CHAPTER 12

Multiple Responses

Only a single response variable was assumed in the previous chapters. In practice, however, there are usually multiple responses, such as multiple product measurements that are of interest and that ideally should be maintained at a target value while, say, maximizing product yield. There are many different types of scenarios that occur in practice. For example, if there are three response variables, the objective might be to maximize two of them and minimize the other one. In this chapter we will look at several examples with various objectives.

When multiple responses are analyzed, a model is first fit for each response. Then some give-and-take must occur in the optimization since a combination of factor levels that would optimize one response almost certainly would not optimize the other responses. Assume for the sake of illustration that there are two response variables, which are of equal importance and each is to be maximized, and each variable is a function of the same two factors. One way to picture the necessary "give-and-take" would be to construct overlaid contour plots. Almost certainly, the indicated maximum on each contour plot would not point to the same combination of levels of the two factors, so some compromise would be necessary. The user could make this compromise based on the visual information. Most practical problems will have either more than two response variables or more than two factors, however, so such an approach has limited usefulness. Nevertheless, this approach is illustrated in Section 12.1.

Certain simplifying but unrealistic assumptions were made in some of the early research on multiple response optimization, such as assuming that all response variables are functions of the same independent variables and, even worse, that the model for each response variable has the same general form (see, e.g., Khuri and Conlon, 1981). While such assumptions simplify the analyses considerably, simplified analyses based on unrealistic assumptions are of little value.

Even if such assumptions were tenable, there is still the problem that response variables are often correlated. When this is the case, it won't be possible to try to change factor levels so as to move in the direction of a desirable value for one response

Modern Experimental Design By Thomas P. Ryan
Copyright © 2007 John Wiley & Sons, Inc.

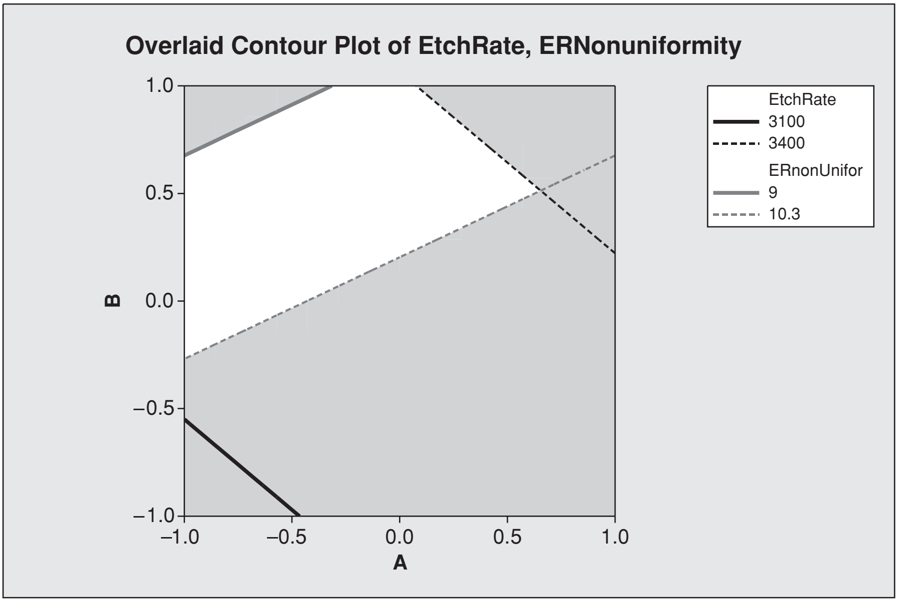


Figure 12.1 Overlaid contour plot for two responses.

variable without moving the expected value of another response variable that may already be very close to its most desirable value.

In the next section we look at a numerical optimization approach for finding the optimal solution, although there is no guarantee that the global optimum has been achieved as there might be multiple local optima. It is also desirable to look at graphs to see how sensitive the optimum solution in terms of the fitted values is to changes in the predictors.

Before multiple response optimization was implemented in statistical software, it was common to try to picture optimum and near-optimum solutions by using overlaid contour plots. Figure 10.12 is an example of a contour plot for a single response variable. If there were two responses, the contour plot for the second response could be overlaid on the contour plot shown in Figure 10.12. Such an approach will work fine when there are two response variables and two predictors, but dimensionality becomes a problem when there are more than two predictors.

12.1 OVERLAYING CONTOUR PLOTS

Figure 12.1 is an overlaid contour plot with two of the response variables in a dataset that is analyzed in Section 12.5. Here, for the sake of illustration, we assume that there are only two responses and two factors. A model for each response must be fit, and we fit a model with only main effects for each response.

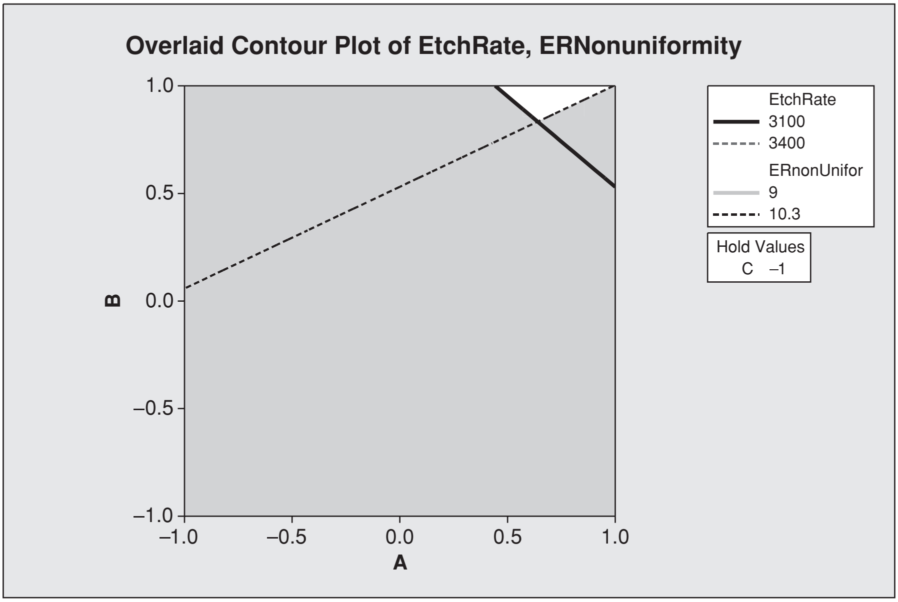


Figure 12.2 Overlaid contour plot with factor *C* held constant.

The plot was constructed by specifying an acceptable range for Etch Rate, which is to be maximized, of 3100 to 3400, and an acceptable range of Etch Rate Nonuniformity, which of course would be minimized, of 9 to 10.3. The region in white is the feasible region in terms of the two factors, A and B. The experimenter would be quite pleased if the Etch Rate were 3400 or above, but below 3100 would be totally unacceptable, with the range between them being "just okay." Similarly, Etch Rate Nonuniformity of 9 or below would be most desirable and above 10.3 totally unacceptable. Notice that the combination of 9 and 3400 cannot be achieved, at least not within the range of factor levels used in the experiment. It might be achieved with a higher level of *B* than was used in the experiment, but such a level might not be feasible. If it were not feasible, then some compromise, as mentioned at the beginning of the chapter, would be necessary.

Figure 12.2 is a contour plot of the data on the same two response variables, but this time the model for each response variable consists of the main effects of all three factors that were in the dataset. Since there are three factors, one of them must be set at a fixed level, and in Figure 12.2 factor *C* was set at the low level.

Notice that the feasible region has shrunk considerably and there is no contour for the most desirable value of Etch Rate Nonuniformity. If factor *C* had been fixed at the high level, a different, and undesirable, contour plot would have resulted, as the reader is asked to show in Exercise 12.8. Thus, contour plots start to lose some value when there are more than two factors.

12.2 SEEKING MULTIPLE RESPONSE OPTIMIZATION WITH DESIRABILITY FUNCTIONS

There are various methods for achieving optimization with multiple responses, with the best known method based on *desirability functions*, as explained in, for example, Derringer and Suich (1980), who modified the desirability function approach given by Harrington (1965), which is discussed in Section 12.6 (see also Derringer, 1994). The general idea is to compute a desirability value, d_i, for each response that is a measure of how close the fitted value with the optimal settings of the factors is to the desired value, and use these values to form the *composite desirability* for k response variables given by

$$D = (d_1 * d_2 * \cdots * d_k)^{1/k} \tag{12.1}$$

Derringer and Suich (1980) sought to maximize D by using a pattern research method similar to the one given by Hooke and Jeeves (1961). Unfortunately, that method is somewhat unreliable. Since that is a “hill-climbing” technique, it is susceptible to producing local solutions rather than global solutions. For this reason, it has been recommended that the procedure be used twice, using the supposed optimum solution as the starting point when the procedure is used the second time. A better approach, in general, when there is a fear of being derailed by local optima is to use many random starting points. Of course this option is available to users only when the algorithm is available for direct use, not when it is part of a statistical software package that controls how the algorithm is used. Algorithms are discussed further in Section 12.7; see http://www.mit.jyu.fi/palvelut/sovellusprojektit/ovi/methods.html#hookejeeves for a comparison of different types of algorithms.

For example, assume that there is only one response variable and the objective is to minimize impurity, with any value of at most 5 considered to be equally desirable, with 5.5 declared to be the largest acceptable value. A fitted value of at most 5 would then have a desirability of 1 because the goal was met, whereas a fitted value of 5.25 would have a desirability, d, of

$$\begin{aligned} d &= \frac{\text{Upper} - \widehat{y}}{\text{Upper} - \text{Lower}} \\ &= \frac{5.5 - 5.25}{5.50 - 5.00} \\ &= 0.50 \end{aligned} \tag{12.2}$$

Clearly the desirability value must be between 0 and 1, inclusive, and this is also true for the composite desirability, which will be 1 only if all the individual desirability values are 1, and will be 0 if at least one individual desirability is 0.

The desirability for a response to be maximized would be similarly computed. For example, if a fitted value of at least 15 is desired with 13.5 being the smallest acceptable value, a fitted value of 14.3 would have a d-value of $d = (\widehat{y} - \text{lower})/(\text{upper} - \text{lower}) = (14.3 - 13.5)/(15 - 13.5) = 0.533$. (Derringer and Suich (1980) presented only the maximization case, pointing out that the minimization of $\widehat{Y}$ is equivalent to the maximization of $-\widehat{Y}$. Loosely speaking, if we replace $\widehat{y}$ by $-\widehat{y}$ in the expression for d in the maximization case, then of course "lower" would be replaced by "upper" in the numerator of the expression because of the change of sign.)

When there is a target value for the response (termed the "most desirable" value by Derringer and Suich, 1980), the d-value is computed as

$$\begin{aligned} d &= \frac{\widehat{y} - \text{Lower}}{\text{Target} - \text{Lower}} \qquad \text{Lower} \le \widehat{y} \le \text{Target} \\ &= \frac{\widehat{y} - \text{Upper}}{\text{Target} - \text{Upper}} \qquad \text{Target} \le \widehat{y} \le \text{Upper} \end{aligned} \tag{12.3}$$

Thus, in each case $d = 1$ if $\widehat{y} = \text{Target}$.

