DRUG DEVELOPMENT AND INDUSTRIAL PHARMACY, 6(4), 311-330 (1980)

USE OF A DESK-TOP COMPUTER IN A SEARCH-TYPE OPTIMIZATION OF TABLET FORMULATIONS

G. R. B. Down, R. A. Miller, S. K. Chopra, J. F. Millar
Pharmacy Research and Development
Merck Frosst Laboratories
Kirkland, Quebec, Canada

ABSTRACT

The use of a desk-top computer in the field of formulation and process optimization is described. To illustrate an application of this technique, a series of tablets has been prepared according to a classical optimization design to determine the optimum levels of process variables which meet pre-established specifications for product variables (hardness, thickness, etc.) based on second order polynomial predictor equations for each measured parameter.

THEORY

The use of computer-assisted optimization techniques in pharmaceutical dosage form development has been the subject of several literature reports (1-3) but apparently has not received wide-spread acceptance and usage. One of the drawbacks is that large-scale computing facilities have been required and these are not always readily accessible to development laboratory personnel. With the recent revolution in the computing equipment industry, low-cost, sophisticated devices are now available to most medium-sized laboratories. A series of programs has been written in this laboratory for use in a desk-top computer[1] to conduct the complex calculations required for formulation and process optimizations.

The purpose of the work contained in this report was to provide, via mathematical modeling of tablet properties, more complete documentation of each formulation and process under development for future reference and also to provide a rapid method of determining the location of the optimum (as defined by the various constraints imposed on the system).

The model should be so constituted that it can be used to predict accurately the performance of the "prototype". The prototype, in turn, is the full-scale physical system which is to be modeled. A second-order polynomial equation is used as the modeling equation. Sharp discontinuities in the prototype - such as capping/laminating in the case of the tablet system - are difficult to accommodate by this model and regions of experimentation where this occurs should generally be avoided.

A systematic series of experiments must first be performed in order to define the interior of a "box" bounded in each dimension by selected limits of experimentation, each dimension representing a single independent variable. The three independent variable system is the easiest to visualize (see Fig. 1). The design matrix (4) for this system is outlined in Table I.

The modeling equation in this case will be of the form:

$$Y = b_o + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_1^2 + b_5X_2^2 + b_6X_3^2 + b_7X_1X_2 + b_8X_1X_3 + b_9X_2X_3 \quad \text{(Eq. 1)}$$

where: Y = measured response (dependent variable)

$b_o, b_1, \ldots$ = estimated regression coefficients

X_1, X_2, X_3 = independent variables

An equation of this type is generated by a multiple linear regression technique for each of the dependent variables measured.

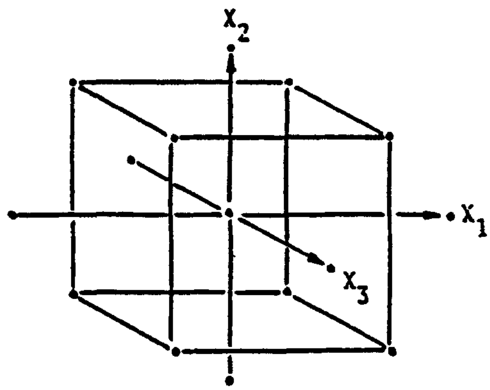

Fig. 1 - A three-factor composite design.

TABLE I

A Three-Factor Composite Design

Trial	Factor Level X_1	X_2	X_3
1	1	1	1
2	1	1	-1
3	1	-1	1
4	1	-1	-1
5	-1	1	1
6	-1	1	-1
7	-1	-1	1
8	-1	-1	-1
9	1.215	0	0
10	-1.215	0	0
11	0	1.215	0
12	0	-1.215	0
13	0	0	1.215
14	0	0	-1.215
15	0	0	0

