International Journal of Pharmaceutics, 23 (1985) 195–207
Elsevier
IJP 00788

Statistical techniques applied to solubility predictions and pharmaceutical formulations: an approach to problem solving using mixture response surface methodology

R.J. Belloto, Jr. [1], A.M. Dean [2], M.A. Moustafa [3], A.M. Molokhia [3], M.W. Gouda [3] and T.D. Sokoloski [1]

[1] College of Pharmacy, Ohio State University, Columbus, OH; [2] College of Mathematics and Physical Sciences, Ohio State University, Columbus, OH (U.S.A.) and [3] College of Pharmacy, King Saud University, Riyadh (Saudi Arabia)

(Received June 18th, 1984)
(Modified version received September 28th, 1984)
(Accepted October 2nd, 1984)

Summary

Mixture response surface methodology is a group of statistical methods which can generate an empirical equation that can be used to quantitatively define the relationship between some response, such as solubility, and the composition of a system, as for example, different solvent blends. The equation can be used to predict response at any proposed or desired mixture. The term response should be viewed in a very generalized sense to include any property that is affected solely by mixture composition and might include many measurable responses that are of interest to pharmacists: cost, half-life, taste, color, tablet hardness, bioavailability, extraction efficiency, chromatographic resolution, and so on. The method is in no way limited by the number of components in the mixture and thus should be viewed as being general in this sense also. The advantages of mixture response surface methodology and the mechanics involved in its use are illustrated through a prediction of the solubility of diazepam and phenobarbital in solvent blends. The approach is entirely empirical. It is based on rigorous statistical design and data analysis and can lead to excellent prediction of solubility. It also is shown how several responses can be

Correspondence: T.D. Sokoloski, College of Pharmacy, Ohio State University, Columbus, OH 43210-1291, U.S.A.

simultaneously analyzed to identify mixtures that strike compromises between maximization of some response (e.g. solubility) and minimization of another (e.g. cost).

Introduction

Formulators of drugs in oral and/or parenteral liquid dosage forms are often challenged with the problem of finding the right blend of solvents which would allow adequate solubility, stability, bioavailability, taste and maintain a reasonable cost. Trial and error approaches are inefficient and costly and extrapolations made from them can be inaccurate. Systematic experimental designs and rigorous data treatment, on the other hand, can lead to the desired information. This study demonstrates how one such method can be applied with obvious increased accuracy and experimental convenience.

As far as solubility is concerned, various data treatments have been applied to the results of solubility experiments and different approaches have been tested for their potential in predicting solubility based on a limited number of data points. These approaches have met with varying degrees of success in several drug systems. For example, Yalkowsky et al. (1972) introduced the log linear equation which relates solubility to the composition of mixed solvent in binary aqueous systems. Subsequently, Yalkowsky and Flynn (1974) found it necessary to expand a term in their equation into a fifth degree polynomial to account for non-linearity across the range of cosolvent (propylene glycol) composition in a study of the solubility of *p*-aminoacetophenone. Martin et al. (1980) and Adjei et al. (1980) on the other hand, introduced an approach to estimate the solubility of drugs in mixed and pure solvent systems based on an extension of the Hildebrand-Scatchard equation of regular solution theory. Martin et al. (1982) later showed that the log linear solubility equation of Yalkowsky et al. (1972) may be derived from the extended Hildebrand approach, thus providing a semi-theoretical foundation for the equation based on solubility parameters.

More recently, Williams and Amidon (1984a) introduced an excess free energy approach to the estimation of solubility. In this approach, an expression for the excess Gibbs free energy of mixing for multicomponent solvent systems was used to obtain parameters characteristic of the interaction between the solvents. An equation was derived to predict the solubility of a solute in a co-solvent system and contains terms for solubility in pure solvents, solvent–solvent interaction contributions and contributions from the solute–mixed solvent interactions. They applied this approach to solubility predictions in ethanol–water (1984b) and ethanol–propylene glycol–water mixtures (1984c). Prediction of solubility using this approach in ternary solvent systems and that of others for binary solvents in part involves an estimation of interaction parameters by statistical regression of experimental solubilities and hence involve fitting equations to data.

In a previous report, the solubility of phenobarbital in propylene glycol–glycerol–water systems was extensively studied (Moustafa et al., 1981). An equation was developed for the prediction of phenobarbital solubility in ternary solvent blends. The equation is an extension of the work of Yalkowsky et al. (1972) and simply includes a higher-order term to account for interactions in the ternary solvent system. The equations takes the form:

$$\log S_t = \log S_0 + \alpha_1 f_1 + \alpha_2 f_2 + \beta f_1^2 f_2 \quad (1)$$

where α_1 and α_2 are constants for propylene glycol and glycerol, respectively and f_1 and f_2 are the volume fractions of both solvents. The term $\beta f_1^2 f_2$ was determined from the fitting of some experimental solubility data to Eqn. 1; it presumably accounts for the increase in solubility due to solvent interactions.

