Organic Process Research & Development **2002**, *6*, 841−846

Application of the Response Surface Methodology for the Design of a Lixiviation Process

José Luis Navarrete-Bolaños,*[,†,‡] Hugo Jiménez-Islas,[†] Enrique Botello-Alvarez,[†] and Ramiro Rico-Martínez[†,‡]

Departamento de Ingeniería Química-Bioquímica, Instituto Tecnológico de Celaya, Av. Tecnológico s/n, C.P. 38010. Celaya, Gto., México, and Universidad Autónoma de Querétaro, Departamento de Investigación y Posgrado en Alimentos (PROPAC), Cerro de las Campanas s/n, C.P. 76010. Querétaro, Qro., México

Abstract:

The design of many separation processes is often hindered by the lack of distribution data between two immiscible phases. Inappropriate or partial designs result in inefficiencies that translate into poor recovery of the products, overdesigned separation equipment, or additional purification and separation stages. In recent years, statistically based experimental design has been successfully used to quantify relevant factors in many biological and chemical processes. In this study, we illustrate the capabilities of one of these approaches, the response surface methodology (RSM) for the optimization of lixiviation processes of marigold flower flour for xanthophyll extraction (oleoresin) using hexane. The method leads to conditions for the maximum recovery of the pigment, allowing the determination of the optimal residence time (14.7 min) and flour-to-solvent ratio (1: 5) at 35 °C. These values can be used in the lixiviation process in any operation mode. In this study, we illustrate the use of a countercurrent process that allows recovery of 97.5% of the oleoresin.

Introduction

The most efficient way to produce many chemical products is to extract them from natural sources. The challenge faced by scientists and engineers is then to develop the separation methods necessary to isolate and purify such products. The separation processes most commonly used in the Biochemical and Chemical industries can be divided into physical, equilibrium-controlled, and rate-controlled separations. While, in general, clear guidelines to select the most appropriate type of process and identify the relevant operating parameters for a given separation are available, the design of the separation process is often hindered by the lack of accurate distribution data. A good prediction of the equilibrium for a given system may allow the recovery of high proportion of the main products, avoiding inappropriate or partial design of the separation train. On the other hand, inaccurate data may result in inefficiencies that translate into poor recovery of the products, overdesigned separations equipment, or additional purification and separations stages. In some cases, the cost of the final product may be doubled because of deficiencies in the separation stages, and one may face recoveries that barely reach 50% of the main product in the original raw materials.

For complex systems, such as the ones formed by substrates directly obtained from natural sources, predicting the distribution of a substance between two immiscible, or partially immiscible phases, is not an easy task. The current engineering practice relies mainly in experimental measurements of the distribution data. However, the inherent nonlinear (nonideal) behavior of such systems often brings about the need for a large number of experiments to achieve reasonable inter- and extrapolations of the experimental data distribution. Under these circumstances, it is desirable to take advantage of experimental strategies based on statistical techniques to minimize the number of experiments required to achieve an accurate description of the equilibrium between phases. Moreover, if these techniques can be used to predict the conditions of maximum recovery, the data thus gathered will be more accurate, precisely for the more relevant experimental conditions. One of such strategies is the response surface methodology (RSM). This methodology comprises a group of statistical techniques for empirical model building and model exploitation. The RSM seeks, via a careful design and analysis of experiments, to relate a response, or output variable, to the levels of a number of parameters, or input variables, that affect such response. The RSM has been used in many fields to improve the efficiency of many processes and optimize design variables.[1−9] However, and despite the success obtained for these applications,

* Author to whom correspondence may be sent. Telephone: (+52) 461 61 1 75 75. Fax: (+52) 461 61 1 79 79. E-mail: jlnb@itc.mx.

† Instituto Tecnológico de Celaya.

‡ Universidad Autónoma de Querétaro.

(1) Saval, S.; Pablos, L.; Sánchez, S. Optimization of a culture medium for streptomycin production using response-surface-methodology. *Bioresour. Technol.* **1993**, *43*, 19−25.

(2) Harker, M.; Tsavalos, A. J.; Young, A. J. Use of response surface methodology to optimize carotenogenesis in the microalga *Haematococcus pluvialis*. *J. Appl. Phycol.* **1995**, *7*, 399−406.

