

CADAE  
COMPUTER ASSISTED DESIGN AND ANALYSIS OF RESPONSE SURFACE EXPERIMENTS

Mark J. Levine, American Cyanamid Co.

ABSTRACT

The CADAE computer program guides non-statisticians through experimental design, data entry/management and statistical analysis of a multi-objective searching scheme based upon Response Surface Methodology (RSM). The program consists of 29 screens written using the SAS/AF\* system on a MicroVAX II.

The searching strategy is to fit a response surface model for each of several performance attributes (dependent variables) over a single factor space and to utilize contour plots and grid searches to identify the factor level combinations which meet objectives on multiple performance attributes simultaneously.

Central composite and  $2^k$  designs are offered with options to block orthogonally and to choose sample size according to the size of the factor effects to be detected. Randomization is applied.

Data entry employs SAS/FSP\* to retrieve the design factor level combinations in the randomized order so that the experimenter enters only the observed performance attribute responses.

Statistical analysis includes model diagnostics, contour plots of any pair of factors at fixed levels of all other factors and multivariate grid searching to identify factor level combinations that produce performance attribute levels which simultaneously meet corresponding project objectives.

GUIDING PRINCIPLES

Like Consulting a Statistical Expert - CADAE guides users through aspects of experimental design and analysis likely to be overlooked by experimenters who have little specialized statistical knowledge. These aspects include sample size selection, blocking, randomization and avoiding extrapolation beyond the experimental region.

Flexibility and Sophistication are Sacrificed for Simplicity - Some examples are: a) Design options include only orthogonally blockable central composite designs (CCD's) and the  $2^k$  factorial designs embedded in them. b) Alpha and Beta errors for sample size evaluations are fixed at 5% and 20%, respectively. c) Only full models are

fit to the data. Models cannot be customized by adding or removing terms.

The resulting simplicity allows users to carry-out the multi-objective response surface searching scheme with only a little expert assistance.

Some Expert Assistance is Required - Model diagnostics are produced with the recommendation that they be examined by a knowledgeable data analyst.

GETTING STARTED

The instructions for using CADAE are to fill in all information requested by each screen, moving among fields within a screen by pressing "key pad 4" and "key pad 5", and then press "PF3".

The initial screen is a menu with the following offerings:

```
Command --->                PURPOSE PROGRAM
                               WELCOME TO "CADAE"
                               A SYSTEM FOR COMPUTER ASSISTED DESIGN AND ANALYSIS OF EXPERIMENTS
                               PURPOSE (CHOOSE BY TYPING "X" IN ONE OF THE SPACES BELOW. THEN PRESS
                               "PF3" TO CONTINUE):
                               _ DESIGN OF STUDY - INITIATE OR REVISE
                               _ DATA ENTRY - ENTER OR ALTER PERFORMANCE ATTRIBUTE VALUES
                               _ ANALYSE PERFORMANCE ATTRIBUTE DATA
                               _ EXIT CADAE
                               _ DESCRIPTION OF AND INSTRUCTIONS FOR CADAE
                               COPYRIGHT 1987 AMERICAN CYANAMID CO.
                               ALL RIGHTS RESERVED
```

PROJECT INITIATION/STUDY DESIGN

A project is defined by the factor levels of an experimental design over which multiple response surfaces are estimated.

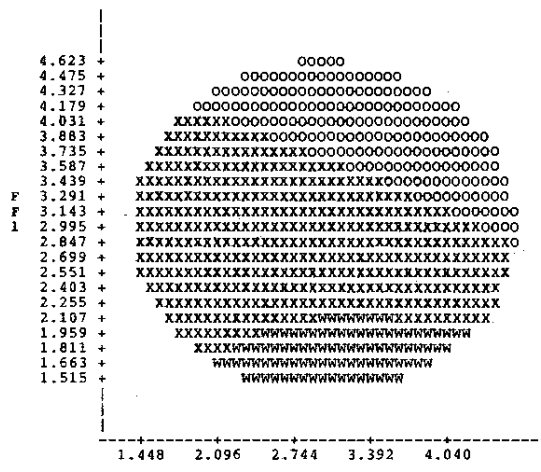
A project is initiated by selecting "Design of Study" in the initial screen. The subsequent screens request a project name, which is used to name a directory of files for the project, and the factor specifications, including the number of factors, the design type (central composite or  $2^k$ ), and the names and ranges of each factor.



The example of user input shown above produced the contour plots shown below. Note that the plotted circular areas change in size to avoid extrapolating beyond the experimental region, which is hyperspherical for CCD's. For  $2^k$  designs, the experimental region is a hypercube, therefore, contour plots are square and equal in size at different levels of the non-axis factors.