The present report describes the application of a statistical approach for dealing with mixtures that had earlier been introduced by Scheffe (1958). The usefulness of the approach to solubility predictions is demonstrated using the results of solubility determinations carried out on diazepam as well as to the previously published data set for phenobarbital solubility which was further enlarged for the purposes of the present study. In the statistical approach, a response such as solubility is measured as a function of mixture proportions (e.g. solvent blend composition). The response–mixture relationship can be described by a q-dimensional surface in space, where q is the number of components in the mixture. This surface can, in turn, be described by a polynomial which is generated to fit the experimental results and subsequently can be used for solubility prediction. Advantages of mixture response surface methodology include better solubility prediction, reduction in the amount of needed experimentation through application of statistical principles, extension to any number of components in a mixture, measuring several responses for simultaneous assessment (stability, cost, etc.) and provision for easy visualization of imposed constraints in the formulation (e.g. blends available where one component cannot exceed a fixed amount).

Experimental

Diazepam solubility

The solubility of diazepam [1] was determined in the 10 mixtures of ethanol 'USP' [2], propylene glycol [3] and demineralized double-distilled water given in Table 1. All chemicals were used as received. Approximately 20 ml of each mixture were placed in a volumetric flask, excess diazepam and a magnetic stir bar were added and the flask sealed and immersed in a jacketed beaker which was attached to a

[1] Diazepam Lot 709060, Hoffmann-LaRoche, Nutley, NJ 07110, U.S.A.
[2] Pure Ethyl Alcohol U.S.P, U.S. Industrial Chemicals, Tuscola, IL 61953, U.S.A.
[3] Propylene Glycol, Lot 795759, Fisher Scientific, Fair Lawn, NJ 07410, U.S.A.

TABLE 1

EXPERIMENTAL AND PREDICTED SOLUBILITIES OF DIAZEPAM AT 20°C IN VARIOUS MIXTURES OF ETHANOL, PROPYLENE GLYCOL AND WATER

Mixture number	Volume fraction of solvent			Solubility (mg/ml)	
	Ethanol	Propylene glycol	Water	Experimental [a] (number trials)	Predicted [b]
1	1.00	0	0	27.8 (2)	29.7
2	0.66	0.17	0.17	28.0 (3)	25.6
3	0.50	0.50	0	27.0 (3)	32.1
4	0.50	0	0.50	6.02 (3)	5.07
5	0.33	0.33	0.33	9.52 (2)	7.81
6	0.17	0.66	0.17	13.0 (2)	7.82
7	0.17	0.17	0.66	0.408 (4)	0.687
8	0	1.00	0	7.42 (4)	8.10
9	0	0.50	0.50	0.610 (2)	0.530
10	0	0	1.0	0.0479 (2)	0.0347

[a] Average of 2–4 determinations.
[b] Using Eqn. 4.

circulating water bath [4]. The temperature was kept at 20°C. Following the procedure of Saleh et al. (1980) the mixture was stirred for 24 h, allowed to equilibrate for another 24 h, filtered and an aliquot of the solution appropriately diluted for spectrophotometric [5] assay. Absorbance was measured at 230 nm and concentrations were determined by referring to individual calibration curves for freshly prepared solutions of diazepam in the same solvent blends. No degradation of diazepam was detected in these solutions during the time of assay.

Phenobarbital solubility

The solubility data used in this report are those determined by Moustafa et al. (1981). Eight more data points were collected from further experiments carried out to test predictability at solvent volume fractions outside that covered by the original set of data. The composition of solvents and solubility values determined for phenobarbital in these mixtures of propylene glycol, glycerol and water are given in Table 2. In each experiment, approximately 30 ml of every mixture were placed in a volumetric flask, excess phenobarbital was added, the flask sealed and shaken in a thermostatically controlled water bath. The temperature was kept at 32°C and control was better than ± 0.2°C. Samples of the solution were pipetted, filtered [6] diluted with the same solvent mixture and assayed spectrophotometrically [7] (Higuchi and Brochmann-Hanson, 1961) for phenobarbital.

[4] Haake, Constant Temperature Circulator, Model FK, Haake Instruments, 244 Saddle River Rd., Rochelle Park, NJ 07662, U.S.A.
[5] Beckman Spectrophotometer UV, Model 5260, Beckman Instruments, Fullerton, CA 92634, U.S.A.
[6] Membrane filter, 0.8 μm, Millipore, Bedford, MA, U.S.A.
[7] Pye-Unicam SP 600, Cambridge, U.K.