The technique can be illustrated by the examination of a multiple regression model with just two independent variables:

$$Y_i = \beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + e_i \quad i = 1 \text{ to } n \qquad \text{(Eq. 2)}$$

where: Y_i = value of the dependent variable in the i^{th} formulation

$\beta_0, \beta_1, \beta_2$ = partial regression coefficients

X_i = value of the independent variables in the i^{th} formulation

e_i = random error associated with the i^{th} formulation

The fitted regression plane

$$\hat{Y} = b_0 + b_1 X_1 + b_2 X_2 \qquad \text{(Eq. 3)}$$

is determined by the method of least squares. The least squares function is:

$$LS = \sum_{i=1}^{n} (Y_i - [\beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2}])^2 = \sum_{i=1}^{n} e_i^2 \qquad \text{(Eq. 4)}$$

The values of β_0, β_1 and β_2 that minimize LS are the least squares estimators, b_0, b_1 and b_2, respectively. The values of b_0, b_1 and b_2 that minimize LS are solutions to the normal equations:

$$nb_0 + b_1 \Sigma X_1 + b_2 \Sigma X_2 = \Sigma Y$$
$$b_0 \Sigma X_1 + b_1 \Sigma X_1^2 + b_2 \Sigma X_1X_2 = \Sigma X_1Y \qquad \text{(Eq. 5)}$$
$$b_0 \Sigma X_2 + b_1 \Sigma X_1X_2 + b_2 \Sigma X_2^2 = \Sigma X_2Y$$

The "goodness of fit" of the predictor equations must be established by statistical means.

The R^2 value, or coefficient of determination, provides an index of the goodness-of-fit of the predictor equation and is computed using the formula:

$$R^2 = \frac{\Sigma (Y-\bar{Y})^2 - \Sigma (Y-\hat{Y})^2}{\Sigma (Y-\bar{Y})^2} \qquad \text{(Eq. 6)}$$

To say that the regression line is "good" because its R^2 is high (close to 1) can be misleading because the regression fit improves as the number of regression terms is increased or the number of terms approaches the number of formulations.

Another test for a significant regression plane is an F test:

$$F = \frac{\Sigma (Y-\bar{Y})^2 - \Sigma (Y-\hat{Y})^2 \; / \; k}{\Sigma (Y-\hat{Y})^2 \; / \; (n-k-1)} \qquad \text{(Eq. 7)}$$

where: n = number of formulations

k = number of independent variables

If $F > F_{\alpha;k;n-k-1}$ then we can reject the hypothesis that all regression coefficients are zero.

Bohidar et al. (5) have shown that all the selected independent variables can be reduced to a small number of process variables based on their contribution to the variation in each of the product variables (tablet properties) considered in the study. They have also shown (6) that, by the method of principal component analysis, one can determine the key tablet properties which are substantially contributing to the overall information about the system. The smaller the number of independent and dependent variables considered, the more efficient is the determination of the optimum.

EXPERIMENTAL

The formulation employed to illustrate the method was one containing 325 mg granular ASA[2] per tablet. Microcrystalline Cellulose[3] (MCC) and Pharmaceutical Corn Starch[4] (PCS) were added as filler and disintegrant. A lubricant blend of Talc[5] and Stearic Acid[6] (in a 1:1 ratio) was employed to provide adequate die wall and punch-face lubrication.

Fifteen formulations, of 1000 tablets each, were prepared (see TABLE II) by dry blending of all ingredients and compressed on an instrumented Manesty rotary tablet press[7] equipped with two sets of 13/32" diameter standard concave tooling. All factors other than compressional force (X_1), MCC level (X_2), and combined lubricant level (X_3) were kept constant. Each tablet also contained 5 mg of PCS.

TABLE II

Tablet Formulations Used In The Optimization Study

Formulation No.	Target Compressional Force (X_1), lbs.	MCC Level (X_2), mg/tab	Lubricant Blend Level (X_3), mg/tab
1	3960	90	15
2	3960	90	5
3	3960	30	15
4	3960	30	5
5	1320	90	15
6	1320	90	5
7	1320	30	15
8	1320	30	5
9	4244	60	10
10	1036	60	10
11	2640	96.45	10
12	2640	23.55	10
13	2640	60	16.075
14	2640	60	3.925
15	2640	60	10

Press Instrumentation

Compression force was monitored from a remote site using pairs of metal foil strain gauges[8] (in Wheatstone bridge configuration) bonded to turned-down sections on opposite sides of the pressure rod. The unbalance in the bridge circuit caused by elongation of the pressure rod during tablet compression was monitored using a carrier amplifier[9], which also served to activate the bridge. Compression events were recorded on a storage-type oscilloscope[10].