(3) Park, J. W.; Testin, R. F.; Vergano, P. J.; Park, H. J.; Weller, C. L. Application of laminated edible films to potato chip packaging. *J. Food Sci.* **1996**, *61*, 766−768.

(4) Smith, N. K.; Gilmour, S. G.; Rastall, R. A. Statistical optimization of enzymatic synthesis of derivatives of trehalose and sucrose. *Enzyme Microb. Technol.* **1997**, *21*, 349−354.

(5) Dona, O. J. Response Surface Methodology: An Analytical Method for Locating Migrated Contaminant Sources. In *Government Reports Announcements and Index*; Air Force Institute of Technology, School of Engineering: Washington, DC, 1997; Issue 16.

(6) El-Masri, H. A.; Reardon, K. F.; Yang, R. S. Integrated approaches for the analysis of toxicologic interactions of chemical mixtures. *Crit. Rev. Toxicol.* **1997**, *27*, 175−197.

(7) El-Halouat, A.; Debevere, J. M. Effect of water activity, modified atmosphere packaging and storage temperature on spore germination of moulds isolated from prunes. *Int. J. Food Microbiol.* **1997**, *35*, 41−48.

(8) Vázquez, M.; Martin, A. M. Optimization of *Phaffia rhodozyma* continuous culture through response surface methodology. *Biotechnol. Bioeng.* **1998**, *57*, 314−320.

10.1021/op0255534 CCC: $22.00
Published on Web 08/16/2002

the RSM has had limited application in other areas, such as in improving the efficiency and design of separation processes. In this contribution, we illustrate the application of the RSM in the design of separation processes using as the model system a colorant extraction from natural sources. The design goal is to establish the optimum value of the operating variables that could be used in the most common modes of the separation equipment operation: batch and continuous, concurrent and countercurrent contacting. In previous work[10] we described a statistically based strategy seeking to achieve the design goal from a single experimental design. Such an approach may prove to be inaccurate and to lack robustness since it is based on extrapolations of a fitted model. In this contribution, we proposed a refinement of this original approach, consisting of two-tier strategy. The initial experimental strategy is used to construct a low-order model that predicts the region of maximum yield. Using this prediction we devise a second experimental design that, statistically, allows a higher-order (and thus more accurate) description of the separation process. The robustness of the obtained result is examined by looking at the agreement between the two approximations. For our model system, the separation process involves the lixiviation (extraction via dissolution in a liquid solvent) of marigold flower flour with hexane.[10−13] The oleoresin pigment (xanthophylls) obtained from this extraction process is used commercially as an additive to poultry feed. In addition, some xanthophylls had demonstrated properties as cancer prevention agents, ligament repair in muscular tissue, aid in enzyme transport, and a prevention agent for age-related macular degeneration.[14,15] These properties have sparked a renewed interest to develop and improve the xanthophylls' extraction methods to increase yield extraction.

Materials and Methods

Flour: Solid to Leach. The material used in the experiments was marigold flower flour (*Tagetes erecta*) with an average particle size of 0.372 mm and humidity content of 10% from Industries ALCOSA S.A de C. V., Guanajuato, Mexico. Flour from a single batch was used for all experiments. Both flour and oleoresin analyses to determine the total concentration of xanthophylls present in the samples were performed according to the AOAC 970.64 (1984) method.[16] All experiments were replicated.

Extraction and Quantification of Total Xanthophyll. The extractions were carried out in batch processes using analytical grade hexane (J. T. Baker) under conditions based in an experimental design. When the extraction was concluded, both phases (light and heavy) were separated.[10] The light phase (liquid) was concentrated and analyzed to determine the xanthophylls concentration[6] and the main components profile was obtained by high performance liquid chromatography.[10] The heavy phase (solid) was desolventized, and the retained solvent volume was determined by difference of weights on an OHAUS digital analytical scale (Explorer, precision ±0.0001 g). Solid free solvent was also analyzed to determine the xanthophylls concentration.[16]

