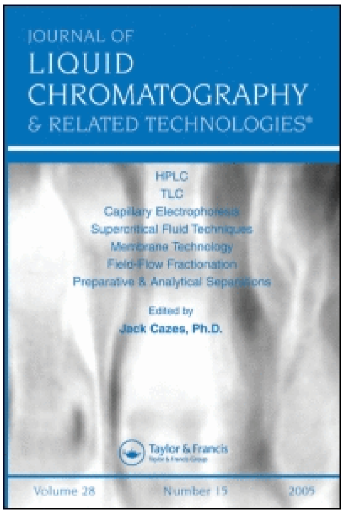

Journal of Liquid Chromatography & Related Technologies
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597273

Optimization of Selectivity in High-Performance Liquid Chromatography Using Mixture-Design Statistical Techniques: Overview and Software for Data Analysis

J. L. Glajch[a]; J. J. Kirkland[a]; J. M. Minor[a]
[a] Central Research and Development and Engineering Departments, E. I. du Pont de Nemours and Company Experimental Station, Wilmington, Delaware

To cite this Article Glajch, J. L. , Kirkland, J. J. and Minor, J. M.(1987) 'Optimization of Selectivity in High-Performance Liquid Chromatography Using Mixture-Design Statistical Techniques: Overview and Software for Data Analysis', Journal of Liquid Chromatography & Related Technologies, 10: 8, 1727 — 1747

To link to this Article: DOI: 10.1080/01483918708066796

URL: http://dx.doi.org/10.1080/01483918708066796

JOURNAL OF LIQUID CHROMATOGRAPHY, 10(8&9), 1727-1747 (1987)

OPTIMIZATION OF SELECTIVITY IN HIGH-PERFORMANCE LIQUID CHROMATOGRAPHY USING MIXTURE-DESIGN STATISTICAL TECHNIQUES: OVERVIEW AND SOFTWARE FOR DATA ANALYSIS

J. L. Glajch, J. J. Kirkland, and J. M. Minor
Central Research and Development and
Engineering Departments
E. I. du Pont de Nemours and Company
Experimental Station, Building 228
Wilmington, Delaware 19898

ABSTRACT

Optimization of separations in high-performance liquid chromatography (HPLC) is dependent on several factors: the proper choice of variables; a definition of criteria for a suitable separation; a rational strategy for examining the effects of these variables; and an efficient analysis of the data using manual methods or computer software. The selection of variables and, strategy for optimization is discussed in this paper, and a detailed description of a versatile software package for data analysis is presented along with provisions for obtaining the actual FORTRAN source code. The original mixture-design statistical software has been expanded and enhanced to accomodate data from both isocratic and gradient elution separations, single or multiple stationary phases, and other variables, such as temperature or pH. The software can be used to predict a single optimum set of conditions for a particular separation, and also to calculate the actual separation quality under any set of conditions bounded by the original experimental data. Examples of these functions are illustrated.

INTRODUCTION

Much of the strength of modern high-performance liquid chromatography (HPLC) is due to the effects of different variables that can be altered to improve the separation. The interaction of many of these effects has led in recent years to several strategies for optimization in HPLC. Although kinetic factors (column packing, flow rate, particle size, etc.) are important parameters, their effects have largely been maximized through small-particle, efficiently-packed columns operated under defined conditions. Therefore, for difficult separations, most systems for optimization have concentrated on thermodynamic or chemical effects for varying band spacing.

In previous work (1), we proposed a mixture-design statistical technique (MDST) for optimizing solvent selectivity in reverse-phase HPLC. This approach was later extended to normal-phase silica (2,3) normal-bonded-phase (4), mixed stationary-phase (5), ion-pair (6) and gradient elution (7,8) HPLC. In addition, a number of other investigators (9-13) have reported on optimized separations using the strategy recommended in Reference 1. Issaq et al have described a mixture-design statistical solvent-optimization method (including software) that is quite similar to that of Reference 1, also based on the solvent selectivity triangle approach (14).

Despite these descriptions and practical applications of MDST by us and others, we find that utility of the approach for the potential user is currently deficient for several reasons. First, although the design and advantages of this statistical approach have been presented in a series of papers, a single coherent, description of the full utility of our procedures has not been given. Second, the software for our method has never been published (although a limited version was in Du Pont's Sentinel(TM) HPLC system that is no longer commercially available). Finally, the current software is much more versatile than the original description (1). Therefore, we provide in this work a complete description of the MDST technique. Further, we document the software for analyzing the data and predicting optimum separations.

