

혼합물실험계획에 의한 커피혼합비율 최적화에 대한 연구 A Case Study on the Optimum Formulation of Coffee by a Mixture Experiment Design

이 종 성* 문 제 응**
Lee, Jong-Seong Moon, Je-Woong

Abstract

Many industrial products such as paints, ink and adhesives are composed of the ingredients of a mixture. In mixture experiments, the characteristics of quality(response) depends only on the proportions of the ingredients and does not depend on the total amount of the mixture.

This article discusses the constrained mixture experimental design, the data analysis, and the optimum formulation of ingredients based on the two quality characteristics - taste and flavor. It is shown that efficient designs can be constructed from D-optimal criterion. Special cubic models were selected as the final mixture response surfaces for both responses. The desirability function was used for the optimization of the two responses.

키워드 반응표면분석, 혼합물 실험, 복수 반응치

Key Words : *Response surface methodology, Mixture experiments, Multiple response optimization.*

1. 서론

혼합물 실험(mixture experiments)계획이란 혼합량(volume)은 문제가 되지 않고 각 성분의 혼합비율(mixing proportion)만이 중요시 되는 경우, 어떠한 성분(components)이 관심이 있는 반응치(responses)에 유의한 영향을 주고 있으며, 그 반응치를 최대 또는 최소로 할 수 있는 최적의 혼합비율이 얼마인지를 알고자 하는 경우에 이용되는 실험계획이다.[2][4]

삼프나 비누 등의 제제, 약품, 음식물, 음료수, 수지, 에칭액 등의 제품들이 이러한 혼합물의 전형

적인 예라 할 수 있다. 본 논문에서는 3가지의 성분 - 커피, 프림, 설탕의 혼합물인 일반적인 음용 커피를 연구 대상으로 하여, 맛과 향이라는 두 개의 반응치를 동시에 최적화 할 수 있는 구성 성분의 혼합비를 탐색하는 다중 반응표면 분석의 혼합물 실험계획과 분석 및 최적화 과정을 수행하였다. 이와 같이 복수의 반응치에 대해 구성 성분의 최적 혼합비율을 탐색하는 다중 반응표면 분석은 일반적인 산업제품의 개발 및 실험단계에서 흔히 시도될 수 있는 실험계획의 전형적 예라고 할 수 있다.[1][3]

일반적으로 이론 상 가능한 모든 혼합비율에서 실험을 실시하는 것이 기술적/경제적인 측면에서 매우 비효율적이므로 본 논문에서는 각 성분의 혼합비율을 실질적 의미가 있는 범위 내로 제한하였으며, 또한 이 제한된 실험영역에서 D-optimality

* 강원대학교 산업공학과 교수, 공학박사

** [주]케이엘 전자

algorithm을 이용하여 실험점을 선택함으로써 추정 회귀계수의 분산을 가능한 최소화 할 수 있도록 하였다 [5][6]

2. 제한된 영역에서의 실험계획

본 커피 실험의 문제에서 실험계획 상의 심플렉스(simplex) 전 영역에서 실험점(design points)을 선택하는 것은 획득된 정보의 실효성 측면에서 바람직하지가 않다 즉, 커피 성분비율이 0인 경우 커피 혼합물이라고 할 수 없으며 지나치게 많거나 적은 설탕과 프림의 성분비율은 관심의 대상에서 제외되어야 하므로, 본 연구에서는 실험의 실제 수행 영역을 다음과 같이 제한하였다.

- 커피: $0.1 \leq x_1 \leq 0.8$
- 설탕: $0.1 \leq x_2 \leq 0.8$
- 프림: $0.1 \leq x_3 \leq 0.6$

위의 영역 내에서 각 성분비율의 합은 100%. 즉, $x_1 + x_2 + x_3 = 1$ 을 만족하여야 한다. 여기서 x_1, x_2, x_3 는 각각 커피, 설탕, 프림의 혼합비율이다.

실험의 반응변수(response variables)는 커피의 맛의 평가치(y_1)과 향의 선호도(y_2)이다. 맛에 대한 평가치는 1부터 13까지의 순위값(ranking value)으로 측정하고, 향에 대한 평가치는 가장 만족하는 향의 정도를 5로 하는 5점 척도로 측정하였다.

실험점의 선정은 제약조건에 의해 만들어진 가능 실험영역의 vertices, edge centers, overall centroid, axial points 등의 후보 실험점(candidate points) 중에서 추정 반응표면 모형의 회귀계수의 분산이 최소가 되도록 D-optimal design algorithm을 사용하여 4개의 vertices, 3개의 edge centers, 1개의 axial point 와 overall centroid 등 9개의 실험점을 선택하였고, 또한, Linear, quadratic, special cubic, 및 full cubic 등의 여러 반응표면 모형을 적합시키며 적합도 검정을 수행하기 위하여 일부 실험점에서 4회의 반복을 추가하여 총 13회의 실험을 실시하였다.