Experimental Design

There are four major elements to be established in the design of processes based on diffusional forces: the number of equilibrium stages or their equivalent, the time of phase contact required, the permissible rate of flow, and the energy requirements.[17−18] These design constraints translate, for our purposes, into the determination of the relevance of four factors: (A) flour sample weight, (B) solvent volume, (C) extraction temperature, and (D) extraction time. The first two are related to the permissible flow rate, the fourth relates directly to the contact time, and the third defines (indirectly) the energy requirements. We performed a factorial fractional experimental design (divided in half) for two levels of each factor (A: 1.0 and 2.0 g, B: 100 and 200 mL, C: 30 and 40 °C., D: 5.0 and 10.0 min.), resulting in a 2^{4-1} factorial design. Under the fractional design one should add experimental tests corresponding to the mean levels of the factors to account for nonadjustable data and allow us to calculate and estimate the prediction error of the statistical model. In our case we added three additional experiments at these mean levels. A fractional factorial design can be interpreted graphically as depicted in Figure 1. The experimental points are located at the corners and center of the cube in the lower left corner. The points in the center correspond to the experiments with the mean values of the factors. The coded values of the factors for the 11 experiments (−1 for the lower level, +1 for the upper level, and 0 for the mean value) are shown in Table 1.

The results from this initial design are used as the basis to predict the factor levels for which one will obtain the largest response value. Using this prediction, we devised a second experimental design to reach a more accurate dependence of the response variable with respect to each factor level. This complementary strategy takes the form of a second-order composite design that has characteristics of rotatability and orthogonality and results very efficiently in the number of runs required. Similar experimental designs

(9) Ooijkaas, L. P.; Wilkinson, E. C.; Tramper, J.; Buitelaar, M. R. Medium optimization for spore production of *Coniothyrium minitans* using statistically based experimental designs. *Biotechnol. Bioeng.* **1999**, *64*, 92−100.

(10) Navarrete-Bolaños, J. L.; Jiménez-Islas, H.; Rico-Martinez, R.; Domínguez-Domínguez, J.; Regalado-González, C. A statistically based strategy for obtaining equilibrium data in xanthophylls lixiviation. *J. Food Sci*. **2001**, *66*, 1143−1146.

(11) Philip, T.; Berry, J. W. A process for the purification of lutein-fatty from marigold petals. *J. Food Sci*. **1975**, *41*, 163−165.

(12) Britton, G. Carotenoids. *Methods Plant Biochem.* **1991**, *7*, 473−518.

(13) Craft, N. E.; Soares, J. H., Jr. Relative Solubility, Stability and Absorptivity of Lutein and Beta-Carotene in Organic Solvents. *J. Agric. Food Chem*. **1992**, *40*, 431−434.

(14) Taylor, R. F. Natural Food Colorants as Nutraceuticals: Markets and Applications. In *Abstracts International Congress and Symposium on Natural Colorants*; Acapulco, México, 1996.

(15) Seddon, J. M.; Ajani, U. A.; Sperduto, R. D.; Hiller, R.; Blair, N.; Burton, T. C.; Farber, M. D.; Gragoudas, E. S.; Haller, J.; Miller, D. T.; Yannuzzi, L. A.; Willett, W. Dietary carotenoids, vitamins A, C and E, and advanced age-related macular degeneration. A multicenter study. *J. Am. Med. Assoc.* **1994**, *272*, 1413−1420.

(16) AOAC. *Official Methods of Analysis*, 14th ed.; Association of Official Analytical Chemists: Washington, DC, 1984.

(17) Treybal, R. E. *Mass-Transfer Operations*, 3rd ed.; McGraw-Hill Inc.: New York, 1987.

(18) Geankoplis, C. J. *Process Transport and Unit Operations*; 3rd ed.; Prentice-Hall: New York, 1993.

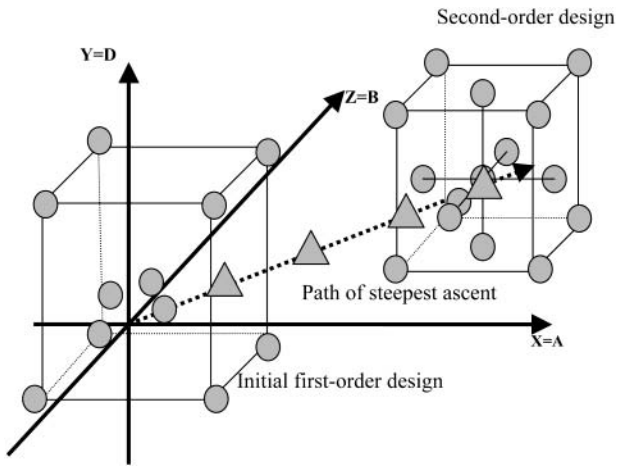

Figure 1. **Graphical representation of the response surface methodology. It includes the first factorial design, the extrapolation given by the steepest ascent path, along with the factorial design for the new search region (second order design).**