THEORY

Variables

In all optimization strategies, the first step is to identify the relevant variables that are likely to have an impact on the quality of the separation. Much of the potential power of HPLC as an analytical technique derives from the fact that there are many possible variables which can influence the separation; this factor can also be a impediment to efficient method development if improper variables are chosen.

Two types of variables have been identified in HPLC (indeed, in many other experimental systems, as well). These two types are discrete and related variables, and refer to whether the value of a variable is necessarily dependent on the value of some other variable in the system. Examples of discrete variables in HPLC are temperature, pH, ionic strength of a buffer, and column type. Examples of related variables include the amount of each type of solvent used in the mobile phase, or bonded-phase type, if mixed stationary phases are used. These latter variables are related, since the sum of all solvent or stationary-phase compositions must total 100%. This distinction is very important since the choice of experimental design strategy will depend upon whether discrete or related variables are being investigated, as discussed below.

An important aspect in optimization strategies are those variables that affect the thermodynamics of the separation (chemical effects) vs. those that influence the kinetics of the separation (physical effects). The latter variables (e.g., particle size, column efficiency, flow rate) certainly can be important to the final result. However, the effect of these variables often can be calculated by using well-known relationships (e.g., see (15)). Chemical effects such as type and amount of mobile and stationary phase, pH, and temperature, are often the most powerful in changing the selectivity of a separation. But unfortunately, these effects are not usually predictable from first principles. Consequently, variables associated with chemical effects are most emphasized in HPLC optimization strategies.

Choosing the retention range (k')

There are two possibilities for developing an optimum HPLC separation: isocratic or gradient elution. Previously, isocratic methods were preferred because of simplicity and concern regarding the lack of reproducibility of gradient elution apparatus. However, current HPLC instrumentation performs gradient separations admirably, and either approach is now viable for method development and appropriate optimization strategies.

The initial criterion in an optimization strategy is to choose a reasonable range of retention (k') for the separation. This retention range is a compromise between providing sufficient time that a separation with good resolution can occur (k' >1), but not so long (k' <20) so as to be inconvenient. If the sample contains materials of interest with a wide retention range, gradient elution is suggested. Samples with all compounds exhibiting retentions of 1<k'<20 are generally best separated by an isocratic operation. In either case, the initial choice of solvent strength or gradient slope can be simplified by the effective techniques described by Quarry, Grob, and Snyder (16). Minor alterations in solvent strength or gradient slope may then be made, if required, before proceeding with the remainder of the optimization procedure.

Selectivity Effects

Once a reasonable retention range has been chosen for either isocratic or gradient elution separations, the most marked improvement in separation generally will result from changing the selectivity or spacing for the peaks of interest. Although there are many ways to change selectivity in HPLC, the most powerful method is usually to change the composition of the mobile phase. Initial work by Snyder (17) has led to the use of solvent selectivity applications in reverse-phase (1,18,19), normal-phase (2,3), ion-pair (6) and normal-bonded-phase (4) chromatography. Many of the early applications were for isocratic systems on a single stationary phase. However, later work has expanded these selectivity effects for applications in gradient elution (8) and multiple or mixed-column systems (5,20). Other variables such as temperature (21) and pH (22) can also be useful in controlling selectivity in certain separations. Fortunately, any of these variables can be incorporated into an optimized method, with the software described below.

Statistical Experimental Design and Predictions

The most important task for any optimization scheme is to identify those parameters most likely to have a major effect on improving the separation. In other areas of analytical chemistry, the major variables in an analysis or measurement are not always apparent, and techniques such as simplex experimental design are useful to identify those variables. However, in HPLC, a separation method often is chosen for practical reasons involving sample type (reverse-phase for water-soluble compounds, ion-pair for ionic materials, etc.), and the most important and convenient variable to consider for optimization is the mobile phase modifier.

With this practical feature as a starting point, it would appear that a useful experimental design method for selecting an optimum mobile phase, based on well-established statistical methods, could be of great assistance in minimizing the number of experiments needed to develop a optimum separation. If designed properly, the experimental design would minimize the number of required experiments, so that a reasonable separation could be defined with the least development time.

EXPERIMENTAL DETAILS

Data

Data used here in were taken from previously published works (1,22), with appropriate corrections noted below. It is our intent in this paper to describe the software in such a way as to allow the reader, using the same data set, to generate results identical to those described.

Software

The software used to analyze the data and predict optimum separations is abstracted in Appendix I. Copies of the actual source code are available from the authors. The software was written in FORTRAN-77 and has been run on PDP-11/60 and VAX-11/780 minicomputers (Digitial Equipment Corp., Maynard, MA), and on an IBM-PC/AT (International Business Machines, Boca Raton, FL). The only changes in the code that were necessary for conversion among these machines, related to differences in the input/output files. The software available from the authors is specifically designed for a VAX-11/780.