Ejection forces were monitored after the manner of Wray *et al*. (7). The existing ejection cam was replaced by two tool-steel cantilever beams. Each beam was instrumented with metal foil strain gauges[8] bonded to either side of the beam (in Wheatstone bridge configuration). Bridge excitation voltage was provided by a DC power supply[11]. Bridge unbalance voltage (caused by bending of each beam) was amplified by one of the differential amplifiers of the oscilloscope.

Both compression and ejection events were recorded directly from the oscilloscope in units of deflection and later converted into physical units.

Methods of Measurement and Equipment

Tablet hardness was determined on 10 tablets using a Schleuniger hardness tester[12].

Friability was measured on 10-tablet samples after 300 revolutions in a friabilator[13].

The thickness of 10 tablets was determined with a dial comparator.

The above attributes were determined immediately after compression.

COMPUTERIZATION OF NUMERICAL METHODS

The computer programs were written to work with only the mean of each set of measured responses. No attempt is made to assign confidence limits to predictions as these will depend not only on the precision of measurement of each response but also on the goodness-of-fit of the data to the second-order polynomial equation.

Driver Program

This program is designed to determine the fitted regression plane (Equation 1) by least squares. The program first establishes a set of normal equations. (In the case of the example, 10 such equations are established.) The equations are then solved by the method of determinants (8). Based on the predictor equation, a response is predicted for each of the 15 experiments. The R^2 and F values are then calculated. The program requires ∿ 15 seconds to assimilate and fit data for one dependent variable. A predictor equation is generated for each dependent variable, in turn, and then the coefficients for each predictor equation are stored on tape for use in the subsequent programs.

Grid Search Program

This is a "brute force" method which searches for the constrained "optimum" by dividing the specified range for each independent variable into 10 increments and solves each point thereby created ($11^3 = 1331$ points in the case of three independent variables) for the attributes in question. Comparison of a predicted value to a specified range of values for an attribute determines whether the program will consider the next attribute or pass to the next point (see Fig. 2). Careful constraining of the attributes will shorten the running time for the program as once an "unsuccessful" prediction is made for a particular attribute, further attributes are not considered for that point. The running time to consider a single attribute 1000 times is ∿ 30 seconds.

The program can be made to print out each "successful" point or simply count the number of successes until the

