or whole plot, which was used to test the three blend factors (BA loading, PA type, PA loading) and their interactions; the other is a “sample”, or sub-plot, which was used to test the process factor, temperature profile and all its interactions with the blend factors.

Because temperature is difficult to vary, all samples to be tested at flatter temperature profile were processed first, followed by all samples to be tested at “rising” temperature profile.

Thus, the experiment could not be completely randomized. The column headed “Run order” was the order in which the runs were performed during the experiment. The column “Standard order” specifies where the particular factor combination would be found in the “design matrix” for the 4-factor 2-level design.

D.6.2 Centre points

A centre point was selected for each PA type and replicated, thus giving four additional blends. From each of these blends, two samples were taken and tested at different temperature profiles. Centre points have blend numbers ranging from 9 to 12 with the BA loading set at 1,8 and the PA loading set at 5.

D.6.3 Replication, repetition

Five rod samples were taken from each of the 24 samples and the expansion ratio was measured for each, thus providing five repetitions from each individual set-up of the experiment.

In this case, this repetition is important both to overcome the effect of measurement uncertainty and to smooth the effect of blend inhomogeneity. The average expansion ratio from the five repetitions was reported and analysed as the response variable. The standard deviation from the repeated measurements was calculated, logged and monitored for assessing control.

Two replicates were obtained from each of the two centre point combinations and the expansion ratio was measured for each. The average expansion ratio was reported and analysed.

D.7 Analysis of results

D.7.1 Data acquired through the experiment and analysis considerations

The results from the experiment are given in Table D.4. They were keyed into a software program called JMP [10].

D.7.2 Estimation of effects: Full model

First, a full model was fit to assess significant effects. Results are shown in Table D.5.

10) JMP is the trade name of a product supplied by SAS Inc. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of the product named. Equivalent products may be used if they can be shown to lead to the same results.

Table D.4 — Experiment results

BA loading	PA type	PA loading	Temp. profile	Standard order	Run order	Blend	Sample	Expansion ratio
1,6	Old	6	Rising	5	1	5	1	2,778
2	Old	4	Rising	2	2	2	1	2,379
1,8	New	5	Rising	18	3	11	1	2,96
2	New	4	Rising	4	4	4	1	2,67
1,6	New	6	Rising	7	5	7	1	2,995
1,8	New	5	Rising	22	6	12	1	2,89
1,8	Old	5	Rising	21	7	10	1	2,65
1,6	Old	4	Rising	1	8	1	1	2,38
1,6	New	4	Rising	3	9	3	1	2,73
1,8	Old	5	Rising	17	10	9	1	2,63
2	New	6	Rising	8	11	8	1	3,185
2	Old	6	Rising	6	12	6	1	2,68
1,6	Old	4	Flatter	9	13	1	2	2,385
1,6	Old	6	Flatter	13	14	5	2	2,732
1,8	New	5	Flatter	24	15	12	2	2,88
1,6	New	4	Flatter	11	16	3	2	2,515
2	New	6	Flatter	16	17	8	2	3,05
2	Old	6	Flatter	14	18	6	2	2,55
2	Old	4	Flatter	10	19	2	2	2,25
1,8	Old	5	Flatter	23	20	10	2	2,46
2	New	4	Flatter	12	21	4	2	2,567
1,8	Old	5	Flatter	19	22	9	2	2,545
1,8	New	5	Flatter	20	23	11	2	2,87
1,6	New	6	Flatter	15	24	7	2	2,891

Table D.5 — Full model: Fixed effect tests

Source	Nparm [a]	DF [b]	DFDen [c]	F Ratio	Prob > F
BA loading	1	1	4	0,056 9	0,823 2
PA type	1	1	4	96,510 5	0,000 6
PA loading(4,6)	1	1	4	90,084 9	0,000 7
Temperature profile	1	1	8	23,208 8	0,001 3
BA loading*PA type	1	1	4	5,793 7	0,073 8
BA loading*PA loading	1	1	4	0,458 7	0,535 4
BA loading*Temperature profile	1	1	8	0,430 5	0,530 2
PA type*PA loading	1	1	4	0,868 0	0,404 3
PA type*Temperature profile	1	1	8	0,102 8	0,756 7
PA loading*Temperature profile	1	1	8	0,016 7	0,900 3
BA loading*PA type*PA loading	1	1	4	2,537 7	0,186 4

[a] N parm = number of parameters.

[b] DF = degrees of freedom

[c] DFDen = denominator degrees of freedom

Because of the split-plot structure, there are two error terms in this experiment, one for each size of experimental units. For a full model, the error term to test whole plot factors (A, B, C and their interactions) has 4 degrees of freedom: 12 total from which 8 are used to estimate whole plot effects. The error term to test the sub-plot factors (D and all its interactions with whole plot factors) has 8 degrees of freedom: 24 total from which 8 are used to estimate whole plot factors and 8 are used to estimate sub-plot factors. This information is shown in Table D.5, in the column “DFDen”.