There are six major sections of code comprising the optimization software. The first program, COLDG, is used to enter the data in a prescribed format. A second program, COLRG, is used to calculate the coefficients of the regression equations that describe the behavior of the data under various conditions. The main optimization program, COLGRD, calculates the optimum separation using the data and coefficients from COLDG and COLRG. COLPLP is used to plot an optimized mobile phase. COLPLP also can output the retention, separation (α or resolution) between any two peaks, or the separation that can be obtained under any set of selected conditions, not just those at the global optimum. COLLIB and COLLB1 are two sets of utility subroutines used by the main programs.

For each new set of data, COLDG requests and stores information on proportions of column and mobile-phase components, elution time, peak width, k', and elution volume information for each peak. The program is designed to calculate those values not explicitly given. For example, retention volumes are calculated from retention times and flow rate, if given. Individual peak widths (expressed as sigma values) can be entered; alternatively, a constant plate number (for isocratic) or a constant peak width (for gradient elution) can be simulated; only those values which are significantly different than the norm need be changed by the operator. After input of all information, the data is stored in a user-named file. It should be noted that this program adequately stores the necessary data, but is relatively cumbersome. Current spreadsheet-type data entry is more versatile and user-friendly. Future versions may incorporate this feature.

Once the primary data has been entered, COLRG performs the necessary regression calculations, using the equations for either quadratic or cubic fit as described by Snee (23). The subroutine TRMS forms a complete quadratic (response surface) set of terms from the components (eg., column, mobile phase, etc.). In this version, the effect of mixed-bed columns on selectivity is assumed to be linear, and mobile phase effects either quadratic or cubic. Higher-order terms could be generated with more chromatographic data. In addition to generating the regression coefficients, COLRG also calculates a residual standard error for each peak, if the number of experimental runs exceeds the number required to exactly fit the equations. For example, using a single column and three mobile phase modifiers, six runs are required to exactly fit the data to a quadratic surface. A seventh run is used to obtain a better fit of the data and to estimate the experimental error of the fit.

Results of the regression analysis can be used to predict elution of any of the compounds within the bounds of the experimental runs. COLGRD uses this principle to predict the elution of all components at each point of a grid covering all possible combination of the variables (mobile phase, stationary phase, etc.). By calculating the resolution of all peak pairs, COLGRD can then predict the optimum set of conditions that will separate all peaks to the best minimum resolution, using the Overlapping Resolution Map (ORM) concept (1). COLGRD also can be used to (a) predict retention of any peak, (b) the resolution of any peak pair, or (c) the minimum resolution for all peak pairs under any specified set of conditions (which may not be the optimum). This feature is extremely useful in final method development, since it allows the user to constrain the separation as needed (to remove or reduce the concentration of one solvent, for example), and predict the resulting separation before actually running further experiments. Further examples of this will be shown below.

The last program is COLPLP, the plotting software to view the results of retention prediction. Maps of retention time, alpha or resolution for any two peaks, or alternatively, a full ORM for all peaks, can be viewed on any alphanumeric terminal, using the alphabet code for output of contour maps. Although the plots in this format are somewhat difficult to visualize, they can be performed on a non-graphics terminal and are a useful general device for plotting. An option in the program allows the data to also be written to a file, which can then be read into more sophisticated graphics routines for better plots.

RESULTS AND DISCUSSION

The data generated in (1) now will be used to demonstrate the use of the programs of Appendix I. In the original report, k' and sigma values were listed. More typically, however, retention times and peak widths are the primary measurements gathered in experiments. Therefore, these data are listed in Table 1 for all nine compounds in the seven solvent mixtures needed for optimization. This information is entered into COLDG along with information on flow rate, an estimate of column plate count (e.g., N=10,000 here), column dead-time (t_0=0.85 min) and descriptive text. If the peak sigma values are unknown, COLDG calculates these for each peak, based on an assumed N=10,000 column, and the peak retention time t_R from the relationship:

$$\text{sigma} = (t_R^2/N) \qquad (1)$$

Once the regression analysis is performed with COLRG, the retention of a peak can be mapped as a function of any solvent over the entire solvent selectivity triangle. Figures 1 and 2

RETENTION TIME DATA FOR SUBSTITUTED NAPHTHALENES ON ZORBAX-C_8 COLUMN

Key: $\overline{\text{MeOH}}$ = methanol-water (63:37): $\overline{\text{THF}}$ = tetrahydrofuran-water (39:61); $\overline{\text{ACN}}$ = acentonitrile-water (52:48).