TABLE 2

EXPERIMENTAL AND PREDICTED SOLUBILITIES OF PHENOBARBITAL AT 32°C IN VARIOUS MIXTURES OF GLYCEROL, PROPYLENE GLYCOL AND WATER

Mixture	Volume fraction of solvent			Solubility (mg/ml)		
	Glycerol	Propylene glycol	Water	Experimental	Predicted using:	
					Eqn. 5	Eqn. 7
1	0.3915	0.5000	0.1085	45.54	50.87	45.91
2	0.7395	0.1000	0.1605	10.96	12.23	11.64
3	0.4350	0.1000	0.4650	4.81	5.25	5.70
4	0.2175	0.2000	0.5825	4.68	5.08	5.47
5	0.0000	0.0000	1.0000	1.29	1.39	1.28
6	0.6960	0.0000	0.3040	5.44	5.86	6.07
7	0.6525	0.1000	0.2475	8.90	9.45	9.52
8	0.5655	0.3000	0.1345	24.36	25.80	23.73
9	0.5220	0.3000	0.1780	21.16	22.08	21.09
10	0.5655	0.0000	0.4345	4.09	4.26	4.59
11	0.5220	0.2000	0.2780	11.32	11.76	12.04
12	0.1740	0.2000	0.6260	4.42	4.58	4.88
13	0.1740	0.3000	0.5260	7.28	7.53	8.03
14	0.3045	0.2000	0.4955	6.14	6.33	6.88
15	0.3915	0.1000	0.5085	4.57	4.71	5.14
16	0.0870	0.2000	0.7130	3.67	3.77	3.87
17	0.0000	0.2000	0.8000	3.08	3.15	3.07
18	0.2610	0.2000	0.5390	5.54	5.66	6.14
19	0.0870	0.3000	0.6130	5.91	6.04	6.27
20	0.1305	0.2000	0.6695	4.09	4.15	4.35
21	0.0000	0.3000	0.7000	4.89	4.94	4.88
22	0.3480	0.5000	0.1520	42.24	42.64	40.06
23	0.3915	0.3000	0.3085	14.10	14.22	14.74
24	0.5655	0.2000	0.2345	13.39	13.47	13.44
25	0.3480	0.2000	0.4520	7.09	7.16	7.70
26	0.2610	0.1000	0.6390	3.47	3.47	3.75
27	0.0000	0.5000	0.5000	13.11	13.10	13.20
28	0.0000	0.1000	0.9000	2.08	2.07	1.96
29	0.2610	0.0000	0.7390	2.26	2.21	2.34
30	0.2610	0.5000	0.2390	31.50	30.55	30.45
31	0.1740	0.5000	0.3260	23.22	22.45	23.10
32	0.7830	0.0000	0.2170	7.61	7.34	7.29
33	0.5220	0.1000	0.3780	6.87	6.57	7.01
34	0.2610	0.3000	0.4390	10.03	9.57	10.25
35	0.4350	0.0000	0.5650	3.39	3.17	3.45
36	0.4350	0.2000	0.3650	9.76	9.07	9.64
37	0.1305	0.0000	0.8695	1.87	1.74	1.74
38	0.0870	0.5000	0.4130	18.25	16.93	17.48
39	0.1740	0.1000	0.7260	3.11	2.88	3.03
40	0.0870	0.1000	0.8130	2.64	2.43	2.44
41	0.6090	0.2000	0.1910	17.86	15.50	15.00
42	0.6525	0.2000	0.1475	22.76	17.90	16.73

Results and Discussion

The observed solubility values of diazepam in various mixtures of ethanol (volume fraction x_1), propylene glycol (volume fraction x_2) and water (volume fraction x_3) are shown in Table 1. Observed solubility values of phenobarbital in various mixtures of glycerol, (x_4), propylene glycol, (x_2), and water, (x_3) are given in Table 2. Following the flow chart of Draper and St. John (1977), different models were examined relating solubility, S, to the volume fractions of the three solvents in the mixtures tested. These ranged from a simple equation were no product terms such as $x_1 \times x_2$ were considered (Eqn. 2) to models that were more complex (Eqn. 3)

$$S = \sum_{i=1}^{3} \beta_i x_i \tag{2}$$

$$S = \sum_{i=1}^{3} \beta_i x_i + \sum_{i<j}^{3} \beta_{ij} x_i x_j + \sum_{i=1}^{3} \beta_{-i} x_i^{-1} \tag{3}$$

In these equations, S is the solubility, x_i is the volume fraction of component i and β is the coefficient to be fitted.