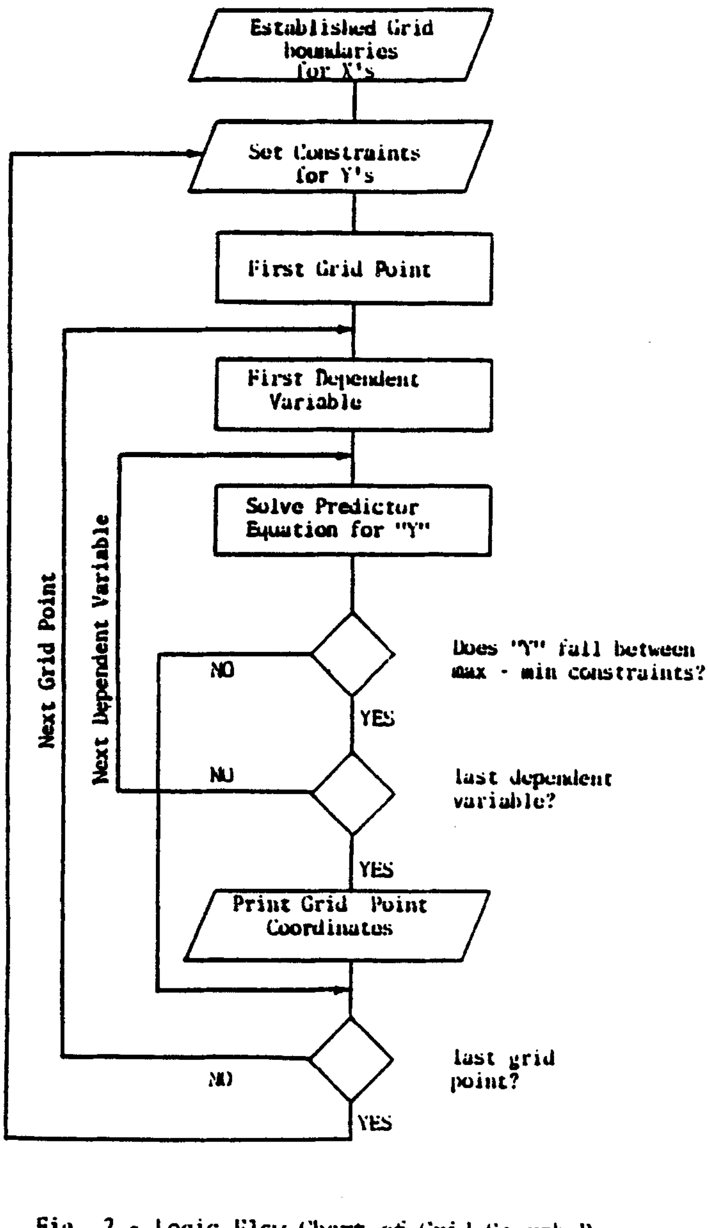

Fig. 2 - Logic Flow Chart of Grid Search Program

attributes have been sufficiently constrained so that only a few successes are obtained. At this point the operator may call for the identification of the successful point(s).

Plotting Program

The plotting program acts as an adjunct to the Grid Search Program by providing visual displays of the models in the form of contour charts, etc. For a graph of two variables, all other variables must be fixed for display purposes. In this case, equation 1 reduces to the general form[14]:

$$AX^2 + BX + C = 0 \qquad \text{(Eq. 8)}$$

One independent variable (ordinate) may be plotted against any other independent variable (abscissa) by dividing the specified range for the abscissa into very fine increments (in our case, normally 200 increments) and solving for the corresponding values on the ordinate where the solution is:

$$X = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A} \qquad \text{(Eq. 9)}$$

If the computed ordinate value falls between the defined axis limits, the x-y coordinate pair is stored and then, after all points are considered, a dot is printed on the coordinate axes corresponding to each stored point. Plots of a particular dependent variable versus a single independent variable (with all others fixed) may of course be performed by simple solution of the predictor equation for Y.

RESULTS AND DISCUSSIONS

Table III compares the actual (mean) responses to the computer-predicted values. Also illustrated are the R^2 and F values for each attribute. In each case the F value is greater than 6.22 which assures, with 99% confidence, that all the regression coefficients are not zero.

Figures 3 to 6 illustrate contour plots of several of the measured attributes. They illustrate the effect of varying

TABLE III Comparison of Actual and Predicted Responses

Exp. No.	Mean Compression Force [15]	Hardness (Y_1), kp		Ejection Force (Y_2), kg		Friability (Y_3), %		Thickness (Y_4), mm	
		Actual	Predicted	Actual	Predicted	Actual	Predicted	Actual	Predicted
1	4136 lbs	14.22	14.57	30.22	31.16	0.414	0.640	5.029	5.037
2	4004	15.24	15.60	39.38	41.05	0.259	0.354	4.947	4.943
3	4030	7.72	7.16	21.56	20.29	5.60	5.12	4.524	4.510
4	4096	6.91	6.62	26.96	26.90	5.39	4.98	4.452	4.453
5	1276	6.31	6.72	20.83	21.41	0.758	1.11	5.316	5.313
6	1342	7.81	8.47	26.72	28.47	0.589	0.991	5.197	5.210
7	1324	5.11	4.97	16.59	15.70	1.31	1.25	4.638	4.642
8	1324	5.11	4.97	19.47	19.27	1.22	1.02	4.593	4.585
9	4281	10.49	10.58	30.75	29.66	2.57	2.98	4.757	4.762
10	1060	5.50	4.92	20.07	19.26	1.44	1.07	5.007	5.003
11	2706	16.23	14.81	39.38	35.50	0.318	-0.570	5.162	5.150
12	2917	7.22	8.20	21.39	23.56	1.72	2.66	4.428	4.442
13	2732	9.95	9.95	24.69	25.39	0.963	0.926	4.788	4.792
14	2662	10.86	10.42	35.66	33.24	0.581	0.673	4.708	4.706
15	2618	9.94	10.66	26.56	29.35	0.644	0.577	4.798	4.795
	R^2 Value	97.1%		93.7%		93.1%		99.9%	
	F Value	124.5		54.9		49.8		4130.8	