표 1에서 y_1 과 y_2 는 각 실험점에서의 맛과 향에 대한 반응치이며, data sheet의 평가치를 평균하여 기록한 것이다.

3. 반응표면분석

Linear, Quadratic, Special Cubic, Full Cubic model 등의 추정되는 혼합물 반응표면모형 중에서 가장 적절한 모형을 찾기 위하여 각 모형을 적합시키고, F-검정, Lack-of-fit Test, MSE, R2 및 Adjusted R2 등을 두 반응치에 대하여 검토해 보았다.

<표 1> D-optimal criterion에 의한 실험계획.

실험점	Components, Mixing proportion			Responses	
	커피 x_1	설탕 x_2	프림 x_3	y_1	y_2
10	0.213	0.563	0.225	3.48	3.32
13	0.325	0.325	0.350	3.32	3.38
1	0.800	0.100	0.100	11.84	1.32
7	0.450	0.450	0.100	5.17	3.34
2	0.100	0.800	0.100	6.26	2.85
4	0.100	0.300	0.600	8.41	2.16
3	0.300	0.100	0.600	8.80	2.85
6	0.550	0.100	0.350	7.13	3.31
5	0.100	0.550	0.350	4.90	2.95
1	0.800	0.100	0.100	11.70	2.01
2	0.100	0.800	0.100	6.10	3.41
7	0.450	0.450	0.100	5.20	2.84
4	0.100	0.300	0.600	8.69	1.90

그 결과 표 2와 같이 y_1 은 Quadratic과 Special Cubic model이 통계적으로 유의한 것으로 나타났는데, R^2 및 Adjusted R^2 값, 그리고 표 3의 Lack-of-fit Test의 결과를 종합하여 보면 Special Cubic Model이 보다 더 유의함을 알 수 있다.

그림 1과 그림 2는 y_1 에 대해 정규성과 등분산성을 확인하기 위한 Normal probability plot과 잔차도로서 모형에 대한 통계적 가정의 타당성을 보여주고 있다.

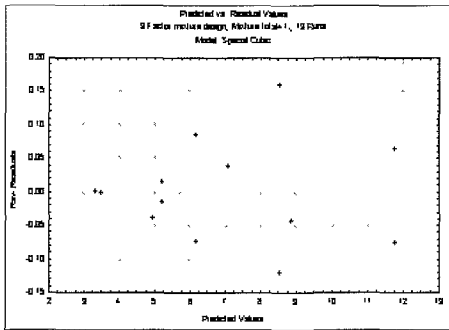
<표 2> Model Summary Statistics (y_1)

EXPERIM. DESIGN	SS Effect	df Effect	MS Effect	SS Error	df Error
Linear	38.294	2	19.147	53.612	10
Quadratic	52.711	3	17.570	0.901	7
Special Cubic	0.833	1	0.833	0.068	6
Total Adjusted	91.906	12	7.659		
EXPERIM. DESIGN	MS Error	F	p	R-Sqr	R-Sqr Adj.
Linear	5.361	3.571	0.068	0.417	0.300
Quadratic	0.129	136.46	1.4E-06	0.990	0.983
Special Cubic	0.011	73.498	1.3E-04	0.999	0.999

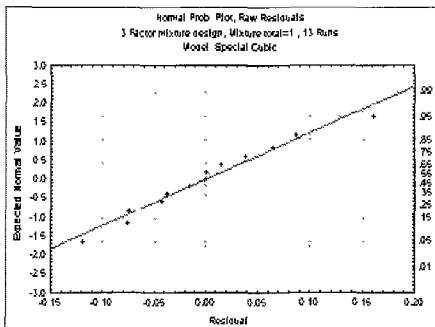
<표3> Lack of Fit Tests (y₁)

MODEL	SS	df	MS	F	p
Linear	53.55	6	8.925	573.49	8.1E-6
	0			2	
Quadratic	0.839	3	0.280	17.972	0.009
Special	0.006	2	0.003	0.186	0.837
Cubic					
Pure	0.062	4	0.016		
Error					