Table 1. **Coded levels for the 11 experiments generated by the initial experimental design with their corresponding response value (total xanthophyll extracted)**

run	*A* (flour)	*B* (solvent)	*C* (temperature)	*D* (time)	xanthophyll concentrations[a]
1	−1	+1	−1	+1	60.80
2	−1	−1	+1	+1	53.12
3	0	0	0	0	76.52
4	+1	−1	+1	−1	88.54
5	0	0	0	0	79.68
6	+1	+1	+1	+1	97.39
7	+1	−1	−1	+1	125.22
8	+1	+1	−1	−1	55.65
9	−1	−1	−1	−1	43.64
10	−1	+1	+1	−1	36.26
11	0	0	0	0	76.52

[a] Grams of xanthophyll/kilogram of oleoresin.

were originally proposed by Box and Behnken in 1960.[19]

Statistical Analysis. We used ANOVA procedures with probability distribution values in the analysis of data (Statgraphics plus V 2.1 software). We analyzed the effects of the principal factors in the maximization of the yield of the oleoresin extraction. Data analysis using least-squares allows us to construct a model that describes the relationship between the factors and the yield.

Multistage Countercurrent Operation-Shanks System. An efficient extraction of the pigment can only be achieved using multiple contact operations. We examine the validity of the results obtained by the RSM methodology on a multistage countercurrent operation, where the solution to be withdrawn is in contact with the freshest solid and the fresh solvent is added to the solid from which most of the solute has already been leached. The system can, of course, be operated with any number of stages. Its purpose is to increase the mass transfer over and above what is possible with a single stage and obtain higher concentrations on the final product.

(19) Montgomery, D. C. *Design and Analysis of Experiments*, 4th ed.; John Wiley & Sons, Inc.: New York, 1997.

Table 2. **Analysis of variance to the first-order model[a]**

source variation	S.S.	D.V.	P.V.
model	5776.77	3	0.002
residual	904.13	7	
total	6680.90	10	

[a] $R^2 = 86.467$; standard error = 11.367. S.S. = sum of squares; D.F. = degrees of freedom; P.V. = probability distribution values.

Table 3. **Points on the path of steepest ascent**

run	coded conditions A	B	D	flower flour (g)	hexane (ml)	time (min)	[*Xt*] g/kg
3,5,11	0	0	0	1.5	150	7.5	72.12
12	1.53	−0.68	1.0	2.26	123.5	10.0	92.93
13	3.12	−1.56	2.0	3.06	71.5	12.5	102.86
14	4.78	−2.57	3.0	3.89	20.65	15.0	125.49
15	6.51	−3.68	4.0	4.71	−32.58	17.5	- - -

Results and Discussion

The results of the experiments generated from the 2^{4-1} design are expressed as a function of the total extraction of xanthophyll (response variable), as shown in the last column of Table 1. The variance analysis show (for a confidence interval of 0.05) that the main factors *A* and *D* are significant, and in a smaller degree, the factor *B* and the interaction $AB = CD$.[10] Using Yates' algorithm[19] it can be shown that the variable *C* (temperature) is not significant in determining the yield of the extraction within the range explored in our experiments. Consequently, we set this variable constant to its mean level for the remaining essays. With this information, it is possible to describe the relationship that exists among the significant factors *A*, *B*, and *D* by constructing a regression-based polynomial model using least-squares. For this case, the regression model obtained is the following:

$$[\text{xanthophyll concentration}] = 72.12 + 21.62A - 7.55B + 14.05D$$

Here, the variables are specified in their coded units. The standard error of the model is 11.367, and the correlation coefficient (R^2) is 86.467. In this equation one implicitly assumes that the dependence of the response variable with respect to the factors is linear and that the effects of these factors are additive. The analysis of variance (Table 2) shows no reason to question the adequacy of these assumptions under a statistical basis.