Although the Derringer and Suich (1980) desirability function approach is undoubtedly the best known approach and is almost certainly the most frequently used approach since it is implemented in the leading statistical software packages, it is not the only transformation approach that has been proposed. Various other methods, including the global criterion method, are described and compared by Tabucanon (1988), which unfortunately is out of print. The global criterion method, which unfortunately does not work for target values, is described in detail in Kros and Mastrangelo (2004).

12.2.1 Weight and Importance

The presentation of desirability to this point has focused on a single response variable. Assume that there are four response variables. The importance of hitting a target value or achieving a maximum or minimum value may vary considerably over the different response variables. If so, different values may be used for Importance, and different values may also be used for Weight.

We will consider Weight first, which determines the shape of the desirability function for each response. Assume that response i is to be maximized and that a value of $\widehat{Y_i}$ more than slightly less than an acceptable maximum value quickly becomes almost unacceptable with movement away from the maximum. More specifically, the desirability is not linear for values less than the acceptable maximum. For such a scenario, d_i^r should be used, with $r > 1$. Since d_i is of course less than 1.0 when the maximum is not achieved, a value of r much greater than 1 could drive d_i^r close to zero, depending on the value of d_i. Similarly, if a response is to be minimized and thus a $\widehat{Y_i}$ value very close to the acceptable minimum is essential, r much greater than 1 would drive d_i^r close to zero. For a target value, two constants could be used for the two equations in Eqs. (12.3), which might reflect that missing the target on the high side is worse than missing it on the low side.

The set of desirabilities with each raised to the appropriate power would then be used in place of the d_i in Eq. (12.1).

The user who does not specify weights for the individual desirability functions is implicitly assuming that each function is a linearly increasing function within the range of specified lower and upper response values if the response is to be maximized, and a linearly decreasing function between the particular specified values if the response is to be minimized. If there is a target value, then linearity is assumed on each side of the target value. The shape of each desirability function is automatically shown when Design-Expert is used, for example.

Whereas Weight is used with each individual response, Importance should be viewed relative to all the responses. For example, if there are four responses ($R1$, $R2$, $R3$, and $R4$), with $R1$ and $R2$ to be maximized and $R3$ and $R4$ to be minimized, the maximization of $R1$ may be more important than the maximization of $R2$, and the minimization of $R3$ may be more important than the minimization of $R4$. It would then be appropriate to assign a greater "Importance" value to $R1$ than $R2$, and $R3$ should be assigned a greater importance value than $R4$. In MINITAB, the Importance values can range from 0.1 to 10, with 1.0 being the default value. In Design-Expert, the possible Importance values range from "+" to "+ + + + +," with "+ + +" being the default.

12.3 DUAL RESPONSE OPTIMIZATION

One of the most visible applications of multiple response optimization during the past 20 years has been *dual response optimization*. The word "dual" in this context does not mean that there are two responses; rather there are two functions of the same response. The first function is that presented in Section 12.1: to maximize or minimize the response, or hit a target value. The second function involves the variance of that response. Specifically, the three dual problems given by Taguchi (see, e.g., Taguchi, 1986), which are approached by modeling the mean and variance separately, were (1) minimize the variance while hitting a target value for the response, (2) maximize the response while keeping the variance constant, and (3) minimize the response while keeping the variance constant.

Significant work in this area includes Vining and Myers (1990), Lin and Tu (1995), who pointed out some deficiencies in the method given by Vining and Myers (1990); Copeland and Nelson (1996), who proposed solving dual response problems with the use of direct function minimization; Tang and Xu (2002), who proposed a goal programming approach; Koksoy and Doganaksoy (2003), who let the mean and variance be on equal footing and used standard nonlinear multiobjective programming techniques; del Castillo and Montgomery (1993), who used a nonlinear programming approach; and Ding, Lin, and Wei (2004), who used a weighted mean squared error approach.

12.4 DESIGNS USED WITH MULTIPLE RESPONSES

It should be noted that multiple responses cause complexities primarily in analysis rather than in design. Multiple responses do force some design considerations, however. In particular, if we relax the assumption that the same model can be used for each response (and later this will be illustrated), we then have to think about the factors that might be related to *at least one* of the response variables. Of course in the extreme case where each response is a function of different variables, multiple response optimization would then reduce to a set of single response optimizations. Such a scenario would undoubtedly be extremely rare, however.

Essentially any design could be used with multiple response optimization, just as any design could be used when there is a single response. The literature contains some articles on specific designs used for which there are multiple responses, such as the analysis of split-plot designs with multiple responses, as discussed by Ellekjaer, Fuller, and Ladstein (1997–1998). Chen, Hedayat, and Suen (1998) presented optimal designs for experiments with multiple responses.

It is important that enough design points be used so that parameter estimates are obtained with good precision. If not, the variance of a predicted response could be large, with the consequence that predicted response values may be obtained with somewhat low precision, which would mean that an optimum solution may also have low precision. We should keep in mind that optimal solutions are functions of data and specifically functions of parameter estimates, rather than obtained from theoretical models. Therefore, “optimal solutions” have a variance, just as do the statistics that produce them.

This suggests that a parsimonious model be fit for each response. Another consideration, however, is what to do when an analysis suggests a nonhierarchical model. Those who advocate hierarchical models would state that nonsignificant effects be included to make the model hierarchical. A nonhierarchical model may have been caused by large interactions. The principle of parsimony should be weighed against the consequences in multiple response optimization if a “full model” of some sort is not used.

12.5 APPLICATIONS

Example 12.1

We consider a dataset given and analyzed extensively by Czitrom, Sniegowski, and Haugh (1998) that was analyzed briefly in Ryan (2000), but which will be analyzed in more detail here, using an approach that is different from that used by Czitrom et al. (1998).

An experiment was performed involving integrated circuits to determine the effects of three manufacturing factors (bulk gas flow (A), CF_4 flow (B), and power (C)) on three response variables: selectivity, etch rate, and etch rate nonuniformity. The objective was to determine factor settings that would maximize selectivity, which is the ratio of the rate at which oxide is etched to the rate at which polysilicon is etched, maximize etch rate (which would also maximize manufacturing throughput), and of course minimize etch rate nonuniformity.

The design that was used was a 2^3 with two centerpoints. We will analyze these data with an eye especially toward determining whether or not the same model form seems appropriate for each of the three response variables. In particular, we will compare R^2 values and also assess the absolute magnitude of those values.

The data that are given in Table 12.1 are in the general (uncoded) form in which the data were given in Czitrom et al. (1998).

Of course analyses cannot be performed with raw-form data, however, as was illustrated in Section 4.1.

TABLE 12.1 Data from a Silicon Wafer Experiment

Run	Bulk Gas Flow (sccm)	CF_4 Flow (sccm)	Power (watts)	Selectivity	Etch Rate (Å/min)	Etch Rate Nonuniformity
1	60	5	550	10.93	2710	11.7
2	180	5	550	19.61	2903	13.0
3	60	15	550	7.17	3021	9.0
4	180	15	550	12.46	3029	10.3
5	60	5	700	10.19	3233	10.8
6	180	5	700	17.5	3679	12.2
7	60	15	700	6.94	3638	8.1
8	180	15	700	11.77	3814	9.3
9	120	10	625	11.61	3378	10.3
10	120	10	625	11.17	3295	11.1

For the response variable Selectivity, the main effects of the first two factors are the only significant effects (at $\alpha = .05$) when a model with all estimable effects is fit. The R^2 value is .9445. This alone is not the best way to select a model, however. Subsequent analysis shows that the AB interaction has a p-value of .037 when used in the model with A and B, so we might use the model with A, B, and AB as the three terms. This gives an R^2 value of .9747. The AB effect estimate is -1.467. Since this is not large relative to the A and B main effect estimates, which are 6.527 and -4.927, respectively, those main effect estimates are not degraded. Since the interaction is somewhat borderline significant and the p-value is well above .01, we will not use the interaction term in the model, also for reasons of simplicity.

For the Etch Rate response variable, the main effects of all three factors are the only significant effects and the R^2 value is .9451.

For the Etch Rate Nonuniformity response variable, the main effects of the three factors are again the significant effects and $R^2 = .9819$.

The fact that the last two response variables would be fit with the same model is not surprising since the two variables are closely related.

It is obvious that all three models fit the data very well, but if we include the main effect of the third factor in the model for the first response so as to have the same model for all three factors, we are placing a term in the model for the first response variable that has a p-value of .236. This is not all that bad, however, because it means that the optimum value of that variable will be relatively insensitive to the value of the third factor in the process of using multiresponse optimization to determine the optimum levels of the three factors for all three response variables, as factors that are not significant have comparatively small coefficients. (Here the coefficient of the third factor in the model for the first response is -1.32, compared to 9.13 and -6.95 for the coefficients of the first and second factors, respectively.) We will see that the (estimated) optimum level of the third factor differs very little over the analyses that follow.

The original settings of the factors were, in uncoded units, $A = 90$, $B = 5$, and $C = 625$, which correspond to -0.5, -1, and 0, respectively, in coded units. These would produce fitted values of Selectivity = 12.43, Etch Rate = 3068, and Etch Rate Nonuniformity = 11.6 from the model that Czitrom et al. (1998) used for each response. Those models are different for Selectivity and Etch Rate from the models selected in this example, however. Specifically, their model for Etch Rate included the AB and AC interactions in addition to the three main effect terms, and their model for Selectivity included the three main effect terms plus the AB interaction.

This led, through their use of contour plots, to the selection of $A = 180$, $B = 15$, and $C = 550$ in the original units, which correspond to 1, 1, and -1, respectively, in the coded units. These settings produce fitted values of Selectivity = 12.45, Etch Rate = 3048 (less than the value for the original settings), and Etch Rate Nonuniformity = 10.3. Note that their optimum solution was one of the 10 design points in the experiment.

Ryan (2000) did not use contour plots but rather used statistical software and arrived at factor settings of $A = 169.07$, $B = 14.09$, and $C = 561.04$. This solution was based on the assumption that maximizing the first two response variables and minimizing the third variable were of equal importance. The fitted values using these settings are Selectivity = 12.48, Etch Rate = 3091, and Etch Rate Nonuniformity = 10.37. This combination of fitted values was superior to that given by Czitrom et al. (1998), so the factor settings would be preferable.

The "optimal" solution to a given problem might be expected to improve over time simply because algorithms and desirability functions have improved over time. Multiple response optimization through the use of desirability functions does have various pitfalls however, as discussed at the beginning of this chapter, and the best solution is to this day not easily obtained. In particular, some well-known statistical software have failed to find the optimal solution for certain problems. For this reason, it is prudent to use different software and compare the results. If the solution obtained with one software package has a higher composite desirability than the solution obtained with another software package, then the first solution is obviously superior, although it might not be the optimal solution. If a user has any doubts about the solution obtained with particular statistical software being the optimal solution, it would be best to obtain, if possible, information about the algorithm that is used. It would also be a good idea to start with the supposed optimal solution and make small changes in the factor levels and see how this affects the composite desirability, provided that this can be performed dynamically, as can be done with JMP, for example.