Compound No. •	Solvent; retention time						
	$\overline{\text{MeOH}}$	$\overline{\text{MeOH}}$-$\overline{\text{THF}}$ (50:50)	$\overline{\text{THF}}$	$\overline{\text{THF}}$-$\overline{\text{ACN}}$ (50:50)	$\overline{\text{ACN}}$	$\overline{\text{ACN}}$-$\overline{\text{MeOH}}$(50:50)	$\overline{\text{MeOH}}$--$\overline{\text{ACN}}$-$\overline{\text{THF}}$(33:33:33)
1	1.49	1.46	1.38	1.35	1.33	1.54	1.44
2	1.61	1.58	1.74	1.66	1.80	1.89	1.67
3	1.99	2.98	3.04	2.42	1.86	2.15	2.55
4	2.93	2.98	3.04	2.82	2.99	3.28	2.99
5	3.62	4.70	4.27	3.87	3.78	4.18	4.28
6	4.53	5.51	4.96	4.49	4.39	4.92	5.06
7	4.53	5.88	5.46	4.86	4.52	4.92	5.34
8	6.90	8.53	6.80	6.44	6.26	7.47	7.73
9	7.89	9.54	6.80	7.27	7.02	8.41	8.89

• Compounds as In Table III, Ref. 1

```
B6
VERSUS
HORIZONTAL U AXIS THF
VERTICAL V AXIS   MeOH
SLUFF HORIZONTAL  ACN
                         V 100.00%         BAND MIDPOINT VALUE
                        /B\                        A    8.87
                       /CCC\                       B    9.10
                      /CDDDE\                      C    9.32
                     /DDDEEEF\                     D    9.55
                    /DDEEEFFFG\                    E    9.78
                   *EEEEFFFGGGH*                   F   10.01
                  /EEEFFFGGGHHHI\                  G   10.24
                 /EEFFFFGGGHHHIII\                 H   10.46
                /EEFFFFGGGGHHHIIIJ\                I   10.69
               /EEFFFFGGGGHHHHIIIJJ\               J   10.92
   0.00% U *EEFFFFFGGGGHHHHIIIJJJ*    0.00% W      K   11.15
             /EEEFFFFFGGGGHHHHIIIJJJK\             L   11.37
            /EEEEEFFFFFGGGGHHHHIIIJJJJ\            M   11.60
           /EEEEEEEFFFFFGGGGGHHHIIIIJJJ\           N   11.83
          /DDEEEEEEEEEFFFFFGGGGGHHHIIIIJJ\         O   12.06
         *DDDDDDDEEEEEEEEFFFFFGGGGHHHHIIIJ*        P   12.28
        /CCDDDDDDDDDDEEEEEEEFFFFFGGGGHHHIII\       Q   12.51
       /CCCCCCCCCCCDDDDDDDEEEEEEFFFFFGGGHHHH\      R   12.74
      /BBBBBBBBBBCCCCCCCCDDDDDDEEEEEEFFFFGGGH\     S   12.97
     /AAAAAAAABBBBBBBBBBBCCCCCCDDDDDEEEEEFFFFG\    T   13.19
    /AAAAAAAAAAAAAAAAAAAAAABBBBBBCCCCCDDDDEEEEF\   U   13.42
W- - - - - - -*- - - - - -*- - - - - -*- - - - - - -U
100.00%               0.00% V                100.00%
RETURN FOR MAX INFO


ABS MAX     11.0317
 PEAKS       6    0
COLS,COMPS  1.0000 0.4500 0.0000 0.5500
```

Figure 1 Retention time map for peak 6 in substituted naphthalene mixture. Conditions as in Table 1.