[15] The mean target values are difficult to achieve exactly because of compression variations caused by tooling and die fill variations.

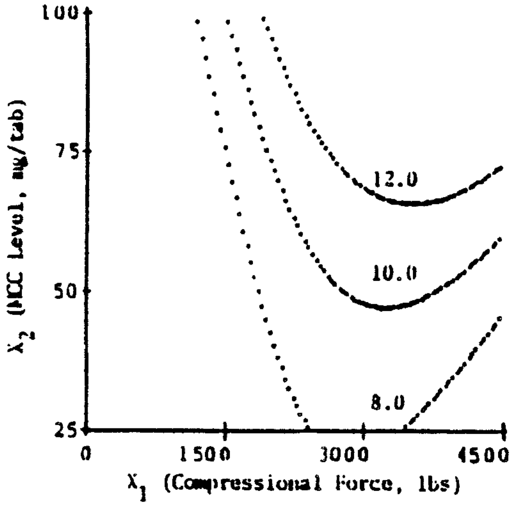

Fig. 3 - Contour plots for Hardness (Y1) in kp as a function of MCC Level and Compressional Force, X_3 (Lubricant Level) = 10 mg/tab.

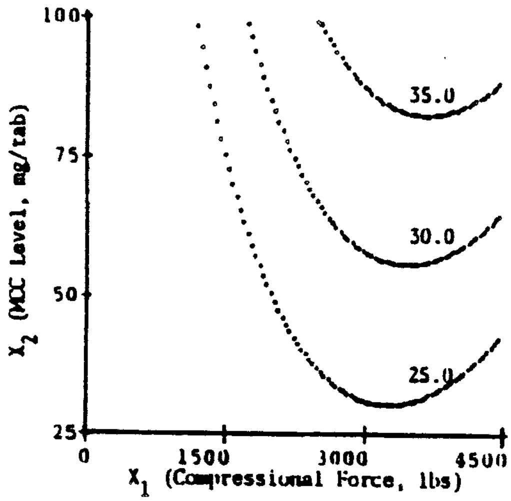

Fig. 4 - Contour plots for Ejection Force (Y2) in kg as a function of MCC Level and Compressional Force, X_3 (Lubricant Level) = 10 mg/tab.

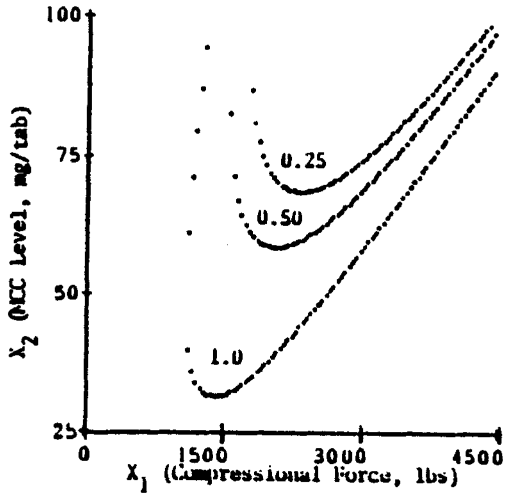

Fig. 5 - Contour plots for Friability (Y3) in % as a function of MCC Level and Compressional Force, X3 (Lubricant Level) = 10 mg/tab.