Given in Figure 12.3 is the JMP output for this example, after specifying minimum, middle, and maximum values of 12.45, 13.725, and 15, respectively, for Selectivity; 3068, 3184, and 3300 for Etch Rate; and 10, 10.15, and 10.30 for Etch Rate Nonuniformity.

(*Note*: Using different models for the responses in JMP creates some problems, which require special treatment, which is why the first row label is of a different form than the other row labels.)

Before discussing these results, it should be noted that running the optimization in JMP using this solution as the starting values will produce a slightly different solution. Therefore, we should always think of any multiple optimization "solution" as just an

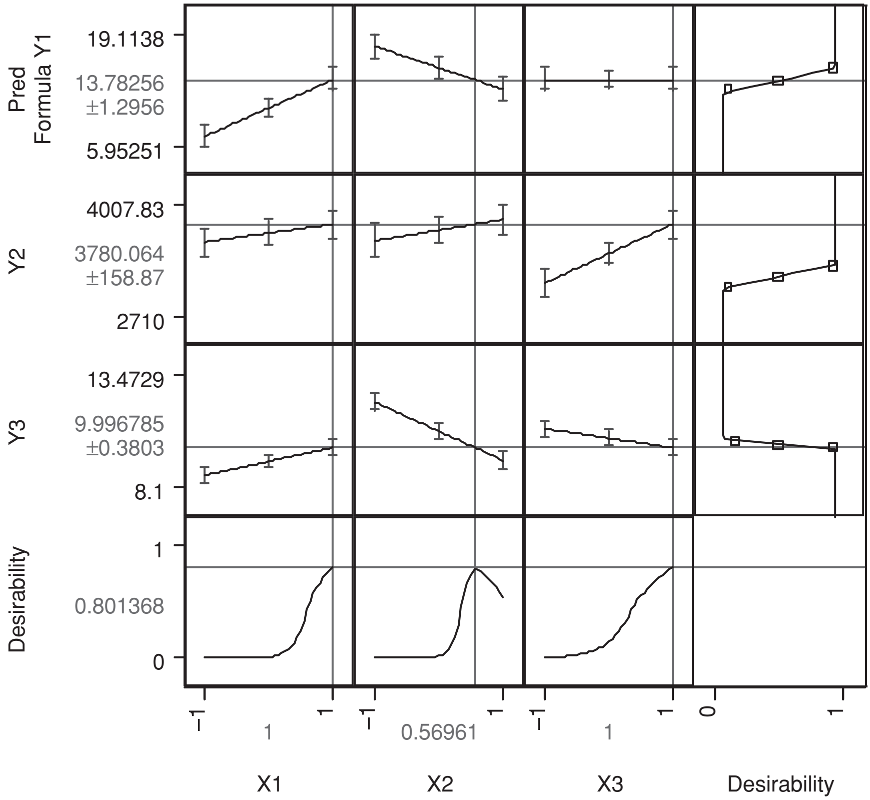


Figure 12.3 JMP desirability output for silicon wafer experiment.

estimate of the optimum solution, and it might be a poor estimate, depending on the nature of the problem and the type of algorithm that the selected software uses.

The individual graph panels in the first three columns of panels show the sensitivity of the composite desirability and the fitted values of each response to changes in each of the three factors. (The top panel in the third column has a horizontal line across the different levels of factor C because that factor was not used in the model for the first response variable.) We can see that the composite desirability is highly sensitive to changes in the second factor, as is indicated by the desirability graph for that factor in the last row. This can also be seen dynamically in JMP by positioning one's cursor at the vertical line, which corresponds to the value of the second factor, and sliding the level one way or the other. This suggests that algorithms that differ slightly could give slightly different solutions for the second factor, and we will see this later for this example.

The panels also suggest that higher levels of the first and third factors may be desirable and the fact that the desirability increases sharply past the midpoint of the range of each factor explains why the optimal solution has each factor set at the highest level used in the experiment.

The 95% confidence intervals on the mean response for each response variable are given at the beginning of each row of panels. The vertical bars in the first three rows and first three columns are 95% prediction intervals for each of the three response variables, for the low, middle, and high values of each of the factors. The coded levels of the factors would of course have to be converted to the actual units.

It is of interest to compare the JMP results with the results obtained using Design-Expert, which are given below.

Constraints

Name	Goal	Lower Limit	Upper Limit	Lower Weight	Upper Weight	Importance
A	is in range		-1	1	1	3
B	is in range		-1	1	1	3
C	is in range		-1	1	1	3
R1	maximize	12.45	15	1	1	3
R2	maximize	3068	3300	1	1	3
R3	minimize	10	10.3	1	1	3

Solutions

Number	A	B	C	R1	R2	R3	Desirability
1	1.00	0.55	1.00	13.9791	3760.47	10.0001	0.843
2	1.00	0.55	0.99	13.9712	3757.67	10	0.842
3	1.00	0.55	0.98	13.96	3753.6	10	0.840
4	0.97	0.53	1.00	13.9247	3756.32	10	0.833

We observe that the solutions produced by Design-Expert and JMP are almost identical, differing only slightly on the level of the second factor. The composite desirabilities differ more than slightly, however, and this is because a different desirability function form is used by each software, as will be explained shortly.

This can be seen as follows. For the Design-Expert solution, the individual desirabilities are computed as follows. For Selectivity, which was to be maximized,

$$d = \frac{\widehat{y} - \text{Lower}}{\text{Upper} - \text{Lower}}$$

$$d = \frac{13.9791 - 12.45}{15 - 12.45}$$

$$= 0.599647$$

The desirability is 1.0 for Etch Rate since the fitted value of 3760.47 is at least equal to and is actually much greater than the value, 3300, that is considered to be completely acceptable. For Etch Rate Nonuniformity, which was to be minimized, the desirability is virtually 1.0 since the fitted value is practically equal to the largest value that is considered to be completely acceptable. Therefore, the composite desirability is simply $(0.599647 * 1 * 1)^{1/3} = 0.843$, which of course was the value given in the Design-Expert output.

We can force the same solution using JMP by fixing the value of the second factor at 0.55. Doing so gives a composite desirability of 0.787, which again we cannot compare with the Design-Expert composite desirability since different functions are used. The fitted values also differed noticeably, part of which was probably due to rounding the level of the second factor.

Regarding the difference in the composite desirabilities, Design-Expert uses the desirability function form used by Derringer and Suich (1980), as we have just verified, whereas JMP does not. The Derringer and Suich (1980) approach does not give smooth functions, which limits the possible shapes of the desirability functions. In contrast, JMP defines the desirability functions based on the lower, middle, and upper values specified by the user, resulting in piecewise smooth functions. For example, moving the middle values for the responses to be maximized close to their maximum values and moving the middle value for the response very close to the minimum value cause a considerable drop in the composite desirability with very small changes in the factor levels.

Regardless of which solution we go by, each of the discussed solutions *appears* to be far superior to the solutions given by Czitrom et al. (1998), and is also superior to the solution given by Ryan (2000), which was obtained using different software from what has been used here. There is one constraint that should be used, however, and we will see how that affects the solution. The solution given by Czitrom et al. (1998) had the second factor set at its highest level, which was 15 in the original units. The manufacturing engineer wanted to increase the CF_4 Flow from 5 to 15 because the manufacturing equipment that was to be used for the etching process is under better control when the CF_4 Flow is at least 15. The value of CF_4 Flow in the solution given by Ryan (2000) was 14.09, as stated previously, which may or may not have been good enough for the engineer.

We know from Figure 12.3 that increasing the value of the second factor is going to cause a sharp decrease in the desirability. A fixed factor level cannot be specified using JMP, other than locking in the "current" value in an analysis, as changes are made by sliding a vertical line in the graph using a mouse, although of course this is an inexact method. Accordingly, if we use a minimum acceptable value, in design units, of 0.999999 for the second factor and a maximum value of 1.0, we might expect to have a difficult time reaching a fitted value of at least 12.45 for the first response because the coefficient of the second factor in the fitted equation is -2.486. Thus, a considerable increase in the value of the second factor will result in a considerable decrease in the fitted value for the first response. The fitted value is 12.71, however, and the composite desirability given by JMP is only 0.537. The other fitted values were 3832.62 and 9.40.

Using Design-Expert, if we try to set the level of CF_4 Flow at 15 (i.e., the coded level set to 1), we obtain the results shown in Table 12.2.

As can be seen from the output, Design-Expert used nine starting points in seeking the optimum solution. Factors A and B are each set at the high level in each of the five solutions, with the level of factor C then obtained in such a way as to produce composite desirability values that happen to be equal. As expected, the value of factor

TABLE 12.2 Solutions Given by Design-Expert

Constraints

Name	Goal	Lower Limit	Upper Limit	Lower Weight	Upper Weight	Importance
A	is in range	−1	1	1	1	3
B	is equal to 1.00		1	1	1	3
C	is in range	−1	1	1	1	3
R1	maximize 12.45	15	1	1	1	3
R2	maximize 3068	3300	1	1	1	3
R3	minimize 10	10.3	1	1	1	3

Solutions

Number	A	B	C	R1	R2	R3	Desirability
1	1.00	1.00	1.00	12.8488	3816	9.375	0.539
2	1.00	1.00	−0.01	12.8487	3473.34	9.83171	0.539
3	1.00	1.00	0.19	12.8487	3541.83	9.74042	0.539
4	1.00	1.00	0.55	12.8487	3662.81	9.57918	0.539
5	1.00	1.00	0.02	12.8487	3486.32	9.81441	0.539

5 Solutions found

Number of Starting Points: 9

A	B	C
-1.000	1.000	1.000
0.000	0.000	0.000
-1.000	-1.000	-1.000
-1.000	-1.000	1.000
1.000	1.000	1.000
1.000	-1.000	1.000
1.000	1.000	-1.000
1.000	-1.000	-1.000
-1.000	1.000	-1.000

C varies directly with $R2$ and inversely with $R3$ since the coefficient of factor C in the model for $R2$ is positive and it is negative in the model for $R3$.

Upon seeing this output, it would be prudent for the user to simply increase the maximum $R2$ value, since $R2$ is to be maximized and all the solutions exceeded the upper bound, and run it again. Raising the maximum to 4000 predictably reduces the number of solutions to one and that solution is the first solution, with a composite desirability of 0.501.

Czitrom et al. (1998) set the first factor at its largest design level and this level is also part of the “optimal” solutions given here, so it is not possible to better their solution for the first response variable without fitting a different model—which they did. They used a significance level of .10 in their analysis, which is a liberal value. A parsimonious model is desirable because using terms that are not significant will inflate the variance of the predicted response, as shown by Walls and Weeks (1969).

It is important to keep in mind that regardless of the desirability form that is used, the composite desirability is a statistic that has sampling variability just as do the parameter estimates in the fitted models and the predicted values. If the predicted responses have a large variance for a given combination of factor levels, then we would expect those variances to translate, in an unknown and undoubtedly complicated way (which has apparently not been investigated in the literature), into a variance of the composite desirability that may be undesirable. (As shown in Figure 12.3, the predicted responses using JMP are given with a 95% prediction interval, as stated previously, so the user can see whether or not the interval is too large. Theoretical results, including the distribution of desirability indices, are given by Govaerts and Le Bailly de Tilleghem, 2005).