The results of this initial design are shown graphically in Figure 1, along with the fitted first-order model. The model can be tentatively accepted as a rough geometrical representation of the underlying response function *over the experimental region explored thus far.* Using this model, we search for the region for which the xanthophyll extraction will be maximized, following the path of steepest ascent, perpendicular to the contour lines, indicated by triangles in Figure 1. A convenient set of points of the steepest ascent path is shown in Table 3. The proposed ascent path suggests conditions (run 14 in Table 3) giving the best yield. Additional points beyond run 14 will not be feasible because

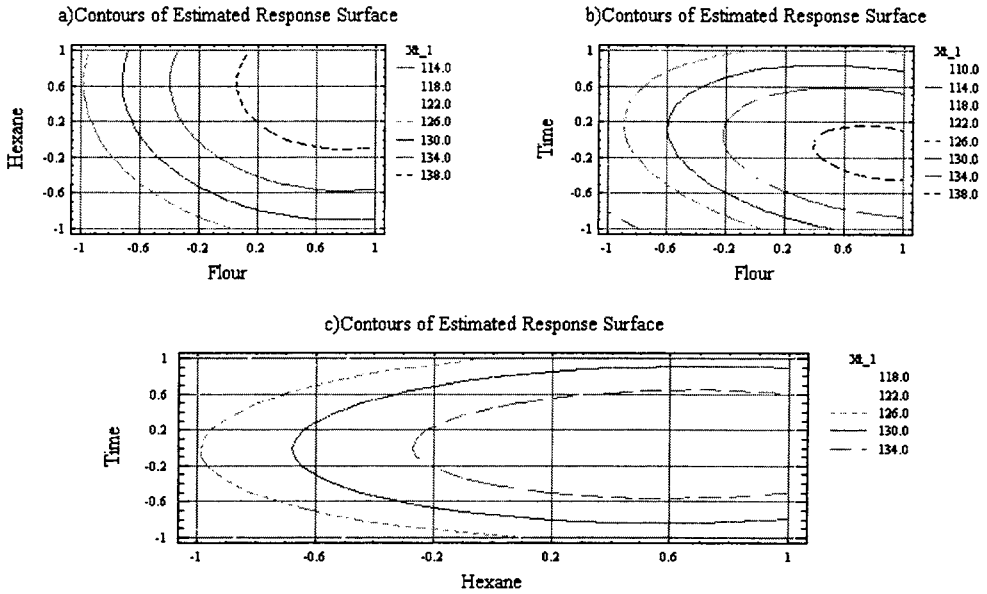

Figure 2. **Plots of response for the operating variables analysis showing an optimum point.**

Table 4. **Box−Behnken design and data obtained from experimental runs**

run	*A* (flour)	*B* (hexane)	*D* (time)	xanthophyll concentrations[a]
1	−1	−1	0	114.98
2	−1	+1	0	118.84
3	+1	−1	0	134.18
4	+1	+1	0	138.68
5	−1	0	−1	113.7
6	−1	0	+1	119.7
7	+1	0	−1	130.2
8	+1	0	+1	121.7
9	0	−1	−1	111.74
10	0	−1	+1	112.82
11	0	+1	−1	128.76
12	0	+1	+1	132.76
13	0	0	0	135.74
14	0	0	0	133.98
15	0	0	0	137.8

[a] Grams of xanthophyll/kilogram of oleoresin.

the solvent volume is reduced to zero (run 15). Graphs of these results suggest that subsequent experiment should be made in the neighborhood of run 14. In this point a new (Box−Behnken) design was made (Table 4) using the levels of the three remaining main variables suggested by the first design and the steepest ascent path (*A*: 3 and 5 g, *B*: 16 and 26 mL, *D*: 12.5 and 17.5 min). Since one would expect that the region described by these new experiments will contain the optimum of the xanthophyll extraction, we set a design that will allow us to describe the system via at least a second-order model to achieve better accuracy.

$$[\text{xanthophyll concentration}] = 135.84 + 7.30A + 5.60B + 0.35D - 4.66A^2 - 4.41B^2 - 9.98D^2 - 0.028AB - 3.75AD + 0.79BD$$