show the actual alphanumeric output from the program for peaks 6 and 7 from Reference 1. In these presentations, the retention time of the peak is identified over the solvent selectivity triangle with the modifier apeces of V=methanol (MEOH), W=acetonitrile (ACN), and U=tetrahydrofuran. (THF). We have found these types of plots to be extremely useful in developing optimum separations, since they generate a simple picture of how a change of variables affects the retention

```
B7
VERSUS
HORIZONTAL U AXIS THF
VERTICAL V AXIS   MeOH
SLUFF HORIZONTAL  ACN
                      V 100.00%         BAND MIDPOINT VALUE
                     /A\                        A   9.18
                    /ABB\                       B   9.47
                   /BBCCC\                      C   9.75
                  /BBCCDDE\                     D  10.03
                 /BCCDDEEEF\                    E  10.32
                *CCCDDEEFFFG*                   F  10.60
               /CCDDDEEFFFGGH\                  G  10.88
              /CCDDEEEFFFGGHHI\                 H  11.17
             /CDDDEEEFFFGGHHHII\                I  11.45
            /CDDDEEEFFFGGGHHIIIJ\               J  11.73
   0.00% U *CDDDEEEEFFFGGGHHIIIJJ*    0.00% W   K  12.02
          /CDDDDEEEFFFGGGHHHIIIJJJ\             L  12.30
         /CCDDDDEEEFFFFGGGHHHIIIJJJ\            M  12.59
        /CCDDDDDEEEFFFFGGGHHHHIIIJJK\           N  12.87
       /CCCCDDDDEEEEFFFFGGGHHHHIIIJJJ\          O  13.15
      *CCCCCDDDDDEEEEFFFFGGGGHHHIIIJJJ*         P  13.44
     /BBCCCCCDDDDDEEEEEFFFFGGGGHHHIIIIJ\        Q  13.72
    /BBBBCCCCCCDDDDDEEEEEFFFFGGGGHHHHIII\       R  14.00
   /BBBBBBBBCCCCCCDDDDDEEEEEFFFFGGGGHHHHI\      S  14.29
  /AAAAABBBBBBBBCCCCCCDDDDDEEEEFFFFFGGGGHH\     T  14.57
 /AAAAAAAAAAABBBBBBBCCCCCCDDDDDEEEEEFFFFGGG\    U  14.85
W- - - - - -*- - - - -*- - - - -*- - - - - -U
100.00%             0.00% V                  100.00%
RETURN FOR MAX INFO


ABS MAX    11.8764
 PEAKS      7    0
COLS,COMPS  1.0000 0.3500 0.0000 0.6500
MORE(1=YES)
1
```

Figure 2 Retention time map for peak 7 in substituted naphthalene mixture. Conditions as in Table 1.

of any compound. If further experiments are to be performed under different conditions, these plots can be used to predict retention before a run is made.

A simple extension of this approach is shown in Figure 3, where the resolution between peaks 6 and 7 is shown in a similar trilinear format. Figure 3 was generated by calculating the resolution between the compounds represented by peaks 6 and 7 at each point within the solvent triangle,

```
B6
B7
VERSUS
HORIZONTAL U AXIS THF
VERTICAL V AXIS   MeOH
SLUFF HORIZONTAL  ACN
                      V 100.00%         BAND MIDPOINT VALUE
                     /A\                        A   0.12
                    /AAB\                       B   0.36
                   /AAABC\                      C   0.60
                  /AAABBCC\                     D   0.84
                 /AAABBCCDD\                    E   1.08
                *AAABBCCDDDE*                   F   1.32
               /AAABBCCDDDEEE\                  G   1.56
              /AAABBCCCDDEEEFF\                 H   1.80
             /AAABBCCCDDEEEFFFF\                I   2.04
            /AAABBCCCDDEEEFFFFGG\               J   2.28
   0.00% U *AABBBCCDDDEEEFFFFGGGG*    0.00% W   K   2.51
          /AABBBCCDDDEEEFFFFGGGGHH\             L   2.75
         /AABBCCCDDDEEEFFFFGGGGHHHH\            M   2.99
        /ABBBCCCDDEEEEFFFGGGGGHHHHHI\           N   3.23
       /ABBCCCDDDEEEFFFFGGGGHHHHHHIII\          O   3.47
      *BBBCCDDDEEEEFFFGGGGGHHHHHIIIIII*         P   3.71
     /BBCCCDDDEEEFFFFGGGGHHHHHIIIIIIIII\        Q   3.95
    /BCCCDDDEEEFFFFGGGGHHHHHIIIIIIIIJJJJ\       R   4.19
   /CCCDDDEEEFFFFGGGGHHHHHIIIIIIIJJJJJJJJ\      S   4.43
  /CCDDDEEEFFFGGGGHHHHHIIIIIIIJJJJJJJJJJJJ\     T   4.67
 /DDDEEEFFFFGGGGHHHHIIIIIIJJJJJJJJJJJJJJJKJ\    U   4.91
W- - - - - -*- - - - -*- - - - -*- - - - - -U
100.00%             0.00% V                  100.00%
RETURN FOR MAX INFO


ABS MAX     2.3947
 PEAKS      6    7
COLS,COMPS  1.0000 0.0000 0.0250 0.9750
MORE(1=YES)
1
```

Figure 3 Resolution map for peaks 6-7.
Conditions as in Table 1.

taking into account peak width data. Thus, the presentation in Figure 3 lists the resolution values for peaks 6 and 7 (as shown by the listing of symbols at the right), and describes (below the plot) the maximum resolution possible for peaks 6 and 7 at position "K" (max. Rs = 2.39). Figure 3 also predicts the solvent mixture providing this maximum resolution (97.5% of the THF-water mixture, 2.5% of the ACN-water mixture and 0% MEOH).