FP3-3

CONTOUR PLOT OF X1\*X2

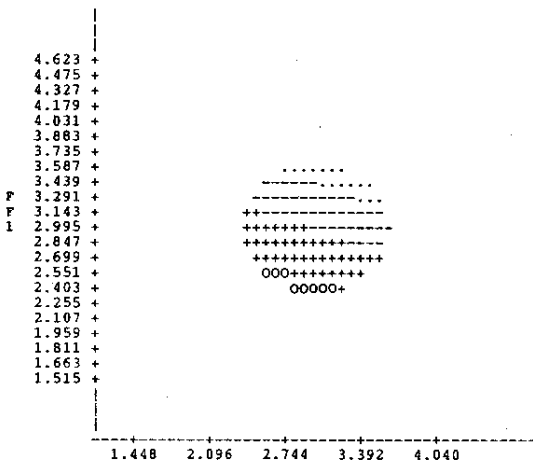


SYMBOL	RESP	SYMBOL	RESP
.....	0.8974285 - 1.4260036	O O O O O	3.5403042 - 4.5974546
-----	1.4260036 - 2.4831539	X X X X X	4.5974546 - 5.6546049
+++++	2.4831539 - 3.5403042	W W W W W	5.6546049 - 6.1831800

CONTOUR PLOT FOR PROJECT: FACT3  
 PERFORMANCE ATTRIBUTE: CLEAN  
 10:22 WEDNESDAY, FEBRUARY 3, 1988

FP3-1.5

CONTOUR PLOT OF X1\*X2

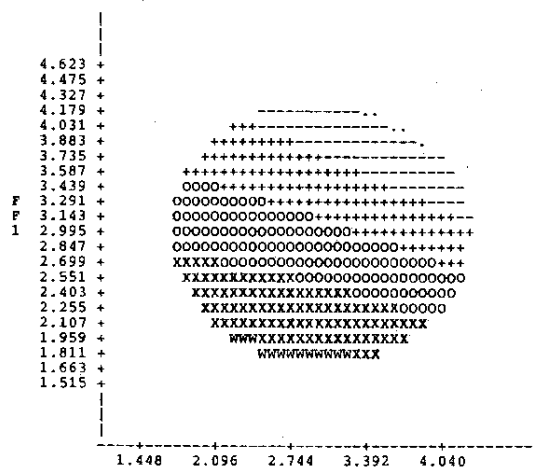


SYMBOL	RESP	SYMBOL	RESP
.....	0.8974285 - 1.4260036	O O O O O	3.5403042 - 4.5974546
-----	1.4260036 - 2.4831539	X X X X X	4.5974546 - 5.6546049
+++++	2.4831539 - 3.5403042	W W W W W	5.6546049 - 6.1831800

CONTOUR PLOT FOR PROJECT: FACT3  
 PERFORMANCE ATTRIBUTE: CLEAN  
 10:22 WEDNESDAY, FEBRUARY 3, 1988

FP3-2

CONTOUR PLOT OF X1\*X2



SYMBOL	RESP	SYMBOL	RESP
.....	0.8974285 - 1.4260036	O O O O O	3.5403042 - 4.5974546
-----	1.4260036 - 2.4831539	X X X X X	4.5974546 - 5.6546049
+++++	2.4831539 - 3.5403042	W W W W W	5.6546049 - 6.1831800

The third item in the analysis menu initiates multi-objective grid searching. The first search screen requests the names of all performance attributes to be simultaneously searched in the session and the approximate number of points to be searched, either 1000 or 5000. Completion of this screen creates a data set containing the factor levels for the selected number of points evenly distributed in the experimental region (without extrapolating beyond it, as described for contour plotting). In addition to the factor levels, each data set record contains values for all selected performance attributes predicted by their regression models at that point. This data set constitutes a "digitized" representation of the response surface system.

SACRINIT PROGRAM

Command ==>

PLEASE SPECIFY:  
 PROJECT: FACT3

AND ALL PERFORMANCE ATTRIBUTES TO BE SIMULTANEOUSLY SEARCHED IN THIS SESSION (NOTE: DATA FOR THESE VARIABLES MUST HAVE BEEN PREVIOUSLY ENTERRED):

PERFORMANCE ATTRIBUTE1: CLEAN      ← USER  
 PERFORMANCE ATTRIBUTE2: CLN2      ← INPUT  
 PERFORMANCE ATTRIBUTES3: DISNF  
 PERFORMANCE ATTRIBUTE4: \_\_\_\_\_  
 PERFORMANCE ATTRIBUTE5: \_\_\_\_\_  
 PERFORMANCE ATTRIBUTE6: \_\_\_\_\_  
 PERFORMANCE ATTRIBUTE7: \_\_\_\_\_  
 PERFORMANCE ATTRIBUTE8: \_\_\_\_\_