In this example the overall desirability for each of the given solutions was not particularly good because the predicted response for selectivity was well short of the goal of 15 (chosen arbitrarily for this example). Reducing the goal value to, say, 12.5 increases the composite desirability considerably.

It is worth noting that the predicted response values are based on some rather strong assumptions, namely, that each model is the true model so that the model holds for points that were not used in the experiment. Furthermore, the variance of the predicted responses will not be small when there are only 10 design points.

There are two sets of values that can be specified when Design-Expert is used: Weight and Importance for each response. A value of 0.1 to 10 for Weight can be assigned to each response in Design-Expert, with the set of these values indicating the relative importance of the response variables. The default value is 1.0. A Weight factor cannot be specified in JMP, but an Importance value can be specified. The weights specified for each response determine the shape of the desirability function when the objective is to hit a target value and when the response is to be maximized or minimized, but they can have very little effect on the solution when the objective is to maximize or minimize the response. For example, using the last variation of Example 12.1, weights of 0.1, 4, and 10 for the three response variables do not cause the optimal solution to change, but the use of different weights will strongly affect the desirability values.

The set of Importance values could have a considerable effect on the solution, however. These values are used to specify the relative importance of optimizing each response. In Design-Expert the importance values can be 1, 2, 3, 4, or 5, as stated in Section 12.2, whereas JMP will apparently accept virtually any positive number.

To illustrate, let's assume that it is far more important to have a small value for Etch Rate Nonuniformity than a large value for the other two variables. In Design-Expert it would be logical to assign an Importance value of 1 to each of the first two response values and 5 to Etch Rate Nonuniformity. This does not affect the solution, however, as the solution given is the first solution listed in Table 12.2. Similarly, the JMP solution is also unaffected when a high Importance value is given to $R3$ and a low value given to the other two responses.

In general, the values assigned for Weight and Importance may not have any effect on the optimal solution, with the effect seen only in the composite desirability. The upper and lower limits specified as well as any factor constraints are likely to exert more influence on the solution than the Weight and Importance values. For example, using Weight values of 1, 0.1, and 10 in Design-Expert for the three responses, respectively, has no effect on the output other than to give a "second-best solution" that has a slightly lower desirability than the optimal solution, which is not affected by the considerable disparity in the weights. Similarly, using various combinations of Weight and Importance does not lead to a different solution. (Again, this assumes that the second factor is set equal to 15 and the upper and lower limits on the fitted response values are unchanged.)

The effect that these have on the composite desirability in JMP can be seen by changing the relative weights in the Prediction Profiler graphical display and observing how the composite desirability changes dynamically. (There is more discussion of software for multiple response optimization in Section 12.8.)

Example 12.2

As a second example, we describe and analyze partially a case study given by Myers (1985). A 3^2 design with three replicates was used and there were four response variables. We will use this as an example primarily for the purpose of determining whether or not optimization *should* be performed, based on the model that is selected for each response.

The experiment involved the drug gentamicin. The manufacturing of two reagents, a particle reagent and an antibody, was not well controlled when new process technology was introduced. Experimentation was needed to determine the volumes of the reagents to be used in an analytical test pack in order to maintain uniform pack performance properties. There were four response variables, Final Blank Absorbance and three rates for calibrators. The data, using the uncoded factor levels, are given below.

Particle Reagent	Antibody	Final Blank Absorbance	Calibrator Rates Blank	4μg/ml	16μg/ml
20	10	.24250	79	8	-1
35	10	.38315	316	-1	32
50	10	.44030	309	326	146
20	25	.37970	85	12	5
35	25	.61950	485	319	48
50	25	.73550	418	490	220
20	40	.51080	83	79	9
35	40	.81405	539	306	78
50	40	.95405	432	542	240
20	10	.23960	70	9	2
35	10	.36490	297	246	29
50	10	.43890	307	329	141
20	25	.38320	84	15	5
35	25	.62520	492	281	59
50	25	.73140	435	428	192
20	40	.50990	77	-2	8

35	40	.81370	544	315	72
50	40	.94520	450	529	236
20	10	.25640	53	8	2
35	10	.37920	310	237	28
50	10	.43530	312	321	129
20	25	.38455	87	14	−2
35	25	.61195	483	337	53
50	25	.73495	435	355	80
20	40	.51400	84	18	9
35	40	.79975	541	35	70
50	40	.95600	425	523	253

Part of the output from Design-Expert is given below.

```
ANOVA for selected factorial model

Analysis of variance table [Classical sum of squares - Type II]
            Sum of              Mean              F        p-value
Source      Squares     df      Square            Value    Prob > F
Model        1.29       8       0.16             4277.82   < 0.0001
 A-A         0.50       2       0.25             6672.88   < 0.0001
 B-B         0.74       2       0.37             9783.36   < 0.0001
 AB          0.049      4       0.012             327.53   < 0.0001
Pure Error 6.791E-004  18       3.773E-005
Cor Total   1.29       26

The "Pred R-Squared" of 0.9988 is in reasonable agreement with the
"Adj R-Squared" of 0.9992.
```

We immediately notice that $R^2 = .999$, even though the design is replicated and there are 8 degrees of freedom for the model terms and 26 degrees of freedom overall. Although such an extremely high R^2 value is not uncommon in a few areas of application, such as calibration, this high value should cause one to question the data. Myers (1985) performed an analysis using the median response values and found that of the eight estimable effects (linear and quadratic for each of the two factors), and four interaction components obtained from all possible combinations of linear and quadratic for each factor, the numbers of effects that were significant at the .05 level of significance were 7, 8, 7, and 5 for the four response variables, respectively. Thus, there were 27 significant effects out of 32 for the four response variables combined. Certainly there should not be such a very high percentage of estimable effects that are significant.

The following output for the second response variable, again from Design-Expert, shows that when the actual response values are used, every estimable effect is significant since each parameter estimate's 95% confidence interval does not include zero, and does not even come close to including zero. Thus, this more or less agrees with the conclusions of Myers (1985). (Note that the "1" and "2" refer to the linear and quadratic effects, respectively.)

Term	Coefficient Estimate	df	Standard Error	95% CI Low	95% CI High
Intercept	304.89	1	1.55	301.63	308.14
A[1]	-226.89	1	2.19	-231.49	-222.29
A[2]	140.33	1	2.19	135.73	144.94
B[1]	-76.78	1	2.19	-81.38	-72.18
B[2]	28.89	1	2.19	24.29	33.49
A[1] B[1]	66.11	1	3.10	59.60	72.62
A[2] B[1]	-60.78	1	3.10	-67.29	-54.27
A[1] B[2]	-21.56	1	3.10	-28.06	-15.05
A[2] B[2]	12.56	1	3.10	6.05	19.06

According to Myers (1985) "The median response values were routinely used in the analysis, to make the fit less sensitive to blunders." Bad data points can cause terms to be significant that should not have been significant, but this won't necessarily inflate R^2. In Exercise 12.2 the reader is asked to determine if there are any outliers and to perform the optimization, if it is practical to do so considering what has been pointed out in this example.

12.6 MULTIPLE RESPONSE OPTIMIZATION VARIATIONS

As stated in Section 12.2, the original desirability function approach was given by Harrington (1965). For maximization he used $d = \exp[-\exp(a - bY)]$; for minimization, $d = 1 - \exp[-\exp(a - bY)]$, and for a target value $d = \exp|\frac{Y-T}{b}|^n$, with $a \le Y \le b$ being the acceptable range for the response variable Y, with n a constant.

Various purported improvements have been presented in the literature since the original work of Harrington (1965) and Derringer and Suich (1980). These were compared by Wurl and Albin (1999) and used in a case study. The work has focused on improved desirability functions. For example, Del Castillo, Montgomery, and McCarville (1996) pointed out that gradient-based optimality algorithms cannot be employed when the individual desirability functions illustrated in Section 12.1 and employed by Derringer and Suich (1980) are used. This is because gradient-based optimization methods, which are superior to hill-climbing techniques such as the Hooke–Jeeves method, as discussed in Section 12.1, require that the objective function have continuous first derivatives, but this requirement is not met by Eq. (12.2).

To remedy this problem, Del Castillo et al. (1996) proposed a piecewise continuous desirability function that corrects for the nondifferentiable points in the desirability functions used by Derringer and Suich (1980) by using a local polynomial approximation. The composite desirability function given in Eq. (12.1) is still used, and is computed using the piecewise desirability function for each response. The authors used the generalized reduced gradient (GRG) algorithm of Lasdon, Waren, Jain, and Ratner (1978), which has been used in Microsoft Excel, in optimizing the function. In the authors' example, however, the GRG algorithm produced practically the same solution as the Hooke–Jeeves algorithm. Undoubtedly, this will not always happen, however, as the solution obtained using the GRG algorithm could differ considerably from the solution obtained with the Hooke–Jeeves algorithm.

Réthy, Koczor, and Erdélyi (2004) proposed a modified desirability function approach that utilized a loss function approach that was in the spirit of Taguchi (loss functions) and Six Sigma, as did Ribardo and Allen (2003). See also the methods of Shah, Montgomery, and Carlyle (2004), Tong and Su (1997) and Ames, Mattucci, Macdonald, Szonyi, and Hawkins (1997). Tabucanon (1988) may also be of interest.

Ortiz, Simpson, Pignatiello, and Heredia-Langner (2004) proposed a genetic algorithm (GA) approach and compared its performance against the GRG algorithm used in conjunction with the composite desirability. The GA algorithm outperformed the GRG algorithm in their study.

Peterson (2004) proposed an entirely different approach, one that takes into consideration correlation between the response variables, which the desirability function approach does not do, and also the precision with which the parameters are estimated. Similarly, Vining (1998) proposed a mean squared error method that also utilized the variance–covariance structure of the multiple responses. Methods that utilize the correlation structure are important because multiple responses will often, if not usually, be correlated, as stated previously, just as measures such as height and weight are correlated in people.

Kim and Lin (2000) suggested an approach using an exponential desirability function that takes into account the differences in predictability of the models for the different responses as well as priorities for the response variables. A good physical interpretation of the objective function value is one of the touted advantages of the proposed approach; the authors do note some possible disadvantages, however.

Various methods have been compared by Kros and Mastrangelo (2001, 2004), with Kros and Mastrangelo (2004) concluding that the combination of (all of) "maximize," "minimize," and "hit the target" for response variables in a problem may dictate how likely the target is to be achieved.

Unpublished work on new desirability functions includes Gibb, Carter, and Myers (2001), who proposed using $d = [1 + \exp(-\frac{Y-a}{b})]^{-1}$, $d = [1 - \exp(-\frac{Y-a}{b})]^{-1}$, and $d = \exp[-\frac{1}{2}(\frac{Y-T}{b})^2]$, for maximization, minimization, and the target value cases, respectively.

Unfortunately, practitioners who would like to use any of these alternative approaches may be stymied by the fact that they are not presently available in statistical software. Regarding software, Ramsey, Stephens, and Gaudard (2005) showed how to perform multiple optimization using JMP.