In Table D.5, low “Prob > F” values (also known as p-values) indicate significant effects. Three main effects (PA type, PA loading, temperature profile) are significant. The interaction (BA loading*PA type) with somewhat a low p-value (0,073 8) was also investigated further. Other effects, with the exception of BA loading needed because it was involved in an interaction, were removed from the model.

D.7.3 Estimation of effects: Reduced model

The model was reduced to a final form by removing unsignificant terms. The summary of fit is given in Table D.6 and fixed effect tests in Table D.7.

Table D.6 — Reduced model: Summary of fit

RSquare	0,979 001
RSquare Adj	0,973 168
Root Mean Square Error	0,046 022
Mean of Response	2,692 583
Observations (or Sum Wgts)	24

Table D.7 — Reduced model: Fixed effect tests

Source	Nparm	DF	DFDen	F Ratio	Prob > F
BA loading	1	1	7	0,050 6	0,828 4
PA type	1	1	7	85,903 3	<,000 1
PA loading(4,6)	1	1	7	80,184 0	<,000 1
Temperature profile	1	1	11	29,859 2	0,000 2
BA loading*PA type	1	1	7	5,156 9	0,057 4

D.7.4 Estimation of factor effects and their standard errors

Estimates for each factor and their corresponding standard errors are given in Table D.8. As indicated by the magnitude of the estimates, PA loading and PA type have the strongest effects. New PA increases the expansion ratio, thus allowing use of a smaller quantity. A higher temperature profile increases the expansion ratio as well.

Table D.8 — Parameter estimates

Term		Estimate	Standard error	DFDen	t Ratio	Prob>|*t*|
Intercept		2,692 583 3	0,017 011	7	158,28	<,000 1
BA loading		−0,004 688	0,020 834	7	−0,22	0,828 4
PA type [Old]		−0,157 667	0,017 011	7	−9,27	<,000 1
PA loading (4,6)		0,186 562 5	0,020 834	7	8,95	<,000 1
Temperature profile [Rising]		0,051 333 3	0,009 394	11	5,46	0,000 2
BA loading*PA type [Old]		−0,047 313	0,020 834	7	−2,27	0,057 4

D.8 Presentation of results — Effects plots

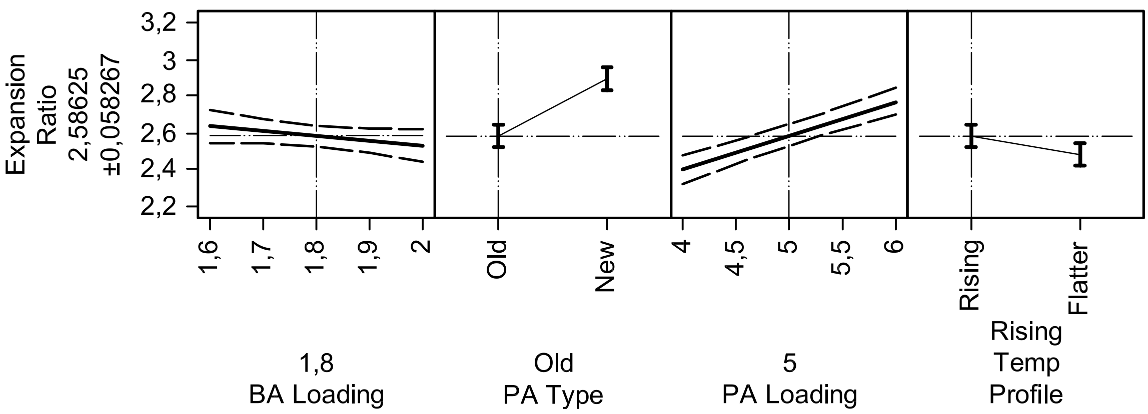


Figure D.2 — Effects plots

The relative slopes of the lines in Figure D.2 are consistent with the significance of the factors (B, C, D).

Figure D.3 is an interaction plot between factor A and factor C (BA loading and PA type). The plot shows a weak interaction, but indicates that the new PA is more efficient with more BA.