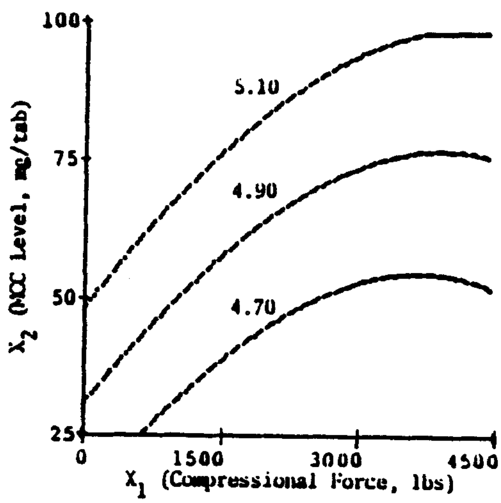

Fig. 6 - Contour plots for Thickness (Y4) in mm as a function of MCC Level and Compressional Force, X3 (Lubricant Level) = 10 mg/tab

compressional force on a tablet system. As the force increases to the point where capping/laminating occurs, hardness and ejection force decrease (i.e., in the case of the example, the system requires more MCC to maintain the same hardness and ejection force) and friability and thickness increase. The apparent thickness increases because of the voids created in the compact of lamination. The Lubricant Level had little effect on the measured attributes except ejection force (see Fig. 7 and 8).

The following constraints were then entered into the grid search program:

hardness	> 9 kp
ejection force	< 30 kg
friability	< 0.9%
thickness	< 4.75 mm

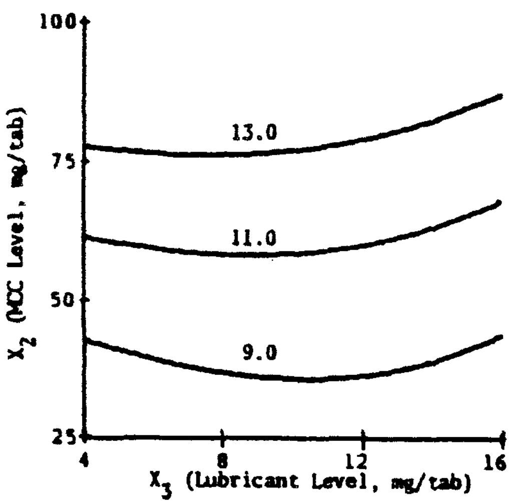

Fig. 7 - Contour plot for Hardness (Y1) in kp as a function of MCC Level and Lubricant Level. X1 (Compressional Force) = 3000 lbs.

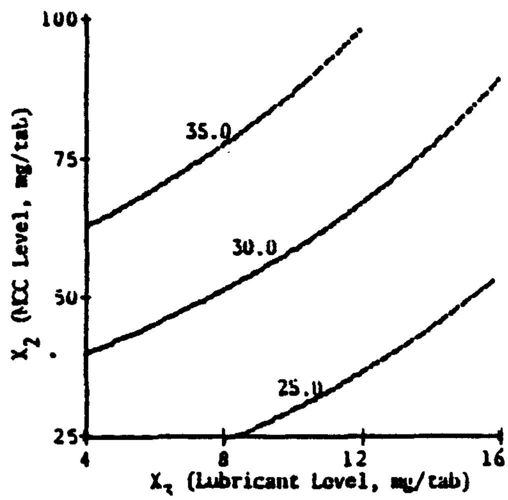

Fig. 8 - Contour plot for Ejection Force (Y2) in kg as a function of MCC Level and Lubricant Level. X1 (Compressional Force) = 3000 lbs.

Using these constraints, the program was able to generate seven sets of conditions for compressional force, MCC concentration, and lubricant level which would give rise to tablets with the specified properties. To test the validity of these predictions, one of the sets of conditions, namely:

X_2 = 50 mg per tablet
X_3 = 10 mg per table

was chosen at random and tablets prepared over a compressional force range from 876 lb to 4057 lb. The actual and predicted results are compared in Table IV and Figures 9 to 12. The overall correlation of actual and predicted values for hardness, ejection force and thickness is good but is somewhat weaker for friability at high compressional force.

As can be seen, the constrained optimum (in both the actual and predicted case) lies somewhere between 2204 and 2811 lbs. compression force (the predicted value from the grid search program being X_1 = 2319 lbs).

