SOFTWARE FOR QUALITY-IMPROVEMENT EXPERIMENTS: AN EVALUATION

David D. Koch and William J. Welch

IIQP Research Report
RR-89-07

August 1989

SOFTWARE FOR QUALITY-IMPROVEMENT EXPERIMENTS: AN EVALUATION

David D. Koch and William J. Welch

Department of Statistics and Actuarial Science
University of Waterloo

ABSTRACT

We take an in-depth look at six packages for the analysis of designed experiments that often arise in quality improvement. The packages are evaluated on a number of test problems, including full and fractional factorials and parameter design for a robust product or process. The goal is to find one or more packages that are easy to use, yet comprehensive enough to handle a wide range of experiments. Finally we recommend some choices.

Key Words: *Design of Experiments, Factorial Design, Fractional-Factorial Design, Parameter Design, Quality Improvement.*

Introduction

Box and Bisgaard (1987) pointed out that a major difference between Japan and North America in the use of statistical methods (including designed experiments) is that "they do it and we don't". Although many corporations have started initiatives to increase the number of designed experiments, there is undoubtedly still a long way to go. We feel that software for design and analysis of experiments can make the process of experimentation less painful and so encourage more experimentation. Our goal, then, is to find an easy-to-use software package that can analyze a wide range of simple experimental designs, as encountered in quality improvement.

Previous software reviews by Hahn (1988) and Nachtsheim (1987) provided valuable guidance to a bewildering and expanding market by summarizing features of various experimental-design packages. Vendors, in their documentation and advertising, however, naturally tend to emphasize what their products can do, rather than what is missing or awkward to use. Only by attempting a wide range of test problems can deficiencies be identified. Our study is also rather more focused on the analysis of factorial, fractional-factorial, and analysis of variance (ANOVA) experiments that might arise in quality improvement, including parameter design for product or process robustness. We do not consider automatic generation of experimental plans (e.g., Welch, 1985), response-surface methodology, or other more specialized topics covered by Hahn and Nachtsheim.

Our original motivation was to find software for integration into a 10-day, design-of-experiments course given to professional engineers. A wide range of tutorial examples was already available. In as much as these examples cover basic quality-improvement experiments that could arise in practice, the results of our study should be of wider interest. We have in mind engineers with some exposure to the basics of statistical design of experiments but modest training in statistics otherwise. As an engineer might use this type of software infrequently, ease of use is of some importance.

The next section outlines the test problems. Then we summarize the performances of six packages on these examples. Finally, we make some specific recommendations about these offerings and suggest some desirable features that we would like vendors to incorporate.

The Test Problems

The test problems cover two-level fractional factorials, including their design, three-level fractional factorials, Taguchi's (1986) parameter design for product and process robustness, and ANOVA. The nine problems, ordered with the most fundamental tasks first, are:

1. Find the levels of five factors in a 2^{5-1} fractional-factorial design that maximize crease retention of paper, a larger-the-better quality characteristic.

2. Analyze a 2^3 factorial design, replicated twice over days, to minimize porosity of castings.

3. Design a 2^{9-4} experiment in two blocks, such that four specified two-factor interactions are estimable. The ultimate objective of the experiment is to maximize a pull strength.

4. Find levels of three factors in a 3^{3-1} fractional factorial experiment to improve the microfinish of a cast aluminum part.

5. Analyze the experiment described by Pignatiello and Ramberg (1985). There are four controllable factors in a 2^{4-1} fractional factorial. For each factor-level combination there are six replicate observations, three each at two levels of an uncontrollable noise factor. The objective is to find levels of the controllable factors such that the free height of a leaf spring is consistently close to a target of eight inches. Thus, the mean free height should be near the target and the standard deviation over the six replicates should be small.

6. Find the levels of two factors in a 3^2 factorial experiment that bring the paint thickness in a powder-painting operation to a nominal value. There are four replicates, corresponding to a noise factor for position on the rack, and, as in the previous problem, we want consistency across replicates.