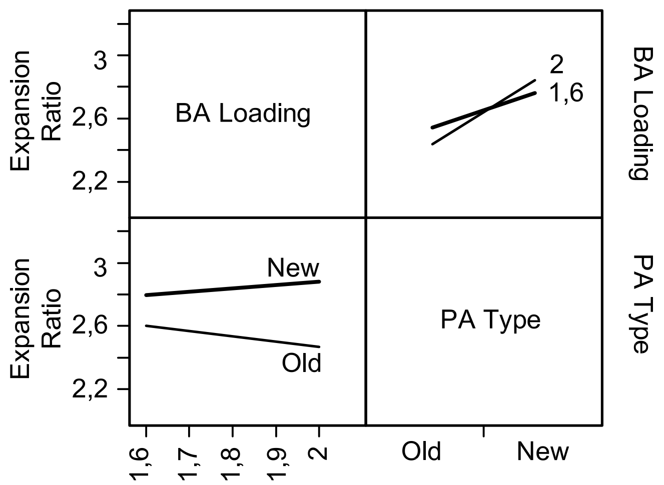


Figure D.3 — Interaction plot — Temperature profile

D.9 Presentation of results — Optimization recommendations

The above analysis confirmed that the new processing aid is more efficient than the old, and that a rising temperature profile is needed to maximize the expansion ratio. It was decided to run the process with the settings shown in Table D.9.

Table D.9 — Settings for a confirmation run

Factor	Level
A: BA loading	2
B: PA type	New
C: PA loading	4,48
D: Temperature profile	Rising

At these settings, as shown in Figure D.4, the fitted model predicts a mean expansion ratio value of 2,83 ± 0,09. This is not the maximum expansion ratio value that can be predicted, but constraints on other responses restricted the ability to perform better.

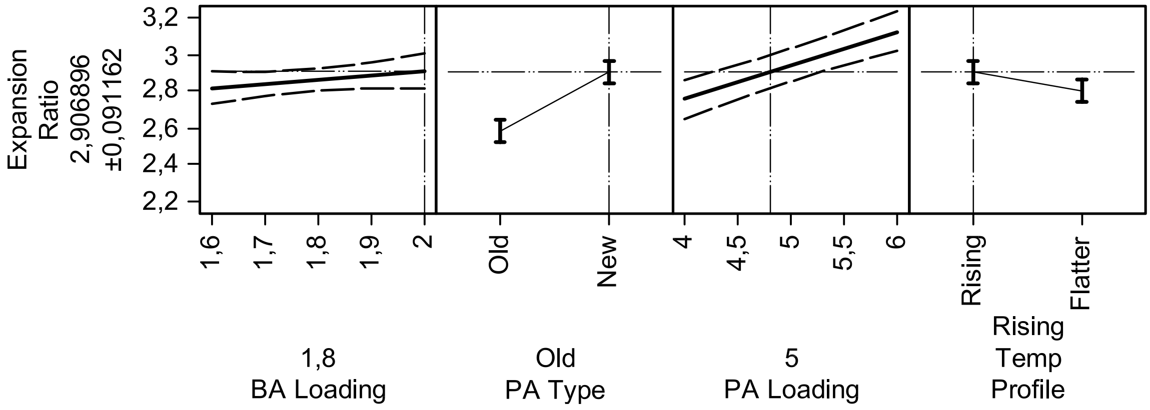


Figure D.4 — Expansion ratio prediction

The contour plots given in Figure D.5 illustrate in a graphical form optimized settings and illustrate to the user changes in the response when the significant factors vary.