DENSITY OF SEARCH: X LOW (APPROX. 1000 PTS(3 MIN) FOR <6 FACTORS)  
 \_ HIGH (APPROX. 5000 PTS(15 MIN) FOR <6 FACTORS)

The desired performance level for each selected performance attribute is provided in the next screen:

```

Command ---->
SRCHOBJ.PROGRAM
PROJECT: FACT3
PROVIDE PERFORMANCE OBJECTIVES FOR EACH PERFORMANCE ATTRIBUTE LISTED BELOW.

```

PERFORMANCE ATTRIBUTE	PERFORMANCE OBJECTIVE
NUMBER	NAME > OR < VALUE
1	CLEAN > USER 5
2	CLR2 < INPUT 4
3	DISNF > INPUT 4
4	_____
5	_____
6	_____
7	_____
8	_____

Objectives of the form "4.5 < CLEAN < 5.5" can be applied to an attribute by listing it twice in the table, "CLEAN < 5.5" and "CLEAN > 4.5". Completing this screen initiates the search of the digitized response surface system for points which simultaneously saticefy all performance objectives.

The outcome of the search is summarized in the Search Review screen:

```

Command ---->
SRCHREV.PROGRAM
PROJECT: FACT3
NUMBER OF POINTS SEARCHED: 930
NUMBER OF POINTS SELECTED: 12
CHOOSE ONE BY PLACING AN X IN THE CORRESPONDING SPACE:
- REVISE PERFORMANCE OBJECTIVES AND SEARCH AGAIN
- PRINT THE SELECTED POINTS, RETURN TO PURPOSE SCREEN
- PRINT THE SELECTED POINTS, REVISE PERFORMANCE OBJECTIVES AND SEARCH AGAIN

```

The digitized response surface system may be searched repeatedly. If the first pass finds no points which meet all objectives, the user may revise the performance objectives to identify available trade-offs. If the first pass finds a very large number of points which meet all objectives, the user may wish to search again using more ambitious objectives.

POINTS MEETING SEVERAL PERFORMANCE OBJECTIVES SIMULTANEOUSLY  
PROJECT: FACT3  
10:35 WEDNESDAY, FEBRUARY 3, 1988

OBS	FF1	FF2	FF3	TYPE OF OBSERVATION	CLEAN	CLR2	DISNF
1	2.18350	1.58578	3.00000	PREDICT	5.16414	2.93891	4.61194
2	2.18350	1.85795	2.16589	PREDICT	5.48912	3.72800	4.88918
3	2.18350	1.85795	2.43806	PREDICT	5.51280	3.69435	4.97505
4	2.18350	1.85795	2.71023	PREDICT	5.42192	3.64430	4.97573
5	2.18350	1.85795	2.98239	PREDICT	5.21646	3.57787	4.89122
6	2.45567	1.73256	2.12594	PREDICT	5.12095	3.42925	4.53988
7	2.45567	1.73256	2.39810	PREDICT	5.21847	3.44003	4.67281
8	2.45567	1.73256	2.67027	PREDICT	5.20141	3.43442	4.72056
9	2.45567	1.73256	2.94244	PREDICT	5.06979	3.41243	4.68312
10	2.45567	2.00473	2.36968	PREDICT	5.14257	3.93020	4.83950
11	2.45567	2.00473	2.64184	PREDICT	5.17452	3.93556	4.91559
12	2.45567	2.00473	2.91401	PREDICT	5.09190	3.92454	4.90649

The key benefit of this "digitized" representation of the response surface system is that it provides an effective way to thouroughly search multiple response surfaces when the number of factors becomes too high for effective use of multiple contour plots.

Using grid searching of the digitized response surface system, CADAE can accommodate from 1 to 8 factors with from 1 to 8 performance attributes, thus greatly extending the complexity and thoroughness of the empirical searches which are typically used by most researchers.

\* SAS/AF and SAS/FSP are registered trademarks of SAS Institute Inc., Cary, NC, USA.

& VAX is a trademark of Digital Equipment Corporation, New York, NY, USA.

Reference:  
Myers, Raymond H., Response Surface Methodology, Virginia Polytechnic Institute and State University (1976)

Mark J. Levine  
Sr. Research Statistician  
Shulton Research Division  
American Cyanamid Co.  
697 Route 46  
Clifton, NJ 07015