Least-squares fitting [8] of the data to such equations resulted in the prediction of solubilities having negative values. This, of course, has no physical meaning. To accommodate this sign dilemma and since the addition of a co-solvent generally increases the solubility in a logarithmic fashion, the natural log of the solubility was regressed against the volume fraction of the components, again by the method of least-squares. Analysis indicated that the logarithmic transformation was a more appropriate choice for fitting the data. This is in agreement with Draper and Smith (1981) who point out that transformation of the data sometimes enables the investigator to fit a lower degree polynomial than would otherwise be required and it is always more convenient to manipulate a lower-order equation. The simplest equation describing the relationship between log of solubility and solvent composition was identified based upon established statistical methods. The criteria for best fit were considerations of correlation coefficient, residual pattern, sum of squared residuals among the various models used and a knowledge of the usual physical behavior of such systems. Terms were assumed insignificant if the standard error of a coefficient was more than twice the absolute value of the coefficient itself.

The equation that best fit the diazepam and phenobarbital data was a reduced form of the special cubic equation of Scheffe (1958), Eqns. 4 and 5, respectively. The correlation coefficient used was an adjusted R-squared (R_A^2) term; it differs from the ordinary R^2 by taking into account the number of parameters needed to fit the data (Marquardt and Snee, 1974). The adjusted value is helpful when comparing different

[8] All fitting was done using the Statistical Analysis System (SAS), SAS Institute, P.O. Box 10066, Raleigh, NC 27605, U.S.A.

TABLE 3

COEFFICIENTS AND STANDARD ERRORS (S.E.) OF THE EQUATIONS BEST FITTING SOLUBILITY TO COMPONENT FRACTION FOR DIAZEPAM AND PHENOBARBITAL

Diazepam			Phenobarbital				
Term	Eqn. 4		Term	Eqn. 5		Eqn. 7	
	Coefficient	S.E.		Coefficient	S.E.	Coefficient	S.E.
x_1	3.39	0.221	x_4	2.60	0.120	2.44	0.079
x_2	2.09	0.152	x_2	5.47	0.293	5.43	0.255
x_3	−3.36	0.192	x_3	0.330	0.055	0.250	0.040
x_1x_2	2.91	0.093	x_4x_2	1.45	0.919	0.934	0.516
x_1x_3	6.44	0.981	x_4x_3	−0.656	0.393	0.151	0.264
$x_1x_2x_3$	8.37	6.06	x_2x_3	−1.31	0.640	−1.03	0.496
			$x_4x_2x_3$	−2.07	2.43	–	–

models using the same data set and moreover, Marquardt and Snee showed that the usual R-squared is not appropriate as it always results in a very misleading high value.

$$\ln(S) = 3.39x_1 + 2.09x_2 - 3.36x_3 + 2.91x_1x_2 + 6.44x_1x_3 + 8.37x_1x_2x_3 \quad (4)$$

$$\ln(S) = 2.60x_4 + 5.47x_2 + 0.33x_3 + 1.45x_4x_2 - 0.656x_4x_3 - 1.31x_2x_3 - 2.07x_4x_2x_3 \quad (5)$$

Values of R^2_A were 0.973 and 0.993 for Eqns. 4 and 5, respectively. These equations can be used to predict the solubility of either drug in any mixture of components. Table 3 gives the standard errors of the coefficients in Eqns. 4 and 5. Tables 1 and 2 list the solubilities predicted using Eqns. 4 and 5. Fig. 1A is a plot of Eqn. 4 that attempts to represent the solubility surface for diazepam in three-dimensional space ($q = 3$). Figures such as 1A enable one to determine whether it is possible to formulate a solution having the desired concentration at a particular combination of solvents. Any concentration above the solubility surface at a particular component mixture represents a supersaturated solution since the surface defines the solubility limits at a particular temperature. Fig. 1B is a contour plot of the same equation where the view is that of looking down along the response axis. Each plateau or line represents a constant solubility. The dotted lines are negative logarithm values.

Another potential response of interest to a formulator is the cost involved in the use of a particular mixture. In generating a cost surface for the diazepam case, the coefficients of the equation relating cost, C (in dollars/ml), to fraction of component in a mixture are given by Eqn. 6 where x_1, x_2, and x_3 again refer to ethanol, propylene glycol and water volume fractions, respectively [9].

[9] The cost of solvents used was obtained from the Ohio State University Laboratory Stores.