Example 12.3

Returning to a more conventional example, we consider the example with two response variables given by Vining (1998). Engineers performed an experiment involving reaction time (A), reaction temperature (B), and the amount of catalyst (C) in an effort to determine the optimum levels of these factors to maximize the conversion of a polymer (Y_1) and achieve a target value of 57.5 for the thermal activity (Y_2). A central composite design was used with $\alpha = 1.682$ and six centerpoints were used. Vining (1998) used this example to illustrate the compromise approach he advocated, which was mentioned earlier in this section. The data are given below.

A	B	C	Y_1	Y_2
-1	-1	-1	74	53.2
1	-1	-1	51	62.9
-1	1	-1	88	53.4
1	1	-1	70	62.6
-1	-1	1	71	57.3
1	-1	1	90	67.9
-1	1	1	66	59.8
1	1	1	97	67.8
-1.682	0	0	76	59.1
1.682	0	0	79	65.9
0	-1.682	0	85	60.0
0	1.682	0	97	60.7
0	0	-1.682	55	57.4
0	0	1.682	81	63.2
0	0	0	81	59.2
0	0	0	75	60.4
0	0	0	76	59.1
0	0	0	83	60.6
0	0	0	80	60.8
0	0	0	91	58.9

We can use this example to illustrate the different solutions provided by different software, along the lines of what was shown in Example 12.1 in Section 12.5.

Vining (1998) stated that the engineers could justify a second-order model with nine terms: the three linear and the three pure quadratic effects, plus all three interaction terms. A preliminary analysis, shown below, suggests that the data do not support such a model, however, as the effects that have p-values less than .05 are B, C, B^2, C^2, AC, and BC. (A separate analysis showed that the ABC interaction was not significant.)

```
Response Surface Regression: Y1 versus A, B, C

Estimated Regression Coefficients for Y1

Term          Coef  SE Coef       T      P
Constant    81.091    1.924  42.153  0.000
A            1.028    1.276   0.806  0.439
B            4.040    1.276   3.166  0.010
C            6.204    1.276   4.861  0.001
A*A         -1.834    1.242  -1.476  0.171
B*B          2.938    1.242   2.365  0.040
C*C         -5.191    1.242  -4.179  0.002
```

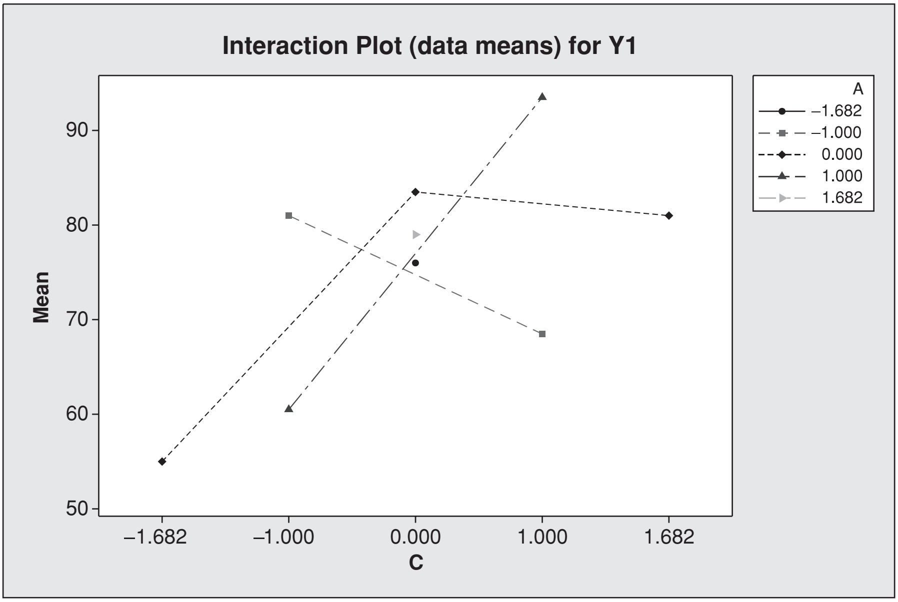


Figure 12.4 AC interaction for Example 12.3 data.

```
A*B          2.125    1.668   1.274   0.231
A*C         11.375    1.668   6.821   0.000
B*C         -3.875    1.668  -2.324   0.043

S = 4.717      R-Sq = 92.0%     R-Sq(adj) = 84.8%
```

We can see that the AC interaction effect is by far the largest effect, being almost twice as large as the next largest effect. This is obviously the explanation for the nonhierarchical model. We can see the problem if we construct the AC interaction plot (Fig. 12.4).

We need only consider the factorial points since the non-factorial points are not used in the computation of the interaction effects. The graph shows that there is a large A effect at the high level of C and a large A effect at the low level of C. There is not much difference in the response values at the two axial point values for A, and the response values at $A = 0$ do not affect the regression coefficient of the linear effect of A if a regression approach is used to determine the significant effects. Hence, the p-values are unaffected. This figure shows that the engineers were justified in assuming that there was an A effect, but the true nature of the effect cannot be seen without looking at this graph and/or computing conditional effects.

Conditional effects were discussed in a mostly general way for response surface designs in Section 10.9. Here the question arises as to whether or not a very large interaction, such as the AC interaction in this example, could mask a real quadratic effect in either A or C. This is an important question since Vining (1998) stated that "the engineers could justify" the model that included the quadratic term in factor

A but the analysis shows the quadratic term to not be significant. For designs with more than two levels, there is not a simple way to decompose interaction effects into the difference of conditional main effects, as was shown for two-level designs in Chapter 4, Appendix A. For example, for designs with three levels, the middle level is used in computing the quadratic effect of a factor, under one set of definitions, but is not used in computing the linear $\times$ linear interaction component(s). We should remember, however, as mentioned in Section 6.1.2 for three-level designs, that there is not just a single way to compute effect estimates when designs have more than two levels, as is also discussed for three-level designs by Giesbrecht and Gumpertz (2004, p. 291).

The situation is more complicated for response surface designs, some of which have five levels. If we use a linear model approach for the analysis of data from such designs, there is not a simple way to express a quadratic effect estimate since quadratic effects are neither orthogonal to each other nor to a column of ones.

Therefore, there is not a clean way of looking at quadratic effect estimates separately, let alone developing a relationship between the magnitude of quadratic effects and the magnitude of interaction effects. Consequently, we will not pursue this further and for this scenario it would have perhaps been prudent to simply rely on the engineers' judgment and include the A^2 term.

Although it may be annoying to use a model with coefficients for main effect terms that cannot be justified with a conditional effects analysis, the relevant question is whether or not the fitted values move in the proper directions as the factor levels are changed.

Continuing with the example, Vining (1998) stated that the engineers were comfortable with a model for Y_2 that contained only the A and C main effects. Although there are some interactions that are large relative to the B main effect, they are small relative to the A and C effects, so a model with only those two terms seems appropriate.

The smallest acceptable value for Y_1 was considered to be 80 and 100 or above was most desirable. The acceptable range for Y_2 was considered to be 55–60, with a target value of 57.5, as stated previously.

Design-Expert gives the following solution(s).

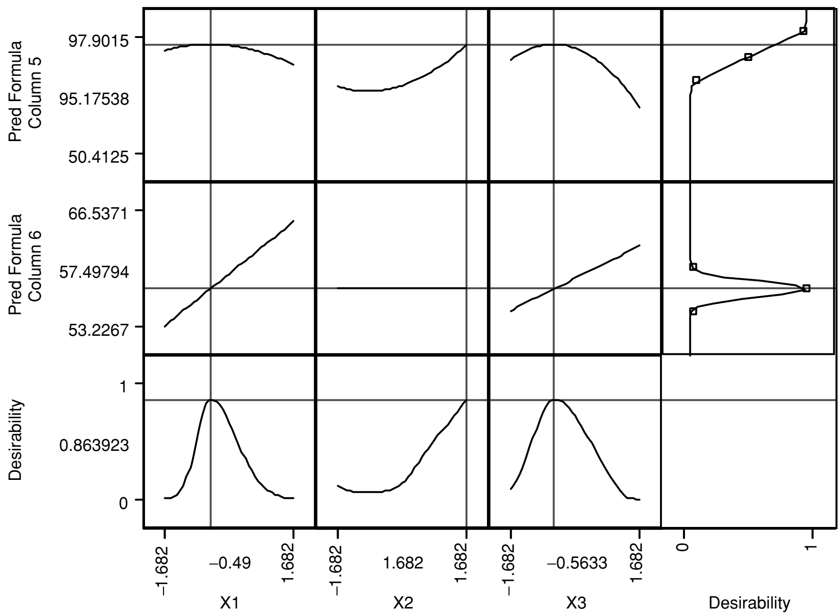
Name	Goal	Lower Limit	Upper Limit	Lower Weight	Upper Weight	Importance
A	is in range	-1.682	1.682	1	1	3
B	is in range	-1.682	1.682	1	1	3
C	is in range	-1.682	1.682	1	1	3
Y1	maximize	80	100	1	1	1
Y2	is target = 57.5	55	60	1	1	1

Solutions

Number	A	B	C	Y1	Y2	Desirability	
1	-0.49	1.68	-0.56	95.1754	57.4999	0.871	Selected
2	-0.49	1.68	-0.56	95.1742	57.5001	0.871	
3	-0.48	1.68	-0.58	95.1738	57.4999	0.871	
4	-0.54	1.68	-0.48	95.0832	57.5001	0.868	
5	-0.91	-1.68	0.11	83.6089	57.5001	0.425	
6	-0.91	-1.68	0.12	83.6081	57.5001	0.425	

The best solution(s) does not differ noticeably from the solutions given by Vining (1998) for the Derringer and Suich (1980) approach, which is what is being used here. This is especially true in regard to the fitted value of Y_1, which Vining gave as 95.21, almost the same as given in the Design-Expert output, and there is agreement with the fitted value of 57.5 for Y_2. Although Vining (1998) did not state that one of the response variable objectives was more important than the other one, changing the relative Importance values in Design-Expert has very little effect on the solution(s). Specifically, giving an Importance value of 1 for Y_1 and 5 for Y_2 gives a fitted value of 97.1753 for Y_1, compared to 97.1754 when they have equal Importance values. The only noticeable difference in the factor levels is that $C = -0.57$ with this unequal weighting, rather than -0.56 with the equal weighting. Reversing the Importance values gives the same solution as when using equal Importance values. Of course the composite desirability values differ greatly as the composite desirability is naturally quite high (.955) when the Importance values are 1 for Y_1 and 5 for Y_2 since the optimal solution hits the target value for Y_2. Reversing the Importance values gives a composite desirability of .794.

When JMP is used, convergence to an optimal solution is not achieved, but the last solution computed is given below.



Despite the nonconvergence, notice that this solution is essentially the same as the solution given by Design-Expert. This is interesting because JMP and Design-Expert use different desirability function approaches, as was explained in Section 12.5.

One way to observe the “non-uniqueness” of the JMP solution is to run “Maximize Desirability” from this solution and notice the small change in the solution that results each time this is done, along with the accompanying message of nonconvergence.