TABLE IV

Comparison Of Actual And Predicted Responses

(X_2 = 50 MG/TAB, X_3 = 10 MG/TAB)

Mean Compression Force (lbs)	Hardness (Y_1), kp		Ejection Force (Y_2), kg		Friability (Y_3), %		Thickness (Y_4), mm	
	Actual	Predicted	Actual	Predicted	Actual	Predicted	Actual	Predicted
876	2.94	3.62	14.87	16.30	--	1.22	4.972	4.928
1399	5.89	6.22	19.58	20.96	1.10	0.795	4.786	4.837
2204	7.72	8.99	24.06	26.04	0.724	0.725	4.657	4.735
2811	9.61	10.08	25.49	28.21	1.24	1.14	4.664	4.687
3463	9.25	10.30	27.09	28.92	1.13	2.04	4.636	4.665
4057	9.94	9.65	28.30	28.13	1.41	3.25	4.609	4.671
Correlation Coefficient (r) [16]	0.98		0.98		0.73		0.97	
Regression Equation [16]	$Y = 0.879 + 0.961\ X$		$Y = 2.069 + 0.977\ X$		$Y = -1.873 + 3.089\ X$		$Y = 1.203 + 0.752\ X$	

[16] obtained by linear regression of predicted values on actual values (see Fig. 9-12)

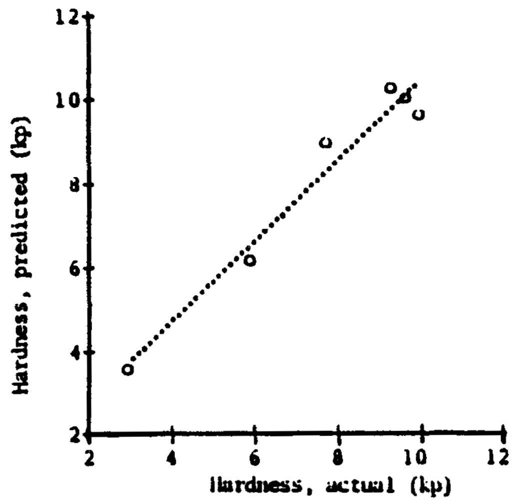

Fig. 9 - Regression plot of predicted hardness values on actual values (see Table IV).

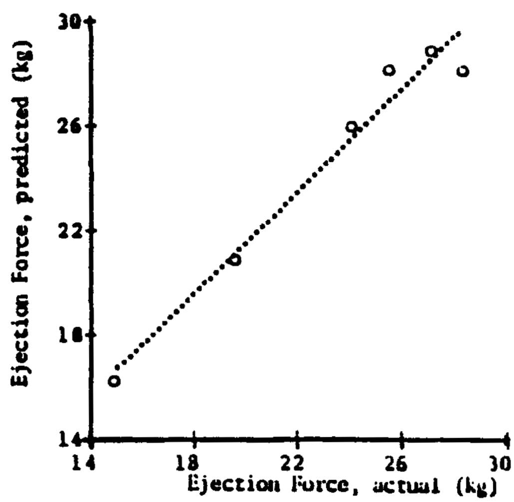

Fig. 10 - Regression plot of predicted ejection force values on actual values (see Table IV).

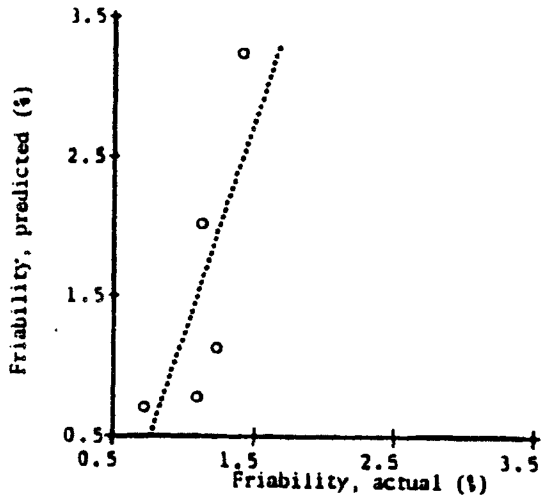

Fig. 11 - Regression plot of predicted friability values on actual values (see Table IV)