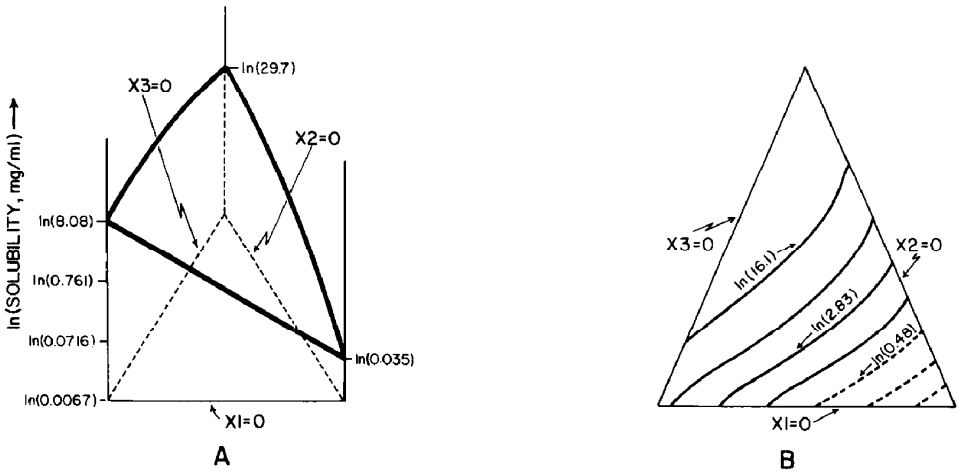

Fig. 1. A: a three-dimensional representation of Eqn. 4 which describes the response surface (solubility) as a function of solvent fraction. B: contour plot of Eqn. 4 where the view is that of looking down along the response (solubility) axis. Each line or plateau represents a constant natural logarithm of the solubility; the dotted lines are negative logarithmic values. In both A and B, x_1 is the fraction of ethanol, x_2 is the fraction of propylene glycol, and x_3 the fraction of water.

$$C = 4.67 \times 10^{-3}x_1 + 4.03 \times 10^{-3}x_2 + 1.34 \times 10^{-4}x_3 \quad (6)$$

Plots of Eqn. 6 are given in Fig. 2A and B. Fig. 2A is a three-dimensional representation of the data showing the cost response surface. It is a plane in

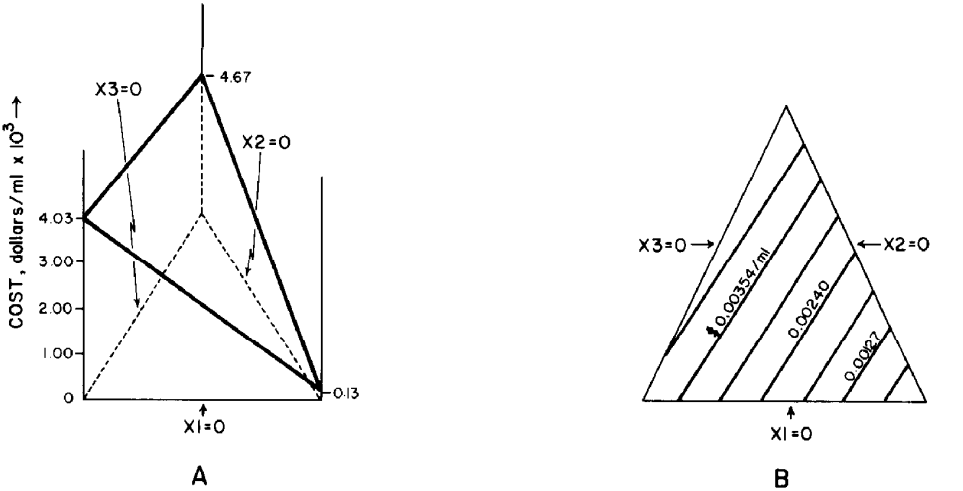

Fig. 2. A: a three-dimensional representation of Eqn. 6 which describes the response surface (cost) as a function of solvent fraction. B: contour plot of Eqn. 6 where the view is that of looking down along the response (cost) axis. Each line or plateau represents a constant cost.

three-dimensional space since only first-order terms in composition are used. Fig. 2B gives the corresponding contour plot for Eqn. 6; the view is that of looking down along the cost axis which extends out of the paper. Each line or plateau of constant cost rises from lower right toward upper left.

To illustrate the usefulness of a simultaneous assessment of responses, consider a desire to identify mixtures of the three solvents maximizing diazepam solubility while minimizing cost. One way of doing this is a graphical method which involves superimposing Figs. 1B and 2B. Fig. 3 has the contour plots for solubility and cost superimposed upon one another. Consider the following example for the use of Fig. 3. Suppose a concentration of diazepam of 1.2 mg/ml was desired and not more than 10% ethanol may be used. Fig. 3 shows that the formulations possible would be bound by the solubility line to the right and to the top by a line drawn at the 10% level of ethanol parallel to the x_1 base (see area outlined in Fig. 3). Another method of finding an optimum formulation is computer-based where, with the appropriate software (Derringer and Suich, 1980), the optimum can be calculated, provided all responses of interest have been fitted correctly and all constraints have been considered. As long as there are not too many components, the graphic approach will probably suffice for most applications.