<그림1> 등분산성 확인을 위한 잔차 plot



<그림2> Normal probability plot

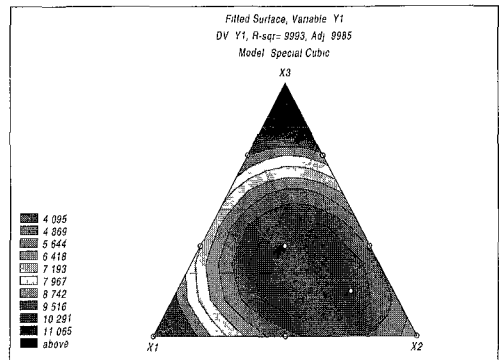


실험 데이터로부터 적합된 Special cubic의 반응 표면 모형은 다음과 같다.

$$\begin{aligned} \hat{y}_1 = & 11.7762x_1 + 6.1737x_2 + 14.8770x_3 \\ & - 15.1604x_1x_2 - 25.2255x_1x_3 \\ & - 18.9223x_2x_3 - 30.6058x_1x_2x_3 \end{aligned}$$

그림 3은 y₁에 대한 special cubic 반응표면의 contour plot이다. 이 그림은 커피(x₁), 프림(x₃)이 비슷한 비율이고 설탕(x₂)이 약간 많이 혼합되었을 때 맛에 대한 선호도가 비교적 높게 나타난다는 것을 보여주고 있다.

<그림3> Response surface contour plot(y₁)



<그림4> Three-dimensional surface plot(y₁)

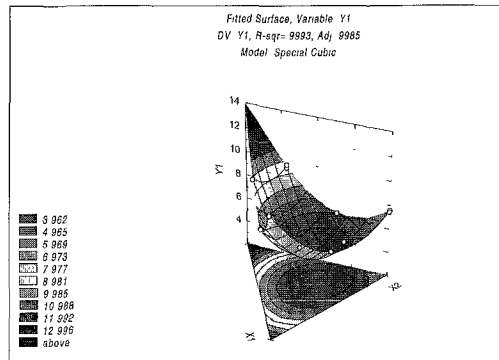


그림 4는 y₁에 대한 3차원 반응표면 plot이다. 이 그림으로부터 정상점이 최소점이며, y₁만을 고려하여 최적의 혼합비율을 찾는다면 x₁ = 0.21277, x₂ = 0.492, x₃ = 0.29523가 된다는 것을 알 수 있다.

같은 방법으로 y₂에 대해서도 분석하여 볼 수 있다. F 검정과 적합도 검정의 결과인 표 4와 표 5로부터 y₂에 대해서는 Quadratic 반응표면 모형이 통계적으로 유의한 것으로 판단된다.

데이터로부터 적합된 y₂에 대한 quadratic 반응 표면 모형의 추정식은 다음과 같다.

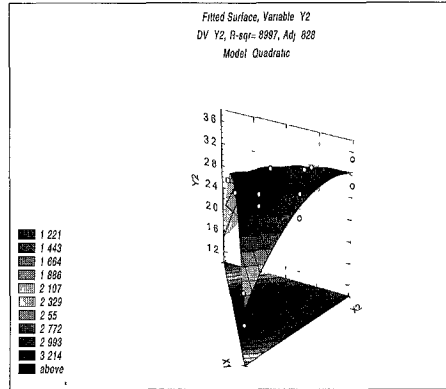
$$\begin{aligned} \hat{y}_2 = & 1.6641x_1 + 3.1352x_2 + 0.8800x_3 \\ & + 2.7378x_1x_2 + 8.4355x_1x_3 + 2.5405x_2x_3 \end{aligned}$$

<표4> Model Summary Statistics (y2)

EXPERIM. DESIGN	SS Effect	df Effect	MS Effect	SS Error	df Error
Linear	1.365	2	0.6823	4.185	10
Quadratic	3.628	3	1.2094	0.557	7
Special	0.001	1	0.0009	0.556	6
Cubic					
Total Adjusted	5.549	12	0.4624		
EXPERIM. DESIGN	MS Error	F	p	R-Sqr	R-Sqr Adj.
Linear	0.418	1.630	0.244	0.2459	0.095
Quadratic	0.080	15.208	0.002	0.8997	0.828
Special	0.093	0.010	0.925	0.8999	0.800
Cubic					

을 알 수 있다.

<그림6> Three-dimensional surface plot(y2)



<표5> Lack of Fit Tests (y2)

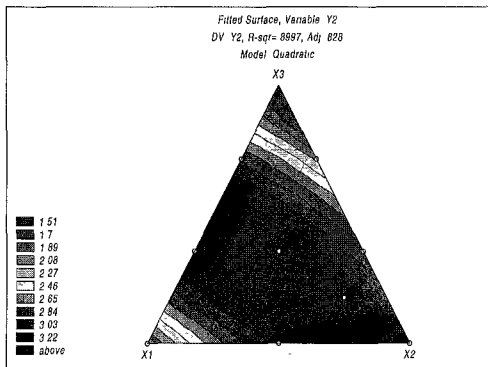
MODEL	SS	df	MS	F	p
Linear	3.631	6	0.605	4.372	0.087
Quadratic	0.003	3	0.001	0.007	0.999
Special	0.002	2	0.001	0.008	0.992
Cubic					
Pure Error	0.554	4	0.138		