7. Identify any differences between nine racks used in a spraying operation (a one-way ANOVA).

8. Identify the effects of two factors arranged in two-way (2×3) ANOVA on the yield of a process. There are two replicates per factor-level combination.

9. Make paired and unpaired comparisons of two gauges.

The factorial and fractional-factorial designs are of primary interest. To analyze the two-level factorial designs (Problems 1 and 2) we expect a package to provide information on the aliasing (confounding) structure, effects estimates (or equivalent means), normal plots of the effects, and an ANOVA table. For presentation purposes, it is also useful to have plots of the level means for main effects, interaction plots showing the effect of factor A at each level of B, say, and possibly cube plots to show the mean yield at any combination of levels of three factors.

Only for two-level designs do we expect advice on generating the experimental plan. Designing the 2^{9-4} fractional factorial in Problem 3 is fairly demanding in this respect. We expect at least the alias structure and help perhaps in assigning factors to the columns of the design so that the desired interactions can be estimated.

In the case of three-level factorials (Problem 4), an analysis of variance table and effects estimates or level means for main effects and two-factor interactions are a bare minimum. A decomposition of the sums of squares for a quantitative factor into linear and quadratic components is also helpful.

In Problems 5 and 6, the objective is to minimize process variability as well as move the mean on target by Taguchi's (1986) parameter design. In these experiments, replicate observations on the quality characteristic are collapsed to a Taguchi S/N (signal-to-noise) ratio. Finding levels of the controllable process parameters that maximize the S/N ratio is aimed at reducing sensitivity to uncontrollable noise (e.g., manufacturing variability). An alternative would be to analyze $\log s$, where s is the usual sample standard deviation computed from the replicates. In both cases, it is also often necessary to compute and analyze $\bar{y}$, the mean across replicates, to move the mean onto target. For these experiments, then, we are looking for a capability to easily compute the S/N ratio, $\log s$, and $\bar{y}$ across replicates. Thereafter, the calculation of effects, normal plots of effects, and so on are much the same as for other factorial designs.

We view the one-way and two-way ANOVA experiments (Problems 7 and 8) and the two-treatment comparisons (Problem 9) as of secondary importance. Nonetheless, a user would not want to learn a new package for such analyses, and we include them to provide an indication of comprehensiveness.

The Packages

Based on the Hahn (1988) and Nachtsheim (1987) reviews, newsletter advertisements, and a software directory in the March 1989 *Quality Progress*, we narrowed the search down to six packages: ANOVA-TM, CADE (Discovery module), Design-Ease, Jass, SCA (Quality and Productivity Improvement module), and STATGRAPHICS. CADE, Design-Ease, and Jass have little or no capability beyond two-level designs and, therefore, cannot tackle some of the above problems at all. We included them, however, in the anticipation that they would excel at their specialization. We only considered packages running on IBM PCs or compatibles, as these machines are the most universally available.

When a capability is missing but one can get around the difficulty with some inconvenience, it may be difficult to draw the line. Therefore, when we claim a task cannot be performed, we mean that we could not find a simple way of achieving the goal. Even after many frustrating days reading the documentation, it is possible that a fine point or two has escaped our attention.

The performances of the six packages on the above test problems are summarized in Table 1. A check mark indicates that a package made a useful contribution to the analysis of a problem; some parts may have been incomplete. Table 2 gives a finer categorization of capabilites, organized by problem type.

We now highlight some of the major strengths and weakesses identified during our experiences applying the packages to the test problems.

ANOVA-TM

As the name implies, ANOVA-TM specializes in the implementation of Taguchi's (1986) methods, including parameter design for product and process robustness to sources of noise. It is convenient to use due to its extensive menu and function key format, and, in contrast to some of the other menu-driven packages considered here, ANOVA-TM allows considerable user flexibility.

ANOVA-TM does not have the capacity to build its own designs based on user information about the numbers of factors, blocks, and runs. The user is expected to choose from a menu of orthogonal arrays (designs), and then assign main effects and interactions to the columns of the array, presumably by use of Taguchi's linear graphs. Because ANOVA-TM allows one to edit any orthogonal array, it is also possible to analyze user-defined two and three-level full and fractional factorials. However, we could not find a way to do the ANOVA problems (7 and 8 on our list).