Interest should focus, however, on the more important issue of trying to minimize the variance of the predicted responses by using parsimonious models. Unless a poor optimization algorithm is being used, and that does not appear to be the case with JMP, a solution that is far from the true optimal solution should very rarely occur. Sampling variability due especially to a small design and/or too many terms in the model should always be considered and minimized to the degree possible.

The approach recommended by Vining (1998) minimizes a multivariate general squared error loss function, under the assumption that the variance–covariance matrix of the response variables is known. The loss function contains a matrix **C** that can be used to represent costs or simply a differential weighting of the relative importance of the various response variables. The two solutions given by Vining (1998), the second of which assigns a very high weight to the second response variable, are both inferior to the best solution(s) given by Design-Expert, as in particular, neither solution is particularly close to the target value for Y_2.

12.7 THE IMPORTANCE OF ANALYSIS

As stated in other parts of the book, the emphasis has been on design and not analysis, although in some places there has been considerable analysis. In multiple response optimization, however, the emphasis is on analysis, which begins with an appropriate model fit to each response variable. A parsimonious model is desirable so as to minimize the variance of the fitted responses, remembering that the "optimal solution" is conditioned on the dataset, and the solution may not even be the optimal solution for the given data.

The variance of the predicted responses should be minimized as doing so will minimize the sampling variance of the optimal solutions, which is quite important, regardless of whether the computing algorithm is effective in identifying the optimal solution for a given dataset.

The datasets should of course also be carefully scrutinized, especially checking to see if there are any bad data points. Accordingly, the reader is asked in Exercise 12.9 to check and see if there are any suspicious data points relative to the second response variable and its fitted model, and if so, to determine the effect that deleting the point has on the optimal solution, especially since the design becomes unbalanced if a deleted point is not a centerpoint.

For additional information, see Chiao and Hamada (2001), who proposed a method for analyzing correlated multiple response data from designed experiments.

12.8 SOFTWARE

One of the unique features of Design-Expert is that up to 30 combinations of factor levels can be produced graphically that all have a desirability value of 1.0, although of course the levels cannot be read exactly from the graph. Another unique feature, and this was pointed out by Reece (2003, p. 375), is that the terms to use in the model for each response can be selected from the half-normal probability plot. What is somewhat odd about this is that the initial plot doesn't show any effects as being significant. That is, neither Lenth's method (see Section 4.10) nor any other method is used to identify significant effects, so this has to be done by the user. (It would be better if the half-normal probability plot showed the significant effects.) When an effect is selected, a line is fit to the remaining points so the line changes dynamically, but there is still no identification of significant effects from the ones that have not been hand selected. Although opinions will likely differ on the value of this setup, Reece (2003, p. 375) stated, "The use of the half-normal probability plot in this manner is a most useful approach to analyzing relatively simple designs without resorting to formal regression analysis".

Although data analysis diagnostics are not being used in this book since the emphasis is on design construction and identification of useful designs, it is worth noting that Design-Expert has excellent (regression) diagnostic capabilities that can be used very easily with multiple response optimization or with any of the other design features. This combined with the other features and the guidance (in blue) that appears with the output when the model for each response is fit essentially put Design-Expert at or at least near the top of the list of statistical software for multiple response optimization.

MINITAB can handle multiple responses through its multiple response optimizer. As a graphical aid, up to 10 contour plots can be overlaid; the use of this capability for analyzing two responses was illustrated in Figures 12.1 and 12.2. Although different models can be fit when the multiple response optimizer is used in menu mode, those models would have to be determined before using that capability because the default is to fit the model with all estimable terms for each response. Once the models are fit in menu mode, they cannot be modified in that mode because the procedure remembers the models that were fit and there is no facility that would permit any changes. Therefore, the models should be selected using regression methods before the optimizer is used. This is true regardless of whether menu mode or command mode is used. If a user does not remember to do this before the optimizer is used in menu mode, then it will have to be done afterward and the models specified in command mode, which is not as convenient. It is not terribly inconvenient, however, as a "trick" can be used of cutting and pasting the main command and subcommands that are displayed when menu mode is used after appropriate modification of the first subcommand.

In discussing MINITAB, Reece (2003, p. 435) stated, "The system does not provide convenient mechanisms for simplifying models to remove terms that are not significant. Therefore, the author simply used the full regression models for each response." Although this was in regard to the general fitting of regression models to multiple responses rather than directly regarding the multiple response optimizer, this may be a typical response and shows that the author used menu mode. Of course models with only significant terms can be fit very easily in command mode just by doing it in two stages.

JMP also has multiple response capability but the same model must be used for each response with the "Fit Model" capability. A workaround is described in the *JMP Statistics and Graphics Guide* when different models are to be fit, but this requires storing the "prediction formula" (i.e., prediction equation) for each response and using the Profiler option in the Graph menu. This is easily done, however, but Design-Expert is slightly easier to use and also has somewhat superior capabilities.

There are other statistical software packages that have multiple response optimization capabilities (see Reece, 2003, p. 334). The software packages that received the highest rating in the category of "Optimizations" were Design-Expert, Echip, Cornerstone (which is not widely known and is actually referred to as a design module), and JMP. (Although the company that produces the Echip software is listed as being out of business, the software is still supported.)

When using software for which the initial fitted model, such as a full quadratic model, is automatically fit for each response, the user must intervene and identify the significant terms, and then fit an appropriate subset model.

It is worth noting that the optimal solution is not guaranteed with any of these software packages. A convergence criterion (or multiple criteria) must be used and it is important that software use a good optimization routine to minimize the risk of identifying local optima as the global optimum. Cornerstone requires a convergence criterion as part of its input, whereas, at the other extreme, such input is not permitted with MINITAB. It is also worth noting that the optimal solution is, at best, optimal only in regard to maximizing the composite desirability. The combination of factor levels that is required for the optimal solution might not be feasible, as the combination may not be in a feasible region or the particular combination may be debarred for some reason. (See Section 13.9 for a discussion of restricted regions of operability and debarred observations.) Myers and Montgomery (1995, p. 252) make a similar suggestion in stating "... the researcher should do confirmatory runs at that condition to be sure that he or she is satisfied that all responses are in a satisfactory region." Although constraints can be specified with Design-Expert, debarred observations cannot be specified in the multiple response optimization routine (or in statistical software in general). Therefore, the user should check to see if the combination of factor levels produced by the optimization routine is feasible.

12.9 SUMMARY

The use of multiple response optimization methods will undoubtedly increase over time as their implementation in statistical software becomes better known to practitioners. No longer is it necessary to use the Solver routine in Excel to solve multiple response optimization problems, as was done in the late 1990s by at least one researcher whose work has been cited in this chapter.

Undoubtedly algorithmic approaches based on the Derringer and Suich (1980) approach are the most frequently used, although alternative approaches were proposed many years ago.

As new and improved methods are developed, it will be interesting to see how fast software developers move to implement them. This will undoubtedly depend on their perception of frequency of use of the routines in their software. Certainly the frequency of use pales in comparison to the use of ANOVA and regression, but in practice experimental design applications usually have multiple response variables, so there is a great need for the best multiple response optimization methods to be implemented in the leading statistical software packages.

The selection of "maximize," "minimize," or "hit the target" is obviously dictated by the problem, but there is some evidence that the combination of these in a particular problem will dictate how likely the target is to be achieved (Kros and Mastrangelo, 2004).

A desirability function approach can also be used when there is only a single response, which of course is a much simpler problem. This is discussed in Section 14.1.

REFERENCES

Ames, A. E., N. Mattucci, S. Macdonald, G. Szonyi, and D. M. Hawkins (1997). Quality loss functions for optimization across multiple response surfaces. *Journal of Quality Technology*, **29**(3), 339–346.

Chen, H., A. S. Hedayat, and C.-Y. Suen (1998). A family of optimal designs for experiments with multiple responses. *Journal of Combinatorics, Information and Systems Sciences*, **23**, 259–269.

Chiao, C.-H. and M. Hamada (2001). Analyzing experiments with correlated multiple responses. *Journal of Quality Technology*, **33**(4), 451–465.

Copeland, K. and P. Nelson (1996). Dual response optimization via direct function minimization. *Journal of Quality Technology*, **28**, 331–336.

Czitrom, V., J. Sniegowski, and L. D. Haugh (1998). Improved integrated circuit manufacture using a designed experiment. In *Statistical Case Studies: A Collaboration Between Academe and Industry* (R. Peck, L. D. Haugh, and A. Goodman, eds.). Philadelphia: Society of Industrial and Applied Mathematics; Alexandria, VA: American Statistical Association.

Del Castillo, E. and D. C. Montgomery (1993). A nonlinear programming solution to the dual response problem. *Journal of Quality Technology*, **25**, 199–204.

Del Castillo, E., D. C. Montgomery, and D. R. McCarville (1996). Modified desirability functions for multiple response optimization. *Journal of Quality Technology*, **28**(3), 337–345.

Derringer, G. (1994). A balancing act: Optimizing a product's properties. *Quality Progress*, **27**(6), 51–58. (http://www.statease.com/pubs/derringer.pdf)

Derringer, G. and R. Suich (1980). Simultaneous optimization of several response variables. *Journal of Quality Technology*, **25**, 199–204.

Ding, R., D. K. J. Lin, and D. Wei (2004). Dual response surface optimization: A weighted MSE approach. *Quality Engineering*, **16**(3), 377–385.

Ellekjaer, M. R., H. T. Fuller, and K. Ladstein (1997–1998). Analysis of unreplicated split-plot experiments with multiple responses. *Quality Engineering*, **10**(1), 25–36.

Gibb, R. D., W. H. Carter, Jr., and R. H. Myers (2001). Incorporating experimental variability in the determination of desirable factor levels. Unpublished manuscript.

Giesbrecht, F. G. and M. L. Gumpertz (2004). *Planning, Construction, and Statistical Analysis of Comparative Experiments*. Hoboken, NJ: Wiley.

Govaerts, B. and C. Le Bailly de Tilleghem (2005). Distribution of desirability in multi-criteria optimization using desirability functions based on the cumulative distribution function of the standard normal. (manuscript, see http://www.stat.ucl.ac.be/ISpub/dp/2005/dp0531.pdf)

Harrington, E. C., Jr. (1965). The desirability function. *Industrial Quality Control*, **21**(10), 494–498.

Hooke, R. and T. A. Jeeves (1961). Direct search solution of numerical and statistical problems. *Journal of the Association of Computing Machinery*, **8**(2), 212–229.

Khuri, A. I. and M. Conlon (1981). Simultaneous optimization of multiple responses represented by polynomial regression functions. *Technometrics*, **23**(4), 363–375.

Kim, K. and D. K. J. Lin (2000). Simultaneous optimization of mechanical properties of steel by maximizing exponential desirability functions. *Journal of the Royal Statistical Society, Series C*, **48**, 311–325.

Koksoy, O. and N. Doganaksoy (2003). Joint optimization of mean and standard deviation using response surface methods. *Journal of Quality Technology*, **35**(3), 239–252.