4. 다중 반응표면의 최적화

Desirability function[1]을 이용하여 최소화 반응 변수인 y_1 과 최대화 반응변수인 y_2 를 동시에 만족시키는 최적 혼합비율을 찾아보고자 한다.

y_1 과 y_2 가 모두 one-sided bounds이므로, 앞절의 개별적인 분석결과에 의거 y_1 과 y_2 의 desirable range를 표 6과 같이 선택하였다.

<그림5> Response surface contour plot(y2)



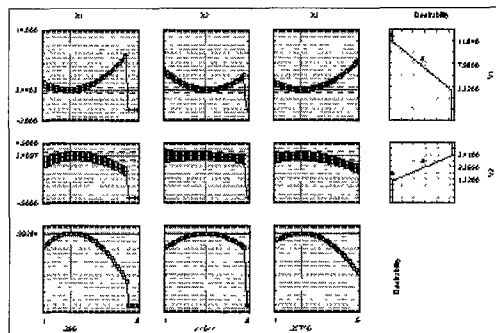
<표6> Desirability function parameters

	low value	desirability	med. value	desirability	high value	desirability
y_1	3.32	1	7.58	0.5	11.84	0
y_2	1.32	0	2.365	0.5	3.41	1

y_2 에 대한 contour plot와 surface graph인 그림 5와 그림 6으로부터 설탕의 혼합비율 x_2 는 향의 선호도 y_2 에 큰 영향을 주지 않으며, 정산점이 최대점임을 확인할 수 있다.

여기서도 y_2 만을 고려하여 최적 혼합비율을 찾는다면 $x_1 = 0.33865$, $x_2 = 0.31$, $x_3 = 0.35135$ 이 됨

<그림7> Desirability function을 이용한 최적화



desirability d_i 의 weights(r,s)를 1로 하고 x_1, x_2, x_3 에 대한 grid를 50으로 하여(그림 8 참조) 최적 혼합비율을 구한 결과, $x_1 = 29.6\%$, $x_2 = 44.6\%$, $x_3 = 25.8\%$ 이었다. 이것을 기준 혼합량에 따라 환산하면 커피 = 4.79g/110ml, 설탕 = 7.14g/110ml, 프립 = 4.13g/110ml이 된다.

5. 결론 및 향후과제

혼합물 실험이 여러 번의 반복된 실험과 시행착오를 거쳐서 이루어지는 경우가 있는데, 이는 경제적인 측면에서 바람직하지 않다. 본 논문에서는 혼합물실험을 계획할 때 고려될 수 있는 실험상의 여러 제약조건이나 복수의 반응치를 갖는 상황에서 적은 횟수의 실험을 통해 실효성 있는 최적 혼합비율에 접근하는 방법을 소개하였다. 본 논문에서 소개된 제약조건을 갖는 다중 반응표면의 혼합물실험의 예는 반응표면의 모형화 과정과 이러한 실험을 계획하는 여타의 산업분야에서 별다른 수정 없이 활용될 수 있을 것이다.

그러나 반응치들 사이의 상관관계가 유의한 경우에는 최적화 결과의 신뢰성 및 로버스트성에 대해 특별한 주의가 요구된다. 앞으로 복수의 반응치(multiple responses) 간에 무시할 수 없는 상관관계가 존재하는 경우의 목적함수의 설정 및 최적화 방법에 대한 연구가 필요할 것으로 생각된다.

참고 문헌

- [1] G. Derringer and R. Suich, "Simultaneous Optimization of Several Response Variables", *Journal of Quality Technology* vol.12, No.4, pp.214-219, 1980
- [2] J. A. Cornell, *Experiments with Mixtures*, John Wiley & Sons, Inc. 1990
- [3] M L Bowles and D. C. Montgomery, "How To Formulate The Ultimate Magarita: A Tutorial on Experiments with Mixtures", *Quality Engineering* vol.10, No.2, pp.239-253, 1997
- [4] R. H. Myers and D. C. Montgomery, *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, John Wiley & Sons, Inc. 1995
- [5] R. D. Snee, "Experimental Designs for Quadratic Models in Constrained Mixture Spaces", *Technometrics*, vol.17, No.2, pp.149-159, 1975
- [6] T. J. Mitchell and C. K. Bayne, "D-optimal Fractions Three-Level Factorial Designs", *Technometrics*, vol.20, No.4, pp.369-383,