Alternatively, the software allows an alpha value plot to be made when only k' information is needed for the peaks. These simple resolution plots can be used when only one specific separation is important in a complex mixture, i.e., when the separation of all components in a mixture is not the optimization goal. Resolution (or alpha) maps can be created only for those peaks of interest (or for one peak from all others). It is likely that the best conditions for this separation will not be the same as the optimal conditions for separating all peaks from each other.

Figure 4 shows the ORM for the mixture, generated from the computer overlap of all 36 resolution maps for the nine peaks in Table I. The presentation is in the same format as Figure 3. Maximum resolution for the constraining peaks 6 and 7 (most difficult to resolve) is again at symbol "K", with a maximum resolution of 2.22 for 65% of the THF-water mixture and 35% of the ACN-water mixture (no MEOH). Although this type of information is often the final goal of an optimized procedure, and this type of plot is useful for defining a specific optimum, the utility of the individual retention time and resolution maps is often overlooked. In particular, we find that these individual "maps" quite useful when slight variations in the separation occur with time. For example, when a new column or new lot of solvent are required, these new variables can change the retention of certain compounds sufficiently to degrade the overall separation. The individual resolution maps quickly indicate the direction for a small solvent adjustment to improve the separation, without the need for performing a totally new optimization.

The solvent optimization example described above was for the simplest case of an isocratic separation with one column. The mixture-design software is capable of incorporating other

```
MIN PLOT
VERSUS
HORIZONTAL U AXIS THF
VERTICAL V AXIS   MeOH
SLUFF HORIZONTAL  ACN
                      V 100.00%                 BAND MIDPOINT VALUE
                     /A\                                 A    0.11
                    /AAB\                                B    0.33
                   /AABBC\                               C    0.56
                  /AAABCCD\                              D    0.78
                 /AAABBCCDD\                             E    1.00
                *AAABBCCDDEE*                            F    1.22
               /AAABBCCDDEEEF\                           G    1.44
              /AAABBCCDDEEEFFF\                          H    1.67
             /AAABBCCDDEEEFFFGE\                         I    1.89
            /AAABBCCDDEEEFFFGFDB\                        J    2.11
   0.00% U *AABBCCCDDEEEFFFGGFDBA*    0.00% W            K    2.33
          /AABBCCCDDEEEFFFGGGFDBAB\                      L    2.55
         /AABBCCDDDEEEFFFGGGGFDBABC\                     M    2.78
        /ABBCCCDDDEEFFFFGGGHGFDCAACD\                    N    3.00
       /ABBCCDDDEEEFFFGGGGHHGFDCAABDE\                   O    3.22
      *BBCCCDDEEEFFFFGGGHHHHHFECBAACDE*                  P    3.44
     /BBCCDDDEEEFFFGGGGHHHHIHGEDCBABCDE\                 Q    3.66
    /BCCDDDEEEFFFGGGGHHHHIIHGFEDBAABBCD\                 R    3.88
   /CCDDDEEEFFFGGGGHHHHIIIIIIHGFDCCBAABCC\               S    4.11
  /CDDDEEEFFFGGGHHHHIIIIIJJJJIHGFEDCBAAAAB\              T    4.33
 /DDEEEFFFGGGGHHHIIIIIJJJJJKJIHGFEDDCCBAAA\              U    4.55
W- - - - - - -*- - - - -*- - - - -*- - - - - - -U
100.00%            0.00% V                    100.00%
RETURN FOR MAX INFO


ABS MAX      2.2203
 PEAKS       7    6
COLS,COMPS   1.0000 0.0000 0.3500 0.6500
MORE(1=YES)
0
```

Figure 4 Overlapping Resolution Map (ORM) for all compounds in substituted naphthalene mixture. Conditions as in Table 1.

variables such as multiple columns, pH or temperature, and gradient elution separations. When utilizing multiple columns, it is assumed that the retention behavior of any compound is a linear function of column type. In other words, if a compound has a k' = 5.0 in 60% methanol/water with an octyl bonded-phase column, and a k'= 6.2 in 60% methanol/water with a cyano bonded-phase column, the compound is assumed to have a k' = (5.0+6.2)/2= 5.6 in 60% methanol/water with a total column consisting of 50% octyl and 50% cyano-bonded phase packings. This approach simplifies the experimental data collection, since data needs only to be collected for the individual columns, and predictions can then be made for any mixed bed.