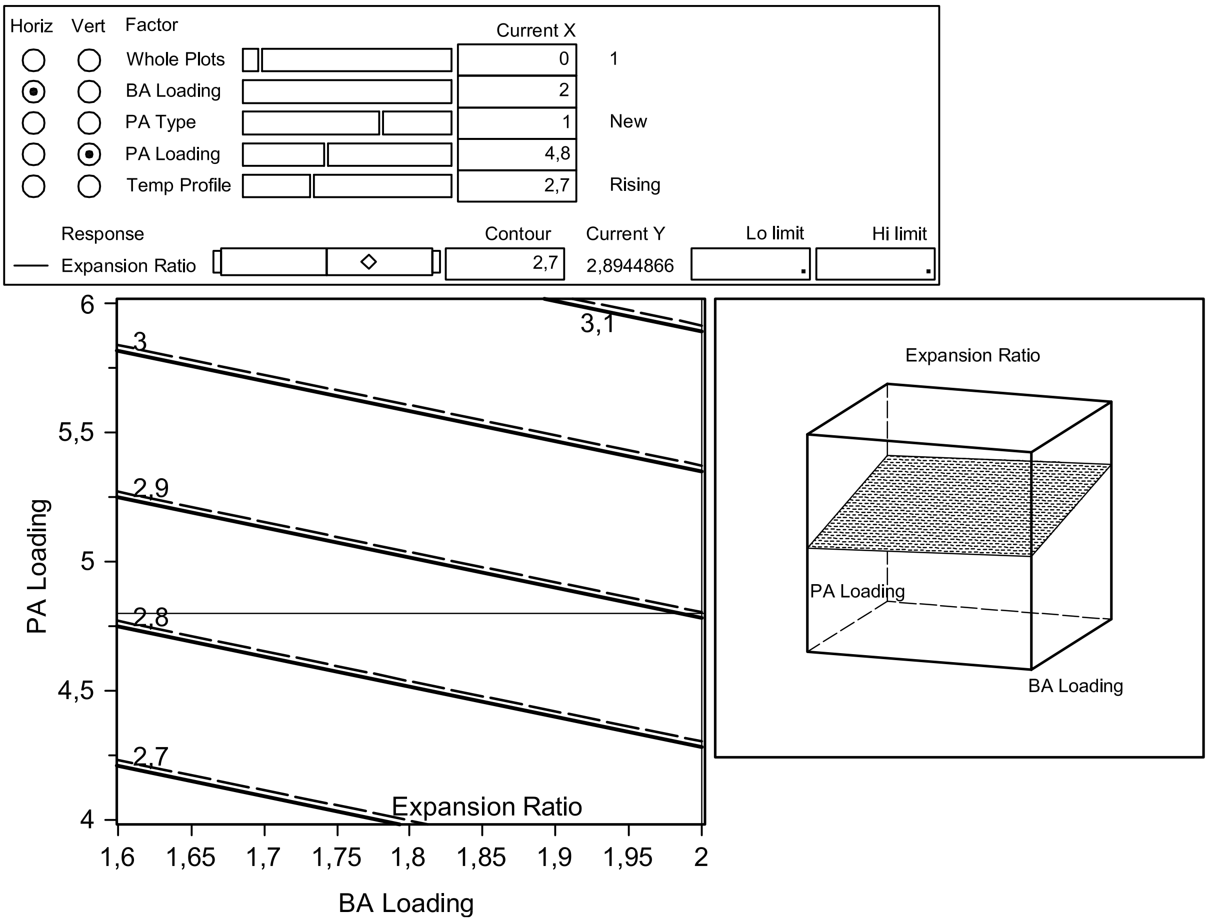


Figure D.5 — Contour plot for expansion ratio

D.10 Confirmation results

The recommended settings were confirmed by an additional experiment where the expansion ratio value fell within the confidence interval.

Thus, it was possible to use less of the new PA product than was necessary with the old product with an improved and optimized processing condition.

Annex E
(informative)

Genetic algorithms for DNA sequencing experiment[11)]

E.1 General

Genetic algorithms have proven useful in solving the DNA sequence assembly problem for small sequence lengths in support of sequencing entire chromosomes. Genetic algorithms are optimization schemes that have a stochastic component for generating potentially improved solutions. The performance of these algorithms could be improved through careful determination of the tuning parameters.

E.2 Overall objective for the experiment

It was hoped in this application that by determining critical settings, larger DNA sequencing problems could be solved. More specifically, the objective was to determine improved settings of the tuning parameters on a 10 000 length DNA sequence assembly problem for use in solving a more challenging problem of length 35 000.

E.3 Description of the process

DNA strands of interest are on the order of 40 000 to 400 000 bases in length. However, DNA strands longer than 1 000 base pairs cannot be accurately sequenced routinely. Consequently, large strands of DNA are broken into smaller pieces for more manageable sequencing. However, taking advantage of automated sequencing of shorter strands generates a new problem — re-assembly of the sequence given that the fragment contains only its base sequence of letters but not the original location, orientation and strandedness. A specific version of a genetic algorithm (a heuristic numerical algorithm for optimization) can be used to arrange the fragments into a full strand. Roughly speaking, a population of permutations of the strands is created and then subsequent generations of solutions are generated using parameterized operators. For assessing performance of a solution, the quality of the solution (fitness value) obtained after 1 000 000 trials was considered. The solution was already known to this problem, but the idea was to determine the settings that expedited its determination. The optimal values of these parameter settings would then be applied to a 35 000 length problem that hitherto had resisted a reasonable solution.

E.4 Response variable

E.4.1 Choice of variable

The response is a fitness value that assesses the performance of the algorithm following 1 000 000 trials.

E.4.2 Measurement of the response variable

The value of the response variable is a direct output of the simulation code.

11) This example was drawn from the article "Genetic Algorithms for DNA Sequencing," by Rebecca J. Parsons and Mark E. Johnson, *American Journal of Mathematical and Management Sciences*, **17**, 1997, pp. 369-396. Included here with permission of the publisher American Sciences Press, Inc., Columbus, Ohio.

E.4.3 Relationship of the response variable to the objective of the experiment

The response variable is an indication of the speed with which the genetic algorithm can make progress toward a solution. Eventually, a specific setting of parameters will lead to a global solution. Ultimately, the solution found at convergence should make biological sense and should be attained in a useful span of time.

E.5 Factors affecting the response

E.5.1 Description of each factor (continuous/discrete) to be varied

The factors chosen to be varied within the experiment were identified by a computer scientist who was an expert in the area of genetic algorithms. These factors were:

A: Inversion rate;

B: Mutation rate;

C: Transposition rate;

D: Crossover rate.

All of these factors are of a continuous type.