Fig. 4 is a triangular graph depicting the experimental design and shows the 42 phenobarbital solubility data points originally run (open circles). Fig. 5 is a contour plot of Eqn. 5 for phenobarbital solubility and may be interpreted in a manner similar to Fig. 1B for diazepam.

Tables 1 and 2 show the agreement of observed solubility values with those generated by the statistical approach adopted in this study for experimental points in both the diazepam and phenobarbital solubility systems. A comparison of the fit

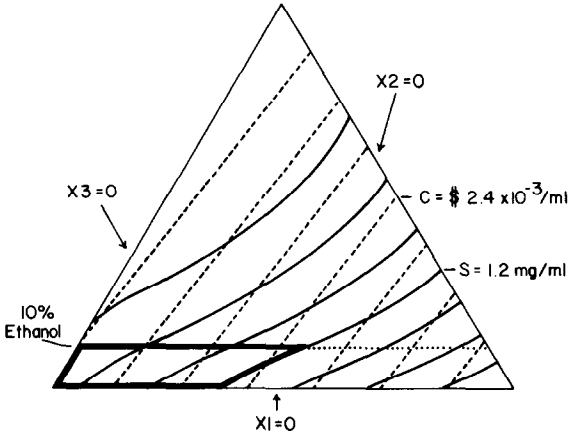

Fig. 3. Superimposition of the natural logarithm solubility surface (solid lines) and the cost surface (dashed lines). The plot is constructed using Figs. 1B and 2B. The volume fractions of ethanol, propylene glycol, and water are x_1, x_2 and x_3, respectively. The figure is used to optimize a formulation with respect to solubility and cost as described in the text under a constraint of there being not more than 10% ethanol in the final product.

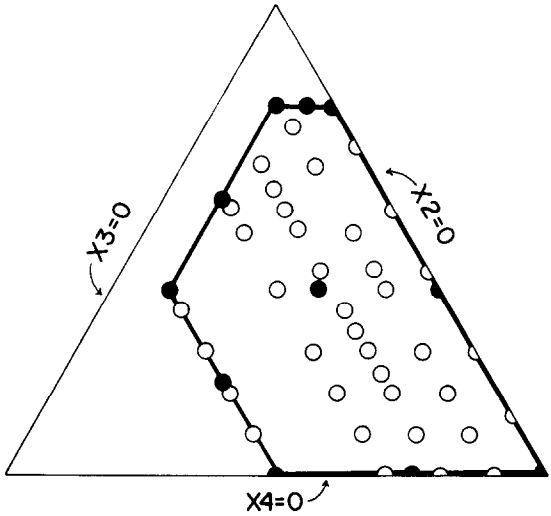

Fig. 4. Triangular graph depicting the experimental design space for phenobarbital solubility in ternary mixtures. Open circles indicate the 42 data points originally run (some closed circles obscure the data points). Closed circles are the data points that are suggested by an extreme vertices selection design as the only ones needed to adequately define the entire solubility surface. x_4 = glycerol; x_2 = propylene glycol; and x_3 = water volume fractions.

to observed solubility values, in the phenobarbital system, by the statistical approach and an extension of the log-linear equation (Eqn. 1) as used by Moustafa et al. (1981) is shown in Fig. 6 at fixed concentrations of propylene glycol. The statistical

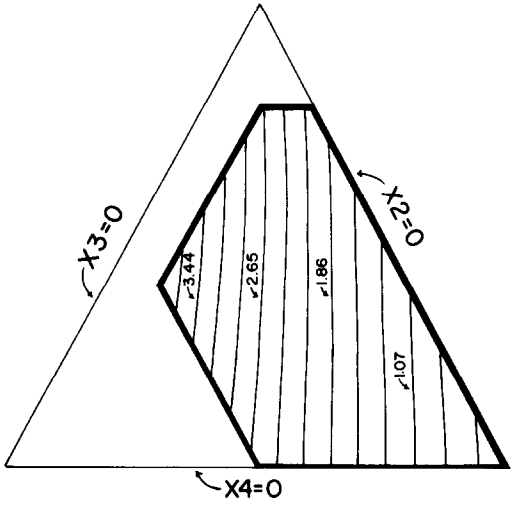

Fig. 5. Contour plot of Eqn. 5 where the view is that of looking down along the response (solubility) axis. Each line represents a constant natural logarithm of the solubility. x_4 = glycerol; x_2 = propylene glycol; and x_3 = water volume fractions.

approach provides a better fit especially when higher glycerol and/or propylene glycol concentrations are used. The mixture response technique like other regression analyses fits data to empirical equations. The equations generated will always provide as good as, and invariably a better, fit than equations based on a theoretical model. This was shown in a comparison of solubility prediction for xanthines in water–dioxane mixtures using this method with that of Martin and coworkers (Ochsner et al., 1985). Prediction of solubility based on first principles is un-doubtedly what is most desirable but this is often very difficult to achieve because a universally acceptable explanation for the non-ideality that is frequently encountered is not available yet. Empirical statistical methods of course require experimental observations of solubility but the equations identified always fit the data regardless of the extent of non-ideality as seen for example in the solubility of xanthines in water–dioxane or phenobarbital in propylene glycol–water–glycerol.