Relative to the other packages considered here, ANOVA-TM generates the most useful data sheets for aid in data recording. There is a data sheet for each run with plenty of space for comments. Also, it is the only package that understands the concept of inner and outer arrays, for control and noise factors respectively, so it can include the levels of the noise factors on the data sheets. The data sheets are not in randomized order, though, so they have to be shuffled manually.

In addition to pre-defined S/N ratios, analysis of $\bar{y}$ and $\log s$ is possible through user-defined S/N ratios. As already mentioned, ANOVA-TM allows considerable user flexibility. There are many options throughout, such as the ability to manually or automatically pool sums of squares into the error term in the analysis of variance table, a feature unique among our test packages. We were a bit surprised that normal-effects plots are not available for two-level designs, only level-means plots for main effects and interactions. For three-level designs, though, this is the only package that can decompose sum of squares into linear and quadratic components. Alias structures and effects estimates are impossible in either case.

On the whole, we were quite impressed with the rather comprehensive capabilities of ANOVA-TM as well as its appealing interface. It is also the only package that can handle both control and noise arrays. A few extra features, like normal-effects plots and residual analysis, would make it even better.

CADE (Computer Aided Design of Experiments)

We only consider the CADE Discovery module here, which specializes in two-level factorial experiments. (The companion Optimization module analyzes response surface designs.) Again, it is largely function-key driven, although a query format is used to input design parameters such as the numbers of runs and factors.

CADE has a very restricted catalogue of designs, nor will it analyze a design chosen by the user. For experiments with 5–11 factors, it automatically selects a fractional factorial with an option to replicate individual runs in the design. For a four-factor experiment, there is a choice of either a full factorial or an eight-run fractional factorial, again with an option for replication of some runs. For designs involving only two or three factors, full factorials with complete replication are available. Only for full factorials is blocking possible, a major disadvantage in our opinion. To illustrate this inflexibility, CADE could not complete Problem 3 to design an experiment for nine variables in 16 runs, divided into two blocks; for nine variables CADE only permits a 32-run design, and blocking is not allowed.

CADE produces a "log sheet" for data retrieval before the analysis phase which must be printed out to view the factor level combinations for a given run. One confusing aspect of data entry is that the manual instructs the user to input responses in the order in which they appear on the log sheet, which is in standard order. However, the screen displays a column of blanks in what seems to be ascending randomized run order from top to bottom.

The analysis is not as complete as we would like to see, either, as no main-effects plots or interaction plots are possible, only two way tables of means for interactions. Furthermore, the half normal plots would not be suitable for presentation. Finally, if $\bar{y}$, $\log s$, or S/N ratios are to be analyzed, the user needs to compute them before entering CADE.

Unfortunately, input is unnecessarily case sensitive, and the query format for entering design parameters is tiresome as there are no defaults. Combined with the lack of flexiblity in choosing designs, we were rather disappointed with this package.

Design-Ease

Design-Ease, like CADE, is aimed at leading the user through all steps in the design and analysis of two-level (fractional) factorials. Being menu and function key driven, it is extremely user friendly.

Many of the packages reviewed here claim some affiliation with the popular text by Box, Hunter, and Hunter (1978) on experimental design. Design-Ease actually begins with a screen version of the table on page 410 of the text for generating fractional factorials, although generators and resolution information do not appear until a later screen. The user selects a design and then requests the program to generate a data sheet which can be printed out at any time for data retrieval. Unlike CADE, Design-Ease suggested an appropriate design for our nine-factor experiment in 16 runs of two blocks, though we had to assign factors to columns by trial and error to obtain a plan that will estimate the four interactions. If the problem had called for four blocks, each of four runs (with fewer interactions), Design-Ease would not have coped, so there is some lack of flexibility.