E.5.2 Selection of levels (related to size of effect to be determined)

The factors to be used in the experiment and their associated levels are shown in Table E.1.

Table E.1 — Factors and their associated levels

Factor	Level 1	Level 2
A: Inversion rate	0,38	0,28
B: Mutation rate	0,14	0,04
C: Transposition rate	0,38	0,28
D: Crossover rate	0,5	0,3

The rates chosen were set to push the boundaries of the conventional wisdom values. These levels were finalized by the computer scientist in consultation with a statistician.

E.5.3 Other factors noted but not incorporated due to issues with relevance

Other factors that might have been considered, but were not included in the experiment, are shown in Table E.2.

Table E.2 — Factors not included in the experiment

Factor	Reason for exclusion
Population size	The number of solutions produced from generation to generation drives the total computation time. Population size should not interact with the tuning parameters. Population size and number of generations determines the total computational effort.

E.6 Full factorial design

E.6.1 Design matrix (standard order, run order)

The design selected was full factorial that allowed for all interactions to be examined.

Table E.3 shows the design.

Table E.3 — Design matrix

A: Inversion rate	B: Mutation rate	C: Transposition rate	D: Crossover rate	Standard order	Run order
0,38	0,14	0,38	0,5	1	1
0,28	0,14	0,38	0,5	2	9
0,38	0,04	0,38	0,5	3	10
0,28	0,04	0,38	0,5	4	2
0,38	0,14	0,28	0,5	5	11
0,28	0,14	0,28	0,5	6	3
0,38	0,04	0,28	0,5	7	4
0,28	0,04	0,28	0,5	8	12
0,38	0,14	0,38	0,3	9	13
0,28	0,14	0,38	0,3	10	5
0,38	0,04	0,38	0,3	11	6
0,28	0,04	0,38	0,3	12	14
0,38	0,14	0,28	0,3	13	7
0,28	0,14	0,28	0,3	14	15
0,38	0,04	0,28	0,3	15	16
0,28	0,04	0,28	0,3	16	8

Randomization of run order was unnecessary here since the results for each combination entailed massive amounts of simulations in the course of running the experiment. The randomness of the uniform random generator was reasonably presumed to be satisfactory regardless of the sub-stream in use.

The run order given was chosen to facilitate analysis after the first eight runs were completed. Subsequently, the other eight runs were made, followed by a full replication. Eight runs required a few days of computation, so that the intermediate analysis allowed an early intervention point if there had been a problem.

E.6.2 Centre points

No centre points were run at this stage of the experiment.

E.6.3 Replication and repetition

Each combination of tuning parameters was run twice.

E.7 Analysis of results

E.7.1 Data acquired through the experiment and analysis considerations

The results from the experiment are given in Table E.4. They were keyed into a software program called JMP [12].

Table E.4 — Experiment results

A: Inversion rate	B: Mutation rate	C: Transposition rate	D: Crossover rate	Standard order	Fitness values
0,38	0,14	0,38	0,5	1	45 281
0,28	0,14	0,38	0,5	2	43 892
0,38	0,04	0,38	0,5	3	46 739
0,28	0,04	0,38	0,5	4	45 565
0,38	0,14	0,28	0,5	5	43 439
0,28	0,14	0,28	0,5	6	44 798
0,38	0,04	0,28	0,5	7	43 866
0,28	0,04	0,28	0,5	8	44 250
0,38	0,14	0,38	0,3	9	48 891
0,28	0,14	0,38	0,3	10	49 191
0,38	0,04	0,38	0,3	11	52 495
0,28	0,04	0,38	0,3	12	52 671
0,38	0,14	0,28	0,3	13	50 393
0,28	0,14	0,28	0,3	14	50 492
0,38	0,04	0,28	0,3	15	53 107
0,28	0,04	0,28	0,3	16	53 116
0,38	0,14	0,38	0,5	17	44 207
0,28	0,14	0,38	0,5	18	43 950
0,38	0,04	0,38	0,5	19	46 755
0,28	0,04	0,38	0,5	20	45 589
0,38	0,14	0,28	0,5	21	44 438
0,28	0,14	0,28	0,5	22	43 026
0,38	0,04	0,28	0,5	23	45 830
0,28	0,04	0,28	0,5	24	49 906
0,38	0,14	0,38	0,3	25	49 173
0,28	0,14	0,38	0,3	26	51 601
0,38	0,04	0,38	0,3	27	52 193
0,28	0,04	0,38	0,3	28	52 378
0,38	0,14	0,28	0,3	29	51 618
0,28	0,14	0,28	0,3	30	52 212
0,38	0,04	0,28	0,3	31	49 795
0,28	0,04	0,28	0,3	32	52 261

12) JMP is the trade name of a product supplied by SAS Inc. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of the product named. Equivalent products may be used if they can be shown to lead to the same results.

All effects were estimated and assessed owing to the replication.