With Eqn. 5, where seven coefficients are being estimated, the use of 42 data points to generate a surface is far in excess of the number needed. From a practical standpoint, one would like to use a number of points that would allow for the

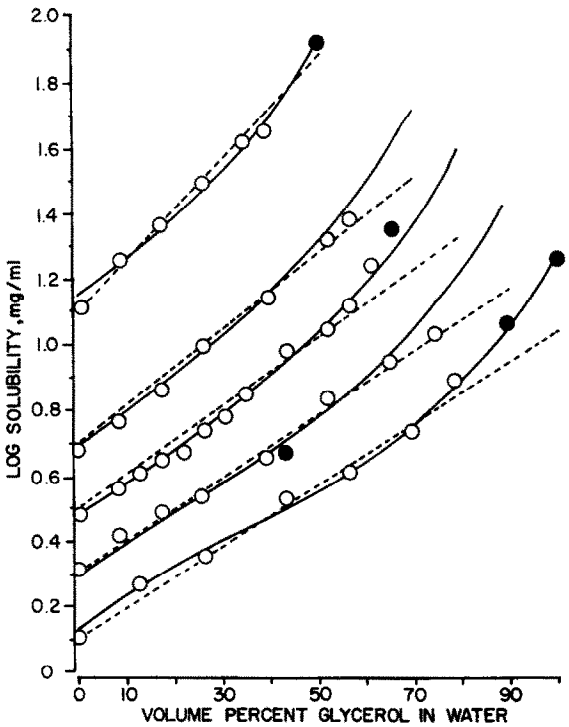

Fig. 6. The relationship between solubility of phenobarbital (mg/ml) and volume fraction of glycerol and water at fixed concentrations of propylene glycol (0, 10, 20, 30 and 50%). Each set of two lines is at a different fixed propylene glycol concentration. Symbols are observed solubilities, dashed lines are solubilities predicted by Eqn. 1 (Moustafa et al., 1981) and the solid line is the solubility predicted by the present method, Eqn. 5: (○) original data points; (●) added data points.

estimation of the desired coefficients while still allowing the investigator some idea of the relative error involved. An analogous situation that most investigators would be familiar with is in the generation of a standard curve in analytical work.

If there is some idea with respect to the form of the equation that will fit the data, then through the statistical concepts of experimental design, it is possible to dictate how many mixtures are needed and what mixtures they should be. While a basic understanding of these selection techniques can be extremely complex to the uninitiated, to use them all that is needed is knowing which model applies and then referring to texts for the appropriate design to be used (Draper and St. John, 1974). In systems such as the one considered here, where there are constraints on one or more of the components, the extreme vertices design as proposed by McLean and Anderson (1966) is particularly useful for the selection of experimental mixtures to be run. Application of the procedures outlined by McLean and Anderson lead to the identification of eleven mixtures (closed circles in Fig. 4) that should be used to generate the predicting polynomial. Data points from among the 42 collected in the solubility study that lay closest to the mixtures generated by the extreme vertices design were used to define again a surface relating solubility to component fraction. Only 10 points were used because when the design generates points that are close to each other the usual procedure is to average the values and run only one point: these points are the two to the right on the top horizontal line of Fig. 4. The actual data points used to generate the new surface based on fewer data points were (referring to Table 2): Mixture Numbers 1, 2, 5, 8, 17, 21, 27, 31, 35, and 36. Using these data points only, the new surface generated is characterized by Eqn. 7.

$$\ln(S) = 2.44x_4 + 5.43x_2 + 0.250x_3 + 0.934x_4x_2 + 0.151x_4x_3 - 1.03x_2x_3 \qquad (7)$$

This surface in theory should be identical to the one derived from the larger data set and indeed one can see that Eqns. 7 and 5 agree quite closely. Table 2 lists the solubilities predicted by Eqn. 7 which compare quite well with those predicted using Eqn. 5 even though Eqn. 7 is a simpler polynomial.

Eight more data points were added extending the volume fraction of each solvent from 0 to 1. Predictions based on equations developed for the 42 point data set were entirely applicable to the 8 additional points. This can be seen in Fig. 6 where the closed circles represent 5 of these added data points whereas the curve was generated from the 42 point data set (Eqn. 5). It should be noted and apparent that the use of equations generated using a particular domain of mixtures need not be applicable in the prediction of solubilities for mixtures outside that original domain.