An added advantage of using multiple stationary phases with multiple solvents is that once the data is gathered, the user can choose to limit the optimization to only a subset of the total variables investigated. For example, if one solvent component is undesirable (such as tetrahydrofuran for trace work at low UV absorbance), the optimum can be calculated for all columns, with only acetonitrile and methanol used as mobile-phase modifiers. Although the local optimum for these solvents may not provide as good a separation as the global optimum with all solvents, the program does allow the prediction of the maximum resolution with a more practical set of operating conditions. A further discussion of this type of optimization can be found in Reference (5).

Previous work (8) has demonstrated that the approaches and software described above also can be used to optimize gradient elution separations, provided the gradient slope (change in solvent strength as a function of time) is constant for all mobile phase systems. The only difference is that the peak width is assumed to be relatively constant over the gradient separation when analyzing the data with COLDG, COLRG, and COLGRD. Therefore, instead of calculating as in Equation 1, peak sigma values are assumed to be constant, and separation is measured as a simple function of retention time difference. As is the case in optimizing isocratic systems, if certain compounds have peak widths that are significantly different from the average, specific peak-width values can be designated.

```
MIN PLOT
VERSUS
HORIZONTAL U AXIS pH
VERTICAL V AXIS   IIR
            *----*----*----*----*----*----*----* MID-BAND VALUES
0.50000E+01*,AFK,.F,AA,FF.KK,,PP..UUU,,,,UUU...P* A= 0.10160E+01
           1,AF.,KF,AAA,F..K,,PP...UUU,,,,UUU...1 ,= 0.10478E+01
           1,AF.,K.F,AA,FF.KK,,PP..UUUU,,,,UUU..1 F= 0.10796E+01
           1,AF.KK.F,AA,,F..KK,PPP..UUUU,,,,UUU.1 .= 0.11114E+01
           1,A,.K,.F,AAA,FF..K,,PP...UUUU,,,,UUU1 K= 0.11431E+01
0.37500E+01*,A,FK,K.F,AAA,FF.KK,,PP...UUUU,,,,UU* ,= 0.11749E+01
           1,A,F.KK.F,,AA,,FF.KK,,PPP..UUUU,,,UU1 P= 0.12067E+01
           1,AA,.K,K.F,AAA,,FF.KK,,PPP..UUUU,,UU1 .= 0.12385E+01
           1FAA,F.K,K.F,AAA,,FF..K,,,PP...UUU,UU1 U= 0.12702E+01
           1F,AA,F.K,K.F,,AAA,FF..KK,,PP...UUUUU1 ,= 0.13020E+01
0.25000E+01*.FAAA,F.K,K.F,,AAA,,FF.KK,,PPP..UUUU* Z= 0.13338E+01
           1.F,AA,,F.K,K.FF,AAAA,FF..KK,,PP..UUU1
           1K.F,AAA,F.KK,K.FF,AAA,,FF..KK,PP...U1
           1,K.F,,AA,,F.KKKK.FF,AAA,,FF..K,,P,.,1
           1P,K..F,AAA,,F..K,K..F,AAAA,FF.,AFKP.1
0.12500E+01*FK,,K..F,,AAA,,F.KK,K.F,AAAA,FF.K,,P*
           1UU.P,,K...FF,AAAA,FF.K,K.FF,AAA,FAF,1
           1UUUUU..PP,,KK.FF,,AAA,,F.K,K,,,AA,FA1
           1UU,,,,,,,,UU..PP,,K..F,,AA,,F,A.F,AA1
           1..UUUUUUUU,,,,Z,.,FA,.PK,A.K.F,AA.AF1
0.00000E+00*,,,,,,,KKKKKK.,AFKKFA,A.K..F,AAAAAAA*
            *----*----*----*----*----*----*----*
           0.36000E+01                    0.60000E+01
RETURN FOR MAX INFO


ABS MAX     1.3179
 PEAKS      6    3
COLS,COMPS  1.0000 1.0000 4.5600 0.2500
MORE(1=YES)
0
```

Figure 5 Minimum alpha plot for separation of disubstituted anilines. Data from Ref. 21.

Other unrelated variables, such as temperature and pH, also can be used with the mixture-design software. An example is taken from the data of Deming et al (22), where the effect of pH and ionic strength on the separation of di-substituted anilines was investigated. Figure 5 shows a minimum alpha plot obtained with data from this work. These results, using only a quadratic fit to the data, are quite similar to those reported in the original work. In this case, a rectangular plot is shown, rather than trilinear, because the variables in this experiment are not related, as discussed above. The mixture-design method software also will accommodate mixtures of related and discrete variables, for example, a four-solvent optimization carried out at two different temperatures.