E.7.2 Estimation of effects

The output in Figure E.1 indicates that factors D (crossover rate) and B (mutation rate) were strongly significant and there appeared to be no significant interactions. Neither inversion nor transposition rate influenced the fitness values significantly. The following normal probability plot of effects (see Figure E.2) also confirms this observation.

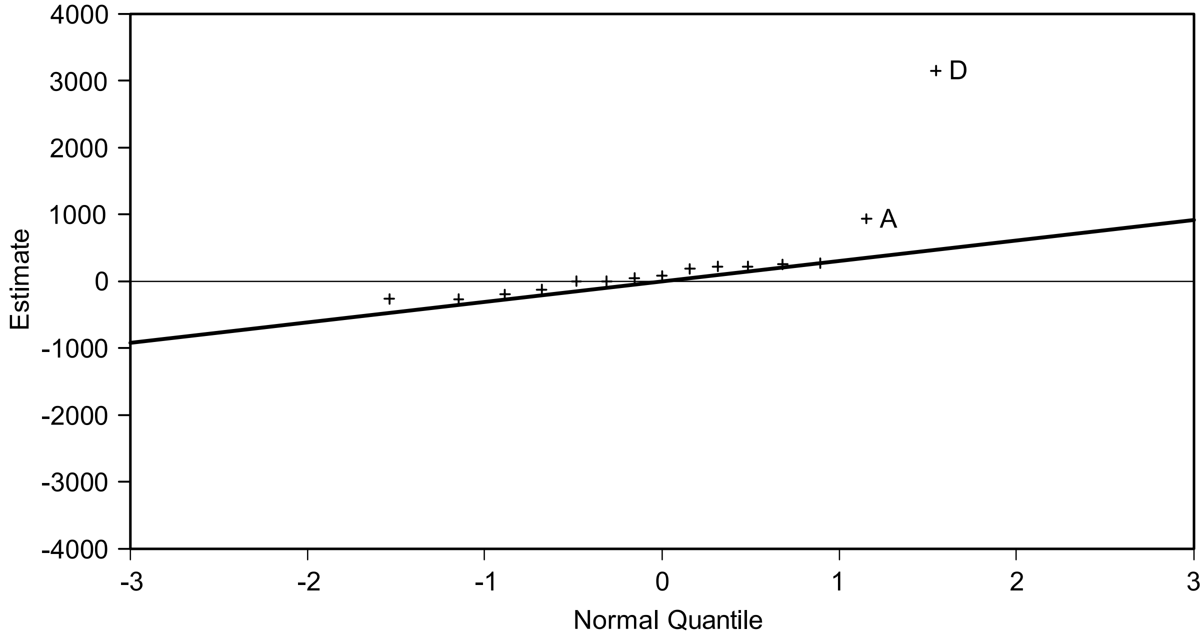
Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	15	356378836	23758589	11.9979
Error	16	31683680	1980230	Prob > F
C. Total	31	388062516		<.0001*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	48222.438	248.7613	193.85	<.0001*
A	208.6875	248.7613	0.84	0.4139
B	934.8125	248.7613	3.76	0.0017*
A*B	101.0625	248.7613	0.41	0.6899
C	61.75	248.7613	0.25	0.8071
A*C	264.75	248.7613	1.06	0.3030
B*C	-202.625	248.7613	-0.81	0.4273
A*B*C	292.375	248.7613	1.18	0.2571
D	3126.75	248.7613	12.57	<.0001*
A*D	182.375	248.7613	0.73	0.4741
B*D	-32	248.7613	-0.13	0.8992
A*B*D	-137.625	248.7613	-0.55	0.5877
C*D	213.3125	248.7613	0.86	0.4038
A*C*D	-259.8125	248.7613	-1.04	0.3118
B*C*D	-254.6875	248.7613	-1.02	0.3212
A*B*C*D	-33.0625	248.7613	-0.13	0.8959

Figure E.1 — Estimation of effects



The slanted line is the Lenth's pseudo standard error (PSE), from the estimates population.

The horizontal line is the root mean squared error (RMSE), from the residual.

Figure E.2 — Normal plot

E.7.3 Pareto chart of effects (see Figure E.3)