The statistical techniques illustrated in this report for diazepam and phenobarbital are readily extended to any number of components. With three solvents, an equilateral triangle describes the experimental design region; with four components, a tetrahedron. It is possible to have five or more components but these are difficult to picture in three-dimensional space although they can be handled easily mathematically and on the computer. There are several other extensions of the use of mixture response methods. A predictive equation may be generated and subsequently used for optimizing a formulation, and it may also help in identifying a

theoretical or mechanistic relationship between response and mixture composition. Thin-layer chromatography and high-pressure liquid chromatography are two areas of possible routine application for mixture experiments where solvent mixtures are sought to maximize resolution. Tablet formulation (Jontschev and Welikowa, 1981) is another area where mixture experiments are applicable. Applications of mixture response surface methodology abound in the pharmaceutical sciences and it is hoped that this paper will encourage other workers to learn and apply these methods.

References

Adjei, A., Newburger, J. and Martin, A., Extended Hildebrand approach: solubility of caffeine in dioxane–water mixtures. J. Pharm Sci., 69 (1980) 659–661.

Derringer, G. and Suich, R., Simultaneous optimization of several response variables. J. Qual. Technol., 12 (1980) 214–219.

Draper, N.R. and St. John, R.C., Models and designs for experiments with mixtures. III. Mixture designs for three and four components model. Mathematics Research Center Technical Summary Report No. 1437, University of Wisconsin, 1974b. (Also issued as Statistics Department Technical Report No. 362.)

Draper, N.R. and St. John, R.C., Designs in three and four components for mixtures models with inverse terms. Technometrics, 19 (1977) 117–130.

Draper, N.R. and Smith, H., Applied Regression Analysis (2nd edn.), John Wiley and Sons, New York, 1981, p. 231.

Higuchi, T. and Brochmann-Hanssen, E., Pharmaceutical Analysis, Interscience, New York, New York, 1961, p. 231.

Jontschev, C. and Welikowa, E., Optimierung der Zusammensetzung einer mehrkomponenten-arzneiform durch mathematische Versuchsplanung und experimentelle Prüfung. Pharmazie, 36 (1981) 32–34.

Marquardt, D.W. and Snee, R.D., Test statistics for mixture models. Technometrics, 16 (1974) 533–537.

Martin, A., Newburger, J. and Adjei, A., Extended Hildebrand solubility approach: solubility of theophylline in polar binary solvents. J. Pharm. Sci., 69 (1980) 487–491.

Martin, A., Wu, P.L., Adjei, A., Lindstrom, R.E. and Elworthy, P.H., Extended Hildebrand solubility approach and the log linear equation. J. Pharm. Sci., 71 (1982) 849–856.

McLean, R.A. and Anderson, V.L., Extreme vertices design of mixture experiments. Technometrics, 8 (1966) 447–454.

Moustafa, M.A., Molokhia, A.M. and Gouda, M.W., Phenobarbital solubility in propylene glycol–glycerol–water systems. J. Pharm. Sci., 70 (1981) 1172–1174.

Ochsner, A.B., Belloto, R.J., Jr. and Sokoloski, T.D., Prediction of xanthine solubilities using statistical techniques. J. Pharm. Sci., 75 (1985) in press.

Saleh, A.M., Khalil, S.A. and El-Khordagui, L.K., Solubility and stability of diazepam in sodium salicylate solution. Int. J. Pharm., 5 (1980) 161–164.

Scheffe, H., Experiments with mixtures. Royal Stat. Soc. J., B, 20 (1958) 344–360.

Williams, N.A. and Amidon, G.L., Excess free energy approach to the estimation of solubility in mixed solvent systems I: Theory. J. Pharm. Sci., 73 (1984a) 9–13.

Williams, N.A. and Amidon, G.L., Excess free energy approach to the estimation of solubility in mixed solvent systems II: Ethanol–water mixtures. J. Pharm. Sci., 73 (1984b) 14–18.

Williams, N.A. and Amidon, G.L., Excess free energy approach to the estimation of solubility in mixed solvent systems III: Ethanol–propylene glycol–water mixtures. J. Pharm. Sci., 73 (1984c) 18–23.

Yalkowsky, S.H., Flynn, G.L. and Amidon, G.L., Solubility of nonelectrolytes in polar solvents. J. Pharm. Sci., 61 (1972) 983–984.

Yalkowsky, S.H. and Flynn, G.L., Correlation and prediction of mass transport across membranes II: Influence of vehicle polarity on flux from solutions and suspensions. J. Pharm. Sci., 63 (1974) 1276–1280.