The program leads the user through the analysis by keeping options hidden until they become relevant. Plots are of satisfactory graphic quality, except that the vertical scales usually have only the minimum and maximum values indicated. A nice feature of the normal-effects plots is the ability to adjust the straight line fitted through the "error" effects after Design-Ease suggests an initial fit. The user may then manually scan the outstanding effects for use in the analysis of variance table later. Again, the user must calculate $\bar{y}$, $\log s$, and S/N ratios manually for Taguchi's parameter design.

In summary, although some flexibility is missing, Design-Ease does its intended job competently—designing and analyzing two-level (fractional) factorials.

Jass

Jass is the third and last package that specializes in two-level factorial designs. It can be used in two modes, query for novices and command for experienced user.

Jass features a "design" section to create a design based on user-given factor and run information. As with Design-Ease, we had to finish the design manually in Problem 3. This is one of the few packages here which allows the user to specify generators for both the treatment and blocking factors. Hence, it is also possible to analyze two-level user-defined designs in the "twolevel" section either by specifying a modification of a system-defined design or by simply reading in a design already performed. One may create report forms for some so-called "nonstandard designs" (with some factors at more than two levels), but, without any capacity for analysis, this seems like an empty feature.

Jass' analysis of two-level designs includes normal plots, interaction plots, and cube plots as well as residual analysis, but we were surprised that it would not generate an ANOVA table. It also seems strange that one should be able to obtain an interaction plot but not the simpler main-effects plot.

Jass seems well thought out in some ways such as the way it intelligently specifies the most natural default values in its query mode, and yet it does not provide some basics such as an ANOVA table and main-effects plots. We are not too excited about its interface either. Although we like the flexibility to specify generators, the reader only interested in two-level designs is probably best referred back to Design-Ease.

SCA

SCA is a command driven package with a comprehensive range of capabilities where commands are entered in "sentence" and "paragraph" formats. We only consider the Quality and Productivity Improvement module.

This package grants the greatest freedom of any here in terms of allowing the user to modify any two-level system-generated designs. After a design is suggested by the system, one may add or drop columns, change signs on one or more columns, and specify new generators for either treatment or blocking factors. This was convenient for designing the experiment in Problem 3, though we had to choose an appropriate blocking scheme. The system will not suggest one itself.

Although cube plots and normal-effects plots may be generated, they are not of presentation quality. Main-effects and interaction plots are unavailable. SCA can conveniently calculate $\bar{y}$, $\log s$, and S/N ratios. All analysis of variance tables, including analysis of three-level factorials, are generated from a regression model, which is a little cumbersome.

We certainly like the flexibility in generating two-level factorial designs, but this does not carry through to the analysis. The command-driven interface makes

SCA very difficult to use relative to some of the other packages, and this offsets its comprehensiveness in our opinion.

STATGRAPHICS

In terms of overall analysis capability, this is the most complete package of any reviewed here. A large menu and several submenus, plus a good selection of prompts with every screen, provide an efficient and friendly environment.

Like many of the other packages, STATGRAPHICS easily generates two-level factorial designs, though once again we had to finish Problem 3 manually. Only when the data have been entered is the alias structure conveniently available. Prior to this, the aliasing can only be determined via a rather verbose correlation matrix, or by making up some temporary data.

Graphic capabilities are superb and suitable for presentation purposes when printed. However, effects are not labelled on normal plots. Main-effects and interaction plots are available, although the latter have the two lines offset rather than overlaid.

Computation of $\bar{y}$, $\log s$, and S/N ratios is possible through the file operations submenu. The multifactor analysis of variance submenu permits analysis of many types of design, but in the case of user-defined two and three-level factorials, estimates of effects have to be calculated by hand from the means table. Effects are only computed for system-generated, two-level factorials.

Finally, by interfacing with STSC's APL*PLUS, extensions to STATGRAPHICS are almost endless to anyone with this computing background. Even without resorting to programming, STATGRAPHICS could tackle nearly all our test tasks.