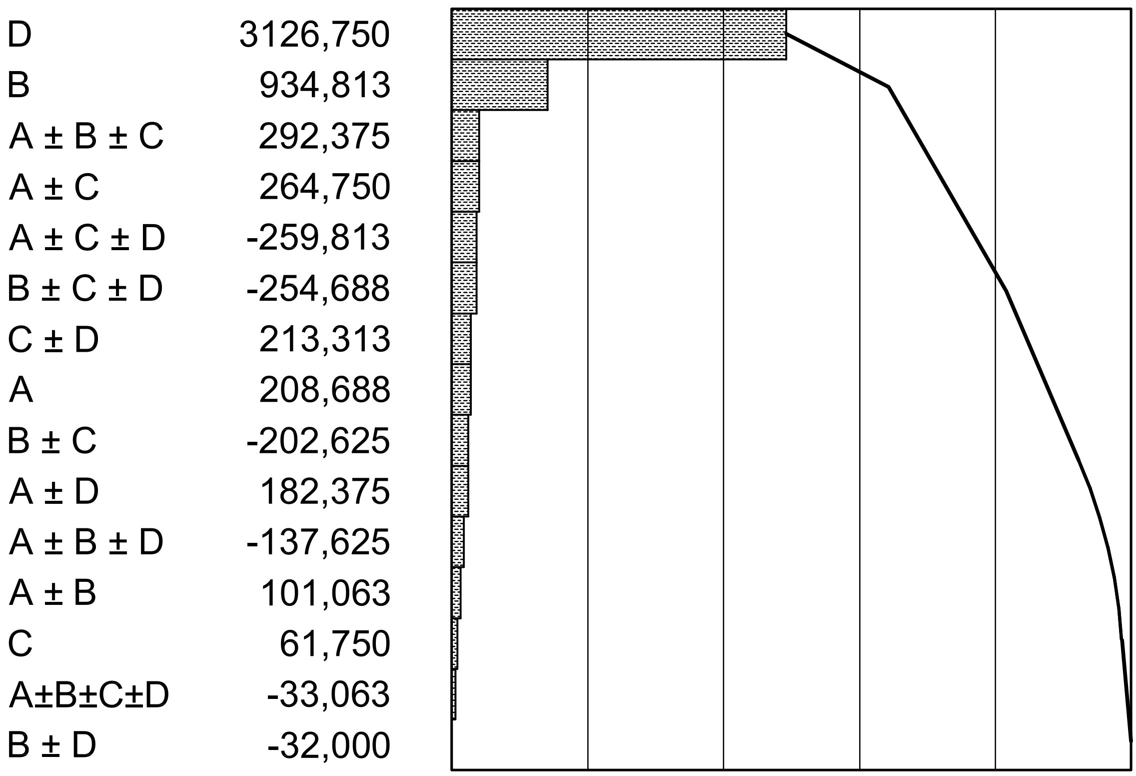


Figure E.3 — Pareto chart of effects

E.7.3.1 Estimation of experimental error and standard error of effects

The standard error of effects was 248,761 3.

E.8 Presentation of results — Effects plots

The relative slopes of the lines in Figure E.4 are consistent with the significance of the factors D and B.

Figure E.5 is an interaction plot between the factors. There were no significant interactions of note since all of the pairs of lines are approximately parallel.