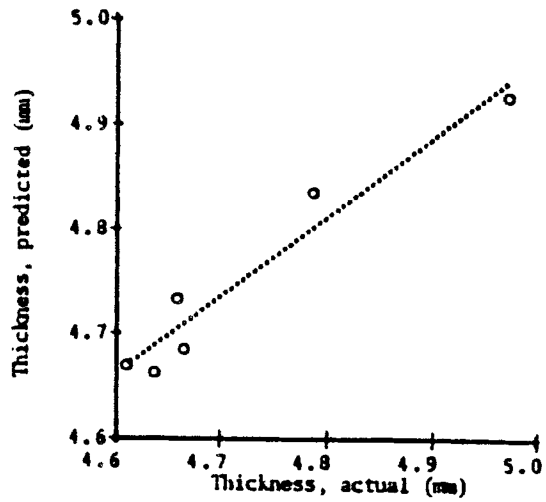

Fig. 12 - Regression plot of predicted thickness values on actual values (see Table IV).

SUMMARY AND CONCLUSIONS

A facile method has been developed for product and process optimization using a desk-top computer. The method offers the advantages of rapid access and versatility.

An example has been illustrated of a three-factor tablet formulation/process optimization. The technique, of course, has wide applicability in the development laboratory and has also been used in the optimization of HPLC assay procedures, etc.

The confidence limits of the predictions and other statistical aspects of the optimization method described will be explored in the next part of this treatise. The method will also be expanded to consider non-search type methods involving techniques such as canonical reduction, principal component analysis and selective regression analysis. These will provide us with some degree of precision of predictions. Further work will be required in implementing these techniques.

ACKNOWLEDGMENTS

The authors wish to thank Dr. N. R. Bohidar for his statistical advice and for reviewing the manuscript.

FOOTNOTES

[1] HP-9825A computer with 9871A printer, Hewlett Packard, Palo Alto, California

[2] Monsanto Ltd., Ruabon, U.K.

[3] FMC Corporation, Philadelphia, PA

[4] St. Lawrence Chemical, Montreal, Canada

[5] Bate Chemical, Montreal, Canada

[6] St. Lawrence Chemical, Montreal, Canada

[7] Manesty B3B, Manesty Machines Ltd., Liverpool, England

[8] FAET 25-B-35-S6-E half-bridge T-rosettes, type SR-4, BLH Electronics, Waltham, Mass.

[9] Sanborn Model 311A, Hewlett Packard Co., Palo Alto, Calif.

[10] Model 5113 with 5B12N time base and two 5A26 dual differential amplifiers, Tektronix, Inc., Beaverton, Oregon

[11] Model 21-200, Calex Mf. Co., Inc., Pleasant Hill, Calif.

[12] Vector Corp., Marion, Iowa

[13] Roche

[14] e.g., for a plot of X_1 (abscissa) vs. X_2 (ordinate), Equation 1 reduces to a form where:

$$C = b_0 + b_1X_1 + b_3X_3 + b_4X_1^2 + b_6X_3^2 + b_8X_1X_3 - Y$$

$$B = b_2 + b_7X_1 + b_9X_3$$

$$A = b_5$$

REFERENCES

1. D. E. Fonner, Jr., J. R. Buck, and G. S. Banker, J. Pharm. Sci., 59, 1587 (1970).
2. J. B. Schwartz, J. F. Flamholz, and R. H. Press, J. Pharm. Sci., 62, 1165 (1973).
3. ibid, 1518 (1973).
4. O. L. Davies, ed., "The Design and Analysis of Industrial Experiments", 2nd edition, Longman, London, 1979, p. 495.
5. N. R. Bohidar, F. A. Restaino, and J. B. Schwartz, Drug Develop. Ind. Pharm., 5, 175 (1979).
6. ibid, J. Pharm. Sci., 64, 966 (1975).
7. J. G. Vincent and P. E. Wray, U.S. Patent 3,388,424.
8. "Simultaneous Equations (Maximum Accuracy)" in "HP9825A General Utility Routines", Hewlett-Packard, Palo Alto, Calif., p. 63.