Recommendations and Discussion

Our goal was to find a package that is easy to use, yet comprehensive enough to handle a wide range of quality-improvement experiments. However, none of the packages could complete every part of every test problem. One would probably have to resort to one of the more powerful general-purpose packages, such as SAS or Systat, to fill in the missing parts. But, for the friendliness of a menu-driven package and a reasonable attempt at all our test problems, STATGRAPHICS is our recommendation. Those who prefer Taguchi's terminology and do not need ANOVA capabilities (e.g., for gauge studies) might favour ANOVA-TM. Design-Ease is competent at its specialization, two-level fractional factorials, but we feel most users will eventually want to run three-level designs.

One glaring weakness shared by these packages is that none could complete Problem 3 to design a two-level fractional-factorial experiment. Given information on the numbers of factors, runs, and blocks, and specified two-factor interactions, a package should be able to generate an experiment plan without the user having to manually assign factors to columns. Franklin (1985) described methodology for generating such plans for p^{n-m} fractional factorials, and Turiel (1988) presented a computer implementation. Thus, two and three-level factorial plans can be generated automatically, but experiments with factors at differing numbers of levels still cause some problems. No doubt we will see these algorithms incorporated before too long.

Acknowledgments

This research was funded by General Motors of Canada, the Natural Sciences and Engineering Research Council, and the Institute for Improvement in Quality and Productivity at the University of Waterloo.

References

BOX, G.E.P. and BISGAARD, S. (1987). "The Scientific Context of Quality Improvement". *Quality Progress* June 1987, pp. 54–61.

BOX, G.E.P., HUNTER, W.G., and HUNTER, J.S. (1978). *Statistics for Experimenters.* Wiley, New York.

FRANKLIN, M.F. (1985). "Selecting Defining Contrasts and Confounded Effects in p^{n-m} Factorial Experiments". *Technometrics* 27, pp. 165–172.

HAHN, G.J. and MORGAN, C.B. (1988). "Design Experiments With Your Computer". *Chemtech* November 1988, pp. 664-669.

NACHTSHEIM, C.J. (1987). "Tools for Computer-Aided Design of Experiments". *Journal of Quality Technology* 19, pp. 132–160.

PIGNATIELLO, J.J. and RAMBERG, J.S. (1985). "Discussion of Off-Line Quality Control, Parameter Design, and the Taguchi Method by R.N. Kackar". *Journal of Quality Technology* 17, pp. 198–206.

TAGUCHI, G. (1986). *Introduction to Quality Engineering.* Asian Productivity Organization, Tokyo.

TURIEL, T.P. (1988). "A FORTRAN Program to Generate Fractional Factorial Experiments". *Journal of Quality Technology* 20, pp. 63–72.

WELCH, W.J. (1985). "ACED: Algorithms for the Construction of Experimental Designs". *The American Statistician* 39, p. 146

TABLE 1. Summary of Package Performances on the Test Problems
($\surd$= useful analysis, $\times$= no analysis)

Problem	ANOVA-TM	CADE	Design-Ease	Jass	SCA	STAT-GRAPHICS
1	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$
2	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$
3	$\times$	$\times$	$\surd$	$\surd$	$\surd$	$\surd$
4	$\surd$	$\times$	$\times$	$\times$	$\surd$	$\surd$
5	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$
6	$\surd$	$\times$	$\times$	$\times$	$\surd$	$\surd$
7	$\times$	$\times$	$\times$	$\times$	$\surd$	$\surd$
8	$\times$	$\times$	$\times$	$\times$	$\surd$	$\surd$
9	$\times$	$\times$	$\times$	$\surd$	$\surd$	$\surd$

TABLE 2. Package Capabilities, Organized by Problem Type

($\surd$= capability present, $\times$= capability absent)