This description was once again obtained via least-squares from the 15 experimental runs of the Box−Behnken design. The model was also used to construct response plots (Figure 2a, b, and c). These plots hint at the presence of a global optimum, a set of operating conditions on the three variables, that leads to maximum xanthophyll extraction. The location of the optimum becomes a simple example of locating an extremum for a multidimensional system via derivation of the second-order model and solution of the resulting set of linear equations:

$$\frac{\partial[Xt]}{\partial A} = 7.30 - (2)4.66A - 0.028B - 3.75D = 0$$

$$\frac{\partial[Xt]}{\partial B} = 5.60 - 0.028A - (2)4.41B + 0.79D = 0$$

$$\frac{\partial[Xt]}{\partial D} = 0.35 - 3.75A + 0.79B - (2)9.98D = 0$$

The system solution shows the following result: $A = 0.83$, $B = 0.62$, and $D = -0.11$ as optimum values for the extraction. Using the transformation (X_i = optimum factor value [(factor upper bound − factor lower bound)/2] + mean factor value) one obtains the decodified factor values. This procedure indicates that the extraction should take place allowing for 14.7 min residence time and handling a flour-to-hexane ratio of 1:5 (see Figure 2 a, b, and c) to maximize the pigment extraction. A verification run around these conditions gives a response of 138.28 g/kg of xanthophyll extracted which compares favorably against the predicted model value of 140.58 g/kg.

Countercurrent Extraction. Using these conditions, we set an extraction battery in countercurrent operation in nine stages until the steady state was reached.[17] At the end, the liquid phase (extract) and the solid phase (leached flour) were separated for analysis from each stage. The extract consists of a mixture of hexane, recovered xanthophyll, and fatty acids. The leached flour consists of a mixture of hexane, unrecovered xanthophyll, and flower flour. The extracts were concentrated on a rotary evaporator, eliminating hexane until an oleoresin-type sample was obtained and the solvent was removed from the leached solids. The analysis of the liquid-

Table 5. **Distribution of xanthophyll in each stage for the Shanks system[a]**

stage	X	Y
1	1.55	0.41
2	1.98	1.12
3	2.03	2.04
4	4.32	8.49
5	12.8	19.95
6	22.7	28.87
7	34.8	47.62
8	48.4	70.67
9	72.3	97.5

[a] X = total xanthophyll percentage in solid phase. Y = total xanthophyll percentage in liquid phase.

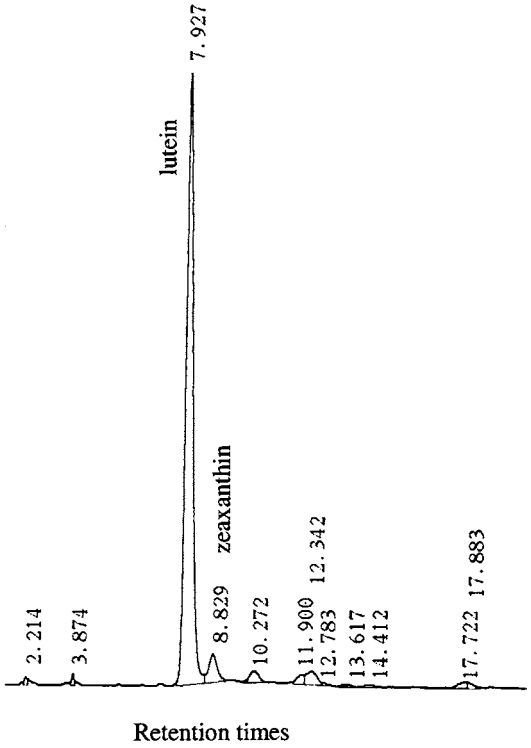

Figure 3. **Chromatograph of HPLC for the saponified oleoresin obtained as final product from the multistage countercurrent operation.**

phase concentrate of the last stage (stage 9) reveals that 97.5% of the original amount of xanthophyll contained in the marigold flower flour is recovered (Table 5). The unrecovered xanthophyll quantified in the first stage represents about 1.55% of the total xanthophyll present in the flour. The mass balances applied to the system indicate a loss of 0.95% of the total xanthophyll due in part to the product deposited in the laboratory equipment.