CONCLUSIONS

Software for mixture-design statistical optimization of variables in HPLC has been detailed. Using an example of an isocratic elution on a single column, the mapping of retention time, capacity factor (k'), resolution, alpha, and overlapping resolution (ORM) or alpha values was illustrated with these programs. Extension of this software to variables other than mobile-phase modifiers has been described, and example plots of this utility have been created with published retention data. The source code for the mixture-design statistical software is described in Appendix I. This software is directly applicable to a VAX-11/780 or equivalent machine.

REFERENCES

1. J. L. Glajch, J. J. Kirkland, K. M. Squire, and J. M. Minor, J. Chromatogr., 199 (1980) 57.
2. L. R. Snyder, J. L. Glajch, and J. J. Kirkland, J. Chromatogr., 218 (1981) 299.
3. J. L. Glajch, J. J. Kirkland, and L. R. Snyder, J. Chromatogr., 238 (1982) 269.
4. P. E. Antle, Chromatographia, 15 (1982) 277.
5. J. L. Glajch, J. C. Gluckman, J. G. Charikofsky, J. M. Minor, and J. J. Kirkland, J. Chromatogr., 318 (1985) 25.
6. A. P. Goldberg, E. L. Nowakowska, P. E. Antle, and L. R. Snyder, J. Chromatogr., 316 (1984) 241.
7. J. L. Glajch and J. J. Kirkland, Anal. Chem., 54 (1982) 2593.

8. J. J. Kirkland and J. L. Glajch, J. Chromatogr., 255 (1983) 27.

9. G. M. Landers and J. A. Olsen, J. Chromatogr., 291 (1984) 51.

10. G. D'Agostino, F. Mitchell, L. Castagnetta, and M. J. O'Hare, J. Chromatogr., 305 (1984) 13.

11. T. Tsuneyoshi, A. Kawamoto, and J. Koezuka, Clin. Chem., 30 (1984) 1889.

12. H. K. Smith, W. L. Switzer, G. W. Martin, S. A. Benezra, W. P. Wilson, and D. W. Dixon, J. Chromatogr. Sci., 24 (1986) 70.

13. M. de Smet, G. Hoogewijs, M. Puttemans, and D. L. Massart, Anal. Chem., 56 (1984) 2662.

14. H. J. Issaq, J. Klose, K. McNitt, J. E. Haky, and G. M. Muschik, J. Liquid Chromatogr., 4 (1981) 2091.

15. L. R. Snyder, J. W. Dolan, and M. P. Rigney, LC/GC, 4(9) (1986) 921.

16. M. A. Quarry, R. L. Grob, and L. R. Snyder, Anal. Chem., 58 (1986) 907.

17. L. R. Snyder, J. Chromatogr. Sci., 16 (1978) 223.

18. S. R. Bakalyar, R. McIlwrick, and E. Roggendorf, J. Chromatogr., 142 (1977) 353.

19. P. J. Schoenmakers, H. A. H. Billiet, and L. de Galan, J. Chromatogr., 185 (1979) 179.

20. J. F. K. Huber, R. Hirz, and P. Markl, 5th International Symposium on Column Liquid Chromatography, Avignon, France, 1981, Paper C20.

21. J. R. Gant, J. W. Dolan, and L. R. Snyder, J. Chromatogr., 185 (1979) 153.

22. B. Sachok, R. C. Kong, and S. N. Deming, J. Chromatogr., 199 (1980) 307.

23. R. D. Snee, CHEMTECH, 9 (1979) 702.

APPENDIX I

Programs Used For Input, Calculations, and Output of Data in the Mixture-Design Statistical Technique

COLDG

Data input for further programs. Includes text to describe system and information on column, mobile phase, and other variables. Retention times are required for each peak as a minimum amount of information. Other data, such as peak width, k', and retention volume can be entered or calculated in the program. User-specified output file stores data for further calculations in other programs.

COLRG

Data regression program. Reads file generated by COLDG and requests further text information. Regression coefficients are generated using quadratic or cubic fit (user choice). Coefficients, along with estimate or residual error for each peak, are stored in a user-specified file.

COLGRD

Program to predict retention of any or all peaks for any set of conditions bounded by the experimental measurements. Normally used to predict the optimum conditions for final run, although this program can predict retention and resultant separations under non-optimum conditions. Data is taken from files specified in COLDG and COLRG.

COLPLP

Program to generate output plots using the same algorithms of COLGRD. Data can be output in alphanumeric plot form, as in Figures 1-5 of this work, or sent to a data file for manipulation and plotting by more sophisticated graphics programs.

COLLIB and COLLB1

Libraries containing various subroutines used by all of the above programs for data input, regression, prediction, and plotting.