(a) Two-Level Factorials

Task	ANOVA-TM	CADE	Design-Ease	Jass	SCA	STAT-GRAPHICS
Menu of designs	$\surd$	$\times$	$\surd$	$\times$	$\times$	$\times$
Build design	$\times$	$\surd$	$\times$	$\surd$	$\surd$	$\surd$
Analyze user's design	$\surd$	$\times$	$\times$	$\surd$	$\surd$	†
Alias structure	$\times$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$
Data sheet	$\surd$	$\surd$	$\surd$	$\surd$	$\times$	$\times$
Effects estimates or means	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$
ANOVA table	$\surd$	$\surd$	$\surd$	‡	$\surd$	$\surd$
Normal plot	$\times$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$
Main-effects plot	$\surd$	$\times$	$\surd$	$\times$	$\times$	$\surd$
Interaction plot	$\surd$	$\times$	$\surd$	$\surd$	$\times$	$\surd$
Cube plot	$\times$	$\surd$	$\surd$	$\surd$	$\surd$	$\times$
Residual analysis	$\times$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$

(b) Three-Level Factorials

Task	ANOVA-TM	CADE	Design-Ease	Jass	SCA	STAT-GRAPHICS
Menu of designs	$\surd$	$\times$	$\times$	$\times$	$\times$	$\times$
Build design	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$
Analyze user's design	$\surd$	$\times$	$\times$	$\times$	$\surd$	†
Alias structure	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$
Data sheet	$\surd$	$\times$	$\times$	$\surd$	$\times$	$\times$
Effects estimates or means	$\surd$	$\times$	$\times$	$\times$	$\surd$	$\surd$
ANOVA table	$\surd$	$\times$	$\times$	$\times$	$\surd$	$\surd$
Lin. / quad. decomposition	$\surd$	$\times$	$\times$	$\times$	‡	‡
Main-effects plot	$\surd$	$\times$	$\times$	$\times$	$\times$	$\surd$
Interaction plot	$\surd$	$\times$	$\times$	$\times$	$\times$	$\surd$
Residual analysis	$\times$	$\times$	$\times$	$\times$	$\surd$	$\surd$

(c) Calculation of $\bar{y}$, $\log s$, and S/N Ratios Over Replicates

Task	ANOVA-TM	CADE	Design-Ease	Jass	SCA	STAT-GRAPHICS
Calculate $\bar{y}$	√	×	×	√	√	√
Calculate $\log s$	√	×	×	√	√	√
Calculate S/N	√	×	×	√	√	√

(d) One-Way and Multi-Factor ANOVA

Task	ANOVA-TM	CADE	Design-Ease	Jass	SCA	STAT-GRAPHICS
ANOVA table	×	×	×	×	√	√
Effects estimates or means	×	×	×	×	√	√
Residual analysis	×	×	×	×	√	√

(e) Paired and Unpaired Comparison of Two Treatments

Task	ANOVA-TM	CADE	Design-Ease	Jass	SCA	STAT-GRAPHICS
Paired test statistic	×	×	×	‡	‡	√
Paired conf. interval	×	×	×	‡	‡	√
Unpaired test statistic	×	×	×	‡	‡	√
Unpaired conf. interval	×	×	×	‡	‡	√

† May be analyzed as a multifactor ANOVA problem and then effects calculated from means table.

‡ May be completed with minor hand calculation.

TABLE 3. Vendor Information

Software	Vendor	Price ($US)
ANOVA-TM Version 2.20	Advanced Systems and Designs, Inc. 27200 Haggerty, Ste. B-4 Farmington Hills, MI 48331 (313) 489-8630	$795
CADE (Discovery) Version 3.1	International Qual-Tech, Ltd. 2820 Fountain Lane North Plymouth, MN 55447 (612) 887-0838	$495
Design-Ease Version 1.2	Stat-Ease, Inc. 3801 Nicollet Ave. So. Minneapolis, MN 55409 (612) 822-5574	$300
Jass Version 2.1	Joiner Associates, Inc. 3800 Regent St. P.O. Box 5445 Madison, WI 53705 (608) 238-6417	$895
SCA (QPI) Version III.3.2	Scientific Computing Associates Lincoln Centre, Ste. 106 4513 Lincoln Ave. Lisle, IL 60532 (312) 960-1698	$745
STATGRAPHICS Version 3.0	STSC, Inc. 2115 E. Jefferson St. Rockville, MD 20852 (301) 984-5000 (800) 592-0050 in U.S.	$895

Note: educational and quantity discounts, site licences, etc. may be available.