Analysis of the Oleoresin with High Performance Liquid Chromatography (HPLC). The extracted oleoresin, obtained in the stage 9, was analyzed by HPLC to determine the concentration of their main components. The resulting profile is shown in Figure 3. The main components are lutein (85.59%) and zeaxanthin (4.54%). They indicate that the concentrations of the original components are conserved and can be compared favorably with those reported previously.[10,20]

Predesign Xanthophyll Extraction Process. Our results allow the predesign of a xanthophyll lixiviation plant for processing marigold flowers flour. They indicate that with the use of a countercurrent system with nine stages the expected recovery is 97.5%. The system should be operated with a 1:5 flour-to-hexane ratio, phase contact time of 14.7 min, and operation temperature of 35 °C. The conditions cited are the basis to determine the equipment cross-sectional area and energy requirements for operation. Such determinations should take into account the stage efficiency and should guarantee high recovery yields and absence of unwanted changes in the final products.

Conclusions

The analysis presented here demonstrates the application of RSM as a robust tool for optimization of the lixiviation process of marigold flower flour. At the core of this work lies the issue of selecting a set of basis points for the construction of a regression-based model that reproduces the dependence of the "response" of the system with respect to the operating variables. The RSM seeks to address this problem by choosing points that lie in regions associated with high probability. In our illustration, the high yield output obtained shows good convergence and robustness properties. Furthermore, different order approximations are obtained from model outputs from different sampling points, and the agreement between successive order approximations is a strong indicator of a good fit. In our illustration, the agreement in the order of magnitude and relevant factor values predicted by the two successive models confirm that the obtained conditions are nearly optimum in a statistical sense. Alternatively, one could use a random choice of experimental points from the entire set of possible experimental conditions, which could lead to poor estimates of the yields of the output values. The RSM, however, presents some limitations. If the probability distributions of the inputs exhibit strong discontinuities, the probability density estimates using RSM are likely not to converge.

The RSM is, in general, useful to address uncertainty analysis of complicated models in many engineering fields. It represents a potentially valuable alternative to the currently prevalent methods of uncertainty analysis that are computationally prohibitive for complex systems.

The use of regression models based on experimental data to describe physicochemical phenomena should be examined carefully. The resulting models could be inaccurate because of experimental errors and a poor selection of "relevant factors" and their operation values. Under these circumstances, the prediction of "optimal" conditions via extrapolation of the models should always be tested. The use of RMS strategy illustrated here provides an alternative to overcome these concerns.

The two-tier strategy illustrated here leads to improved extraction conditions that require 16% less solvent and results in 2% higher purity of the recovered pigment with respect

(20) Quackenbush, F. W.; Miller, S. L. Composition and analysis of the carotenoids in marigold petals. *J. Assoc. Off. Agric. Chem*. **1972**, *55*, 617−621.

to the cases for which only a single experimental design is used.[10] While these improvements may appear small, they can be important in determining the economical feasibility of the process.

When the amount of solvent per unit of feed is fixed, the number of stages required decreases as solvent rate or solvent-to-solid ratio increases. Since the capacity of the equipment for handling the larger liquid flow must at the same time increase, the cost of equipment must then pass through a minimum. The extracts solutions become more dilute as solvent rate is increased; consequently, the cost of solvent removed increases. The total cost, which is the sum of fixed and operation costs, reaches a minimum at the optimum value of the operation variables.

The response surface methodology is mainly applied to laboratory or pilot-plant operations. However, the condition that was optimum to these levels may not be optimum for the full-scale process. This "scale-up" usually results in distortion of the optimum conditions. Even if the full-scale plant begins operation at the optimum; it will eventually "drift" away due to changes in raw materials, environmental regulations, equipment, and operating personnel. Thus, full-scale processes are also susceptible of profiting from the use of RSM.

The RSM technique can be straightforwardly extended in the study of other types of separations used in biochemical or chemical processes whenever the data distribution between phases cannot be described from fundamental principles.

Acknowledgment

We gratefully acknowledge financial support by the Sistema de Investigación Miguel Hidalgo (SHIGO) and the Consejo del Sistema Nacional de Educación Tecnológica (COSNET).

Received for review May 27, 2002.

OP0255534