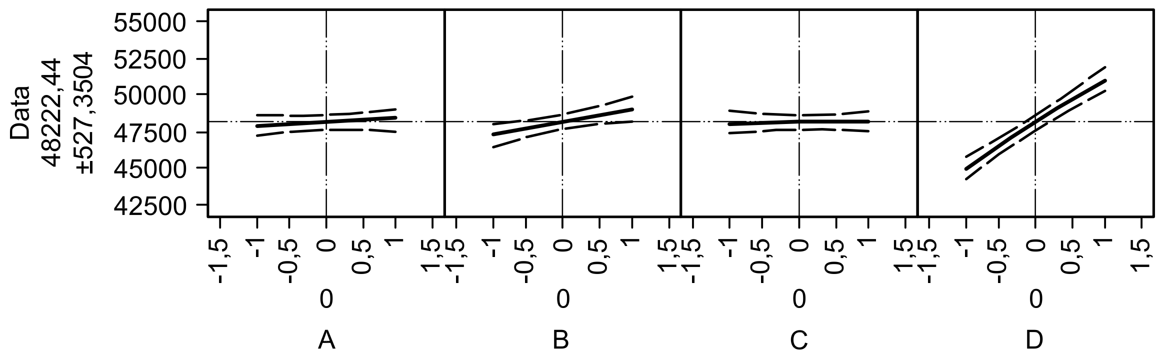


Figure E.4 — Effects plots

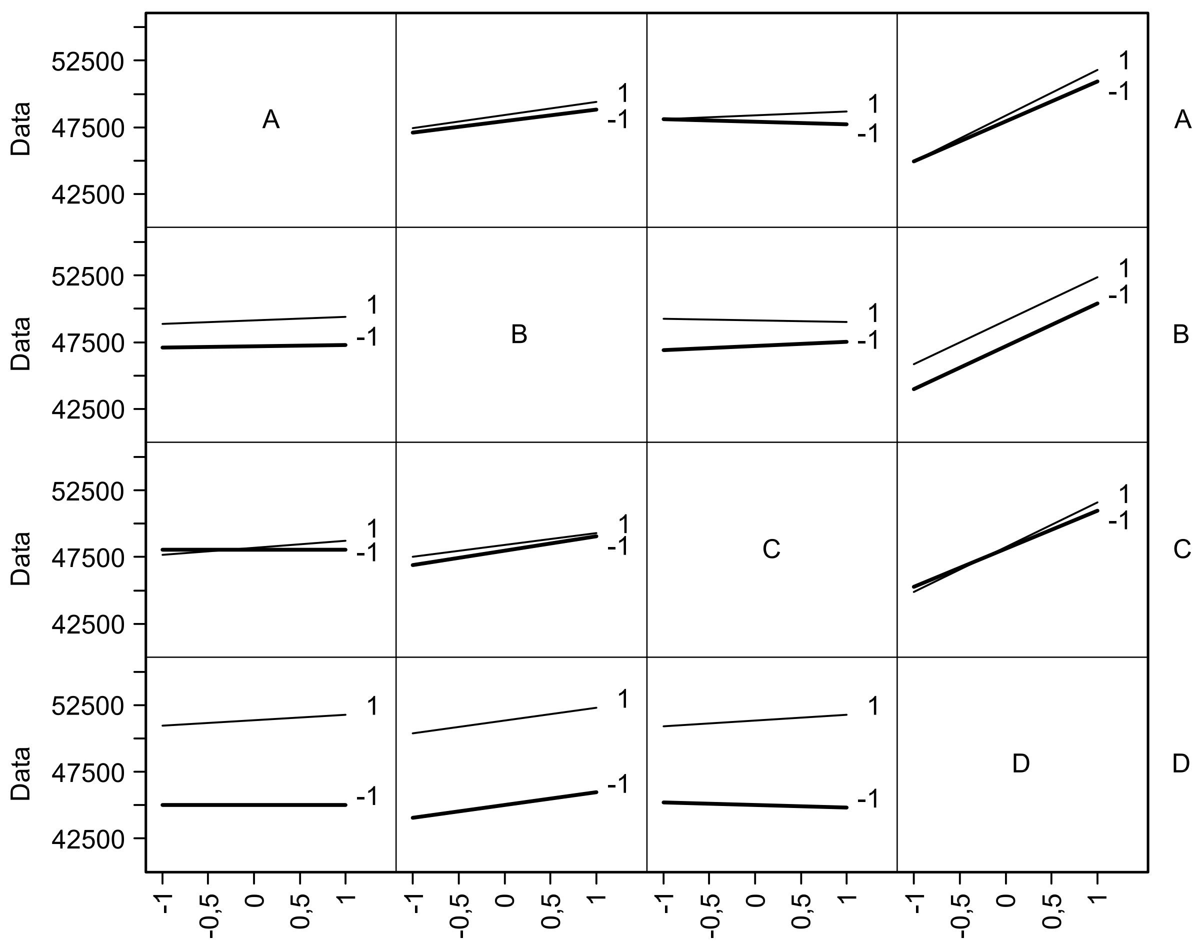


Figure E.5 — Interaction plot between the factors

E.9 Presentation of results — Optimization recommendations

The preferred settings from this experiment are given in Table E.5.

Table E.5 — Preferred settings from experiment

Factor	Setting
A: Inversion rate	+1 (0,28)
B: Mutation rate	+1 (0,04)
C: Transposition rate	−1 (0,38)
D: Crossover rate	+1 (0,30)

The settings of B and D are driven by their effect estimates. Although factors A and C were not significant as main effects, the replication indicated smaller variability in fitness values for the high level of A and low level of C.

E.10 Confirmation results

The full factorial experiment identified two tuning parameters having a major influence on the performance of the genetic algorithm. A follow-up experiment was conducted to further refine the optimal tuning parameters. In particular, a central composite design centred at a mutation rate of 3,2 % and a crossover rate of 26,8 % identified a reasonably flat region in the fitness performance space that then facilitated further investigation of the population size parameter. Following this order of magnitude improvement to solving medium-sized sequence assembly problems, large size problems were considered and were either solved completely or near-optimal solutions were obtained.

Bibliography

[1] BOX, G.E.P., HUNTER, J.S. and HUNTER, W.J. *Statistics for Experimenters – Design, Innovation, and Discovery.* Wiley, New York, 2005

[2] WU, C.F.J. and HAMADA, M. *Experiments: Planning, Analysis, and Parameter Design Optimization*. Wiley, New York, 2000

[3] OEHLERT, G.W. *A First Course in Design and Analysis of Experiments.* W.H. Freeman and Company, New York, 2000

[4] DEAN, A. and VOSS, D. *Design and Analysis of Experiments.* Springer-Verlag, New York, 1999

[5] MILLIKEN, G.A. and JOHNSON, D.E. *Analysis of Messy Data, Volume 1: Designed Experiments.* Chapman & Hall, New York, 1992

ICS 03.120.30

Price based on 52 pages