Kros, J. F. and C. M. Mastrangelo (2001). Comparing methods for the multiple-response design problem. *Quality and Reliability Engineering International*, **17**, 323–331.

Kros, J. F. and C. M. Mastrangelo (2004). Comparing multi-response design methods with mixed responses. *Quality and Reliability Engineering International*, **20**, 527–539.

Lasdon, L. S., A. D. Waren, A. Jain, and M. Ratner (1978). Design and testing of a generalized reduced gradient code for nonlinear programming. *ACM Transactions on Mathematical Software*, **4**(1), 34–50.

Lin, D. and W. Tu (1995). Dual response surface optimization. *Journal of Quality Technology*, **27**, 34–39.

Myers, G. C. (1985). Use of response surface methodology in clinical chemistry. In *Experiments in Industry: Design, Analysis and Interpretation of Results* (R. D. Snee, L. B. Hare, and J. R. Trout, eds.). Milwaukee, WI: American Society for Quality Control.

Myers, R. H. and D. C. Montgomery (1995). *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*. New York: Wiley. (2nd ed. in 2002)

Ortiz, F., Jr., J. R. Simpson, J. J. Pignatiello, and A. Heredia-Langner (2004). A genetic algorithm approach to multiple-response optimization. *Journal of Quality Technology*, **36**(4), 432–450.

Palamakula, A., M. T. H. Nutan, and M. A. Khan (2004). Response surface methodology for optimization of and characterization of limonene-based coenzyme Q10 self-nanoemulsified capsule dosage form. *AAPS PharmsciTech*, **5**(4), 1–8.

Peterson, J. J. (2004). A posterior predictive approach to multiple response surface optimization. *Journal of Quality Technology*, **36**(2), 139–153.

Ramsey, P. J., M. L. Stephens, and M. Gaudard (2005). Multiple optimization using the JMP® statistical software. Presentation made at *the Kodak Research Conference*, May 9, 2005. (available at www.northhavengroup.com/pdfs/MultipleOptimization.pdf)

Reece, J. E. (2003). Software to support manufacturing systems. In *Handbook of Statistics*, Vol. 22, Chap. 9 (R. Khattree and C. R. Rao, eds.). Amsterdam: Elsevier Science B.V.

Réthy, Z., Z. Koczor, and J. Erdélyi (2004). Handling contradicting requirements using desirability functions. *Acta Polytechnica Hungarica*, **1**(2), 5–12.

Ribardo, C. and T. T. Allen (2003). An alternative desirability function for achieving 'Six Sigma' quality. *Quality and Reliability Engineering International*, **19**, 227–240.

Ryan, T. P. (2000). *Statistical Methods for Quality Improvement*, 2nd ed. New York: Wiley.

Schmidt, R. H., B. L. Illingworth, J. C. Deng, and J. A. Cornell (1979). Multiple regression and response surface analysis of the effects of calcium chloride and cysteine on heat-induced whey protein gelation. *Journal of Agricultural and Food Chemistry*, **27**, 529–532.

Shah, H. K., D. C. Montgomery, and W. M. Carlyle (2004). Response surface modeling and optimization in multiresponse experiments using seemingly unrelated regressions. *Quality Engineering*, **16**(3), 387–397.

Tabucanon, M. T. (1988). *Multiple Criteria Decision Making in Industry*. New York: Elsevier.

Taguchi, G. (1986). *Introduction to Quality Engineering*: *Designing Quality into Products and Processes*. White Plains, NY: Kraus International Publications.

Tang, L. C. and K. Xu (2002). A unified approach for dual response optimization. *Journal of Quality Technology*, **34**(4), 437–447.

Tong, L.-I. and C.-T. Su (1997). Optimizing multiple response problems in the Taguchi method by fuzzy multiple attribute decision making. *Quality and Reliability Engineering International*, **13**, 25–34.

Vining, G. G. (1998). A compromise approach to multiresponse optimization. *Journal of Quality Technology*, **30**(4), 309–313.

Vining, G. G. and R. H. Myers (1990). Combining Taguchi and response surface methodologies: A dual response approach. *Journal of Quality Technology*, **22**, 38–45.

Walls, R. C. and D. L. Weeks (1969). A note on the variance of the predicted response in regression. *The American Statistician*, **23**, 24–26.

Wisnowski, J. W., G. C. Runger, and D. C. Montgomery (1999–2000). Analyzing data from designed experiments: A regression tree approach. *Quality Engineering*, **12**(2), 185–197.

Wurl, R. C. and S. L. Albin (1999). A comparison of multiresponse optimization: Sensitivity to parameter selection. *Quality Engineering*, **11**(3), 405–415.

EXERCISES

12.1 A desirability function analysis was used in the article "Stencil printer optimization study" by J. V. Stephenson and D. Drabenstadt (*Surface Mount Technology*, November 1999), which is available on the Internet at http://smt.pennnet.com/Articles/Article_Display.cfm?Section=Archives&Subsection=Display&ARTICLE_ID=111949&KEYWORD=printing. Is this the same type of desirability function that was presented in Section 12.2. If not, explain what is actually being done.

12.2 Consider Example 12.2 and the motivation for using medians that was given by Myers (1985). Recall that an unrealistically high fraction of estimable effects was significant, with all the estimable effects for the second response being significant. Does there appear to be any outliers on any of the responses that could have caused this result? If not, would you suggest that optimization be performed? Explain. If optimization seems reasonable, perform the optimization and, if possible, compare your results with those given in Figure 5 of Myers (1985).

12.3 One of the sample datafiles that comes with MINITAB is FACTOPT.MTW, which contains data for a 2^3 design with two replications, for a chemical reaction experiment. The three factors were reaction time, reaction temperature, and type of catalyst, and the two responses were yield and cost. The former is of course to be maximized and the latter is to be minimized. Use lower limit = 35 and upper limit = 45 for yield and lower limit = 28 and upper limit = 35 for cost and determine the optimal levels of the three factors. The data are as follows.

Row	Time	Temp	Catalyst	Yield	Cost
1	50	200	A	48.4665	31.7457
2	20	200	A	45.1931	31.0513
3	50	200	B	49.2040	36.8941
4	50	150	B	45.5991	32.6394
5	20	150	A	42.7636	27.5306
6	50	150	A	44.7592	29.3841
7	20	200	B	44.7077	34.6241
8	20	150	B	43.3937	30.5424
9	50	200	A	49.0645	32.3437
10	50	150	B	45.1531	33.0854
11	50	200	B	48.6720	37.4261
12	20	200	B	45.3297	35.2461
13	50	150	A	45.3932	28.7501
14	20	150	B	43.0617	30.2104
15	20	150	A	43.2976	28.0646
16	20	200	A	44.8891	30.7473

(a) Determine the optimal factor settings, assuming that yield and cost are equally important from a desirability function perspective.

(b) Now assume that the importance of maximizing yield is given a weight of 0.7 and of minimizing cost is given a weight of 0.3 and determine the optimal factor settings.

12.4 The MINITAB sample datafile RSOPT.MTW contains data on four factors: hot bar temperature, dwell time, hot bar pressure, and material temperature. The experiment focused on the seal strength of a bag, which was one of the response variables, with the variability in the seal strength being the other response variable. The objective was to determine the settings of the four factors so as to hit a target of 26 pounds for the seal strength, with 24 as the minimum acceptable value and 28 as the upper limit. The variance of the seal strength is to be between 0 and 1. Assume that these objectives are of equal importance. The data are as follows.

Row	HotBarT	DwelTime	HotBarP	MatTemp	Strength	VarStrength
1	200	1.00	50	110	12.4470	5.28200
2	225	0.75	100	90	24.7047	1.63296
3	125	0.75	100	90	20.6865	1.34636
4	175	0.75	100	90	29.1000	0.95000
5	175	0.75	100	50	27.4284	1.69595
6	200	0.50	150	110	30.3010	3.45200
7	150	1.00	150	110	23.2990	2.21400
8	175	0.75	100	90	28.3000	0.83000
9	175	0.75	0	90	25.9942	0.83801
10	200	0.50	50	110	27.6490	1.07600
11	175	0.75	100	90	28.2000	0.91000
12	175	0.75	100	90	28.7000	0.92000
13	175	0.75	100	90	27.4000	0.97000
14	150	0.50	150	70	12.0010	2.99600
15	200	1.00	150	110	15.6990	6.49000
16	200	1.00	50	70	8.2510	6.74600
17	150	0.50	50	110	12.2010	4.14400
18	200	0.50	50	70	26.7490	0.80000
19	175	1.25	100	90	21.3752	2.92266
20	150	0.50	50	70	10.5010	4.46000
21	150	1.00	50	70	15.6990	0.99400
22	175	0.75	100	90	28.9000	0.86000
23	200	1.00	150	70	13.7030	4.49400
24	175	0.25	100	90	25.5021	1.49727
25	150	0.50	150	110	15.7010	3.54000
26	175	0.75	200	90	30.0581	1.14190
27	175	0.75	100	130	30.0516	2.58797
28	200	0.50	150	70	28.4010	0.84000
29	150	1.00	150	70	21.5990	2.13800
30	175	0.75	100	90	28.5000	0.94000
31	150	1.00	50	110	19.7990	2.69000

It can be observed, although perhaps not too easily, that the design used is a central composite design with axial points of $\alpha = \pm 2$ in coded form and seven centerpoints. Fit a model to each response and perform the optimization, using the information on the objectives given at the start of the problem.

12.5 Derringer and Suich (1980, references) gave an example of an experiment with four response variables and three factors, and a central composite design with six centerpoints was used. The experiment involved the development of a tire tread compound, with the four response variables being PICO Abrasion Index (Y_1), 200 percent Modulus (Y_2), Elongation at Break (Y_3), and Hardness (Y_4). The first two were to be maximized, with 120 being the smallest acceptable value for Y_1 and a value of at least 170 being perfectly acceptable. For Y_2, the minimum acceptable value was 1000 and 1300 or more was completely acceptable. The value of Y_3 was to be between 400 and 600 and the target value was assumed to be 500. Similarly, Y_4 was to be between 60 and 75, and 67.5 was taken as the target value.

The three ingredients (i.e., factors) used in the experiment were hydrated silica level (A), silane coupling agent level (B), and sulfur level (C). The data are given below.

A	B	C	Y_1	Y_2	Y_3	Y_4
-1	-1	-1	103	490	640	62.5
1	-1	-1	120	860	410	65.0
-1	1	-1	117	800	570	77.5
1	1	-1	139	1090	380	70.0
-1	-1	1	102	900	470	67.5
1	-1	1	132	1289	270	67.0
-1	1	1	132	1270	410	78.0
1	1	1	198	2294	240	74.5
-1.633	0	0	102	770	590	76.0
1.633	0	0	154	1690	260	70.0
0	-1.633	0	96	700	520	63.0
0	1.633	0	163	1540	380	75.0
0	0	-1.633	116	2184	520	65.0
0	0	1.633	153	1784	290	71.0
0	0	0	133	1300	380	70.0
0	0	0	133	1300	380	68.5
0	0	0	140	1145	430	68.0
0	0	0	142	1090	430	68.0
0	0	0	145	1260	390	69.0
0	0	0	142	1344	390	70.0

The authors fit the same full quadratic model for each response and obtained a composite desirability of 0.583. This resulted from the following fitted values: $\widehat{Y_1} = 129.5$, $\widehat{Y_2} = 1300$, $\widehat{Y_3} = 465.7$, and $\widehat{Y_4} = 68$. These values resulted from using $X_1 = -0.05$, $X_2 = 0.145$, and $X_3 = -0.868$. Do you agree with their analysis? In particular, is the full quadratic model appropriate for each response variable? If not, give your solution and compare it with the authors' solution. If your solution is superior, what would you recommend to the authors?

12.6 Palamakula, Nutan, and Khan (2004, references) used a Box–Behnken design with three factors with the objective of optimizing self-nanoemulsified capsule dosage form of a highly lipophilic model compound, coenzyme Q10 (CoQ). The factors studied were R-(+)-limonene (A), surfactant (B), and cosurfactant (C). There were five response variables, although the authors studied only the first two. The two variables that are to be maximized are the percentage of the drug dissolved after the first 5 minutes (Y_1), and the percentage that is dissolved in 15 minutes (Y_2). The turbidity (Y_3) and particle size (Y_4) values should be under 150. This did happen in every run of the experiment except run #4, for which particle size exceeded 1000. The zeta potential (Y_5) should be under 25. The data are given below, with the factor values given in raw form rather than coded form.

A	B	C	Y_1	Y_2	Y_3	Y_4	Y_5
81	57.6	7.2	44.4	99.6	63.2	10.9	12.1
81	7.2	7.2	6.0	1.34	140	49.8	62.3
18	57.6	7.2	3.75	13.1	9.73	14.9	70.4
18	7.2	7.2	1.82	1.44	12.9	>1000	58
81	32.4	12.6	18.2	36.1	16.8	32.6	35.2
81	32.4	1.8	57.8	72.9	9.37	39.3	27
18	32.4	12.6	68.4	89.9	5.13	38.6	16.8
18	32.4	1.8	3.95	76.08	5.13	24.7	12.9
49.5	57.6	12.6	58.4	87.9	14.3	16.0	9.09
49.5	57.6	1.8	24.8	39.7	4.53	31.3	16.8
49.5	7.2	12.6	1.6	2.97	108	26.0	74.9
49.5	7.2	1.8	12.1	26.5	41.7	123	49.7
49.5	32.4	7.2	81.2	94.6	7.67	13.3	5.02
49.5	32.4	7.2	72.1	88.2	7.70	29.5	4.11
49.5	32.4	7.2	82.06	95.4	8.23	14.0	6.45

Maximize Y_1 and Y_2, assuming that they are of equal importance, while checking to make sure that the constraints on Y_3, Y_4, and Y_5 are met. Note the following. The authors' solution had $\widehat{Y_1} = 81.6$ and $\widehat{Y_2} = 95.9$, resulting from, in coded form, $A = 0.0344$, $B = 0.216$, and $C = 0.240$. Although the authors initially used a full quadratic model for each response variable, the article does not clearly indicate whether or not this model was also used in the optimization. (a) Do you agree with the authors' results? If not, support your solution and compare it to the authors' solution. (b) Are there large interactions that are creating problems relative to the main effect coefficients? Explain. If so, what you suggest be done about the problem?

12.7 Del Castillo, Montgomery, and McCarville (1996, references) gave an example of a wire bonding process optimization problem from the semiconductor industry to illustrate the improved optimization approach that they proposed. A Box–Behnken design was used for three factors: flow rate (A), flow temperature (B), and block temperature (C). Wires were bonded between two positions and the maximum, beginning, and finish bond temperature recorded for each position, for a total of six response variables. The data are given below with the factor levels given in the raw units and the first three response variables are "maximum," "beginning," and "finish," respectively, for the first position, followed by these variables in the same order for the second position.

A	B	C	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6
40	200	250	139	103	110	110	113	126
120	200	250	140	125	126	117	114	131
40	450	250	184	151	133	147	140	147
120	450	250	210	176	169	199	169	171
40	325	150	182	130	122	134	118	115
120	325	150	170	130	122	134	118	115
40	325	350	175	151	153	143	146	164
120	325	350	180	152	154	152	150	171
80	200	150	132	108	103	111	101	101
80	450	150	206	143	138	176	141	135
80	200	350	183	141	157	131	139	160
80	450	350	181	180	184	192	175	190
80	325	250	172	135	133	155	138	145
80	325	250	190	149	145	161	141	149
80	325	250	180	141	139	158	140	148

There is a target value for each response variable, as well as a minimum and maximum for each. These are given below.

Response variable	Minimum	Target	Maximum
Y_1	185	190	195
Y_2	170	185	195
Y_3	170	185	195
Y_4	185	190	195
Y_5	170	185	195
Y_6	170	185	195

Thus, the maximum bond temperature at each of the two positions has the same target value, and the other four response variables also have the same target value. The responses were considered to be equally important. The authors fit the same model to the last three response variables and the same model to the second and third variables. The model for the first response variable differed from these two models so there were three different models altogether. The authors obtained the following solution: $A = -0.0074$, $B = 1.0$, and $C = 0.7417$, with fitted values for the response variables of 186.9, 173.0, 170.1, 190.0, 170.9, and 182.4. Thus, the target value for the fourth response variable was hit exactly and the fitted value for the last response variable was close to the target value. Fit appropriate models for each response variable and perform the optimization. How does your solution compare to the authors' solution, especially in regard to the target values?

12.8 Consider the example in Section 12.1 with the data given in Table 12.1. Fit the model with only main effect terms in factors A, B, and C for the response variables Etch Rate and Etch Rate Nonuniformity. Then construct the contour plot, analogous to Figure 12.2, but this time fix factor C at the high level.

Compare your graph to Figure 12.2 and comment. In particular, does your graph show a feasible region?

12.9 Consider Example 12.3. Determine if there are any unusual data points (i.e., outliers) when the main effects of factors A and C are the only terms in the model. If there are any such points, what action would you suggest be taken regarding the point(s)? In particular, the design becomes unbalanced if deleted points are not centerpoints. Does this create a major problem? Note that Design-Expert still produces a fitted model and an optimal solution when there is a missing factorial point or axial point (or both). What does this suggest about how the data are used in obtaining the fitted model? If you find one or more suspicious data points and believe that they should be deleted, redo the optimization with the deleted point(s). If you deleted any points, does the deletion have much effect on the optimal solution?

12.10 Consider Example 6.3 in Section 6.4. There were three response variables and 30 observations per treatment combination used in a 2×3 design. (See the journal article for the data.) When multiple optimization is performed, it is worthwhile to consider that not only is the optimal solution a bit "random" even for a fixed set of data, as indicated in Section 12.5, relative to JMP solutions, but also of course the solution depends upon the response values, which are naturally random. In view of this but considering the number of observations per treatment combination, how well determined would the optimization likely be? If possible, read the article that is cited in Section 6.4 and try to determine an "optimal" solution.

12.11 Derringer (1994, references) gave some illustrative examples, one of which was a discrete case application in which there were four product formulations and three formulation properties: tensile strength, hardness, and elongation. There was a target value of 2000 pounds per square inch for tensile strength, with a small value of hardness desired, as well as a large value of elongation. Hardness was considered to be twice as important as tensile strength and elongation was considered to be four times as important as tensile strength. The four formulations with the individual desirabilities are given below.

Candidate formulation	Tensile strength	Desirability	Hardness	Desirability	Elongation	Desirability
1	1750	0.40	45	0.10	550	0.15
2	2000	1.00	30	0.97	500	0.00
3	1600	0.07	35	0.87	525	0.05
4	1600	0.07	30	0.97	585	0.47

(a) Determine the best candidate formulation, using the information that is given.

(b) How sensitive is the selection of the best formulation to the stated weights?

12.12 Wisnowski, Runger, and Montgomery (1999–2000, references) gave an example with four response variables and we will use the first three. The latter are the percentages of three bacterial populations killed in the production of a soap product. The design, given below along with the response values, was a 2_{V}^{5-1} design with four centerpoints, except that the centerpoints were not such on factor B since only two levels for that factor were used.

A	B	C	D	E	Bacteria 1	Bacteria 2	Bacteria 3
0	-1	0	0	0	25.53	18.60	88.38
-1	-1	-1	1	-1	54.90	12.80	99.95
1	-1	-1	-1	-1	36.53	32.76	99.96
0	-1	1	1	1	23.47	27.47	94.31
1	1	1	1	1	20.92	52.05	97.47
0	1	0	0	0	61.70	36.63	99.90
1	-1	-1	1	1	41.48	28.34	64.62
-1	1	-1	-1	-1	80.23	62.57	99.99
-1	1	1	1	-1	61.70	51.07	99.99
0	1	0	0	0	46.13	36.36	99.17
1	-1	1	1	-1	74.44	67.44	99.52
-1	-1	-1	-1	1	31.89	20.49	87.86
1	1	1	-1	-1	56.77	92.23	99.99
-1	1	-1	1	1	3.33	29.51	99.82
-1	1	1	-1	1	20.00	38.90	98.50
1	1	-1	1	-1	77.62	98.76	99.99
-1	-1	1	1	1	9.46	-8.28	93.90
-1	-1	1	-1	-1	80.26	47.77	99.98
1	-1	1	-1	1	35.28	24.84	95.58
1	1	-1	-1	1	38.55	13.69	99.99

(a) Does there appear to be any bad data points. If so, determine what correction action to take.

(b) Recognizing that each of the response variables should be maximized, determine the best combination of each factor to use to accomplish this after fitting a model for each response.

12.13 Consider again Example 12.1. The AB interaction was not used in the model for the first response variable for two stated reasons. Use this term in the model and redo the optimization, using appropriate software. Does this make much difference? Explain.

12.14 Khuri and Conlon (1981, references) gave two examples, one of which was from Schmidt, Illingworth, Deng, and Cornell (1979, references). The latter investigated the effects of cysteine (A) and calcium chloride (B) combinations on the textural and water-holding characteristics of dialyzed whey protein concentrate gel systems. There were four texture characteristics: hardness (Y_1), cohesiveness (Y_2), springiness (Y_3), and "combustible water" (Y_4). Each response was to be maximized. The data from the use of a central composite design were as follows.

A	B	Y_1 (kg)	Y_2	Y_3 (mm)	Y_4 (g)
-1	-1	2.48	0.55	1.95	0.22
1	-1	0.91	0.52	1.37	0.67
-1	1	0.71	0.67	1.74	0.57
1	1	0.41	0.36	1.20	0.69
-1.414	0	2.28	0.59	1.75	0.33
1.414	0	0.35	0.31	1.13	0.67
0	-1.414	2.14	0.54	1.68	0.42
0	1.414	0.78	0.51	1.51	0.57
0	0	1.50	0.66	1.80	0.44
0	0	1.48	0.66	1.79	0.50
0	0	1.41	0.66	1.77	0.43
0	0	1.58	0.66	1.73	0.47

Khuri and Conlon fit the same (quadratic) model to each response variable. Would you do the same? If not, fit appropriate models to each response and perform the optimization.