



## 다변량 통계 분석법을 이용한 2성분계 혼합물의 인화점 예측

김성영 · †이범석 · 정창복\* · 최수형\*\*

경희대학교 화학공학과, \*전남대학교 응용화학공학부, \*\*전북대학교 화학공학부  
(2006년 9월 29일 접수, 2006년 12월 13일 채택)

## Prediction of Flash Point of Binary Systems by Using Multivariate Statistical Analysis

S. Y. Kim · †Bomsok Lee · C. B. Chung\* · S. H. Choi\*\*

*Dept. of Chemical Engineering, KyungHee University*

*\*Faculty of Applied Chemical Engineering, Chonnam National University*

*\*\*Faculty of Chemical Engineering, Chonbuk National University*

*(Received 29 September 2006, Accepted 13 December 2006)*

### 요 약

화학공정 설계에서 공정의 위험성 판단은 중요한 부분이다. 실제 화학공정에 사용되는 가연성 물질의 화재 및 폭발 위험성을 판단하는 인화점에 대한 예측은 그 방법 중의 하나이다. 본 연구에서는 2성분계 가연성 물질의 인화점에 대한 실험 자료를 이용하여 다변량 통계 분석법(partial least squares(PLS), quadratic partial least squares(QPLS))을 이용하여 2성분계 혼합물의 인화점을 예측하였고, 기존의 Raoult의 법칙과 Van Laar 식에 의한 예측값과 비교해 보았다.

**Abstract** – Estimation of process safety is important in the chemical process design. Prediction for flash points of flammable substances used in chemical processes is the one of the methods for estimating process safety. Flash point is the property used to examine the potential for the fire and explosion hazards of flammable substances. In this paper, multivariate statistical analysis methods(partial least squares(PLS), quadratic partial least squares(QPLS)) using experimental data is suggested for predicting flash points of flammable substances of binary systems. The prediction results are compared with the values calculated by laws of Raoult and Van Laar equation.

**Key words** : Flash point, Multivariate statistical analysis, Partial least squares, Quadratic partial least squares

### I. 서 론

화학공정에 사용되는 가연성 물질들의 인화점 예측은 화학공정의 위험성 검사를 위해 매우 중요하다. 이는 실제 공정에서 취급 부주의로 인한 화재 및 폭발의 위험성을 지니고 있기 때문이다.

인화점은 공기 중에서 연소할 수 있는 혼합물을 형성하는 충분한 증기를 발산하는 물질의 온도로, 휘발성 물질에서 발생하는 증기압이 주위의 공기와 가연성 혼합 기체를 만들만큼 높아졌을 때의 최저온도로 정의된다.

순수한 물질의 인화점은 MSDS(Material Safety Data

Sheets), SFPE(Society of Fire Protection Engineers) handbook[10] 등의 자료를 통해 확인할 수 있지만, 실제 공정에서 생산되는 혼합물의 인화점은 조성의 변화가 다양하고 측정에 많은 시간이 요구되므로 인해 안전에 중요한 연소 특성치 임에도 불구하고 확인하기가 어렵다. Affens 등[4]이 Raoult의 법칙을 이용하여 2성분계 탄화수소 용액의 인화점을 예측하였고, Thome[5]은 Van Laar 식과 Clausius-Clapeyron 식을 통해 활동도 계수를 도입하여 혼합물의 인화점을 예측하였다. Gmehling 등[6]은 활동도 계수의 계산에 UNIFAC 기법을 도입하였다. 최근 Ha 등[7]은 밀폐식 장치를 이용하여 측정된 2성분계의 인화점을 Raoult의 법칙과 Van Laar 식을 이용해 추정하였고, Liaw 등[8]은 활동도 계수의 계산에 Wilson, NRTL, UNIQUAC 식을 적용하여

†주저자: bslee@khu.ac.kr

혼합물의 인화점을 예측하고 서로 비교하는 연구를 발표하였다.

본 논문에서는 위 연구들에서 사용한 수식 모델들과는 다르게, 증기압, 활동도 계수, 폭발하한계, 혼합물의 조성등의 자료들을 partial least squares(PLS), quadratic partial least squares(QPLS) 등의 다변량 통계 분석법[2,3]에 적용하여 기존의 수식 예측 모델의 계산을 단순화시키고 예측 신뢰성을 향상 시켰다. 다변량 통계 분석법은 입출력 데이터를 동시에 분석하는 기법으로 실제 실험 데이터에 의존한 실험적 모델링(empirical modeling)이다. PLS, QPLS 방법을 2성분계 혼합물의 인화점 실험데이터에 적용하여 회귀 모델을 만들고 새로운 값을 예측하였다. 이 결과들을 기존의 Raoult의 법칙과 Van Laar 식에 의해 예측된 결과와 실제 실험값과 비교하였다.

## II. 이 론

### 2.1. 다변량 통계 분석법

입출력 데이터가 복잡한 화학공정의 다변량 데이터를 분석하고 물성을 예측하는데 다변량 통계 분석법은 매우 유용하다. 다변량 통계 분석법은 입출력 데이터를 동시에 분석하는 통계적 기법으로 실제 실험데이터에 의존한 실험적 모델링(empirical modeling)이다. 사용되는 화학반응식이나 반응상수들이 많아서 예측 모델의 구성이 힘든 경우, 실제 실험데이터를 기반으로 모델을 세울 수 있는 장점이 있다. 본 논문에서는 다변량 통계분석법 중 PLS와 QPLS 방법을 사용하여 M.E.K. (methylethylketone) + Toluene, n-propanol + n-propionic acid[7,9] 등의 2성분계의 인화점 실험데이터를 기반으로 실험적 모델을 수립하였다.

### 2.2. Partial least squares(PLS) 방법

PLS 방법은 새로운 잠재 변수들(latent variables)을 찾아내 모델에 사용하는 방법으로 PLS에 사용되는 잠재변수들은 서로 선형적으로 독립이다. 입력 잠재변수와 출력 잠재변수 사이에는 매우 높은 상관관계를 갖게 되며 입력 변수 중에 분산(variance) 값이 높은 순서대로 잠재 변수로 채택된다[2]. 입력 변수 데이터 블록(X)은 score 행렬  $t_h$ 와 loading 행렬  $p_h$ 의 곱으로 나타내어지고, 출력 변수 데이터 블록(Y)은 score 행렬  $u_h$ 와 loading 행렬  $q_h$ 의 곱으로 표시되어진다.

$$X = \sum_{h=1}^a t_h p_h^T + E \quad (1)$$

$$Y = \sum_{h=1}^a u_h q_h^T + F \quad (2)$$

$$u_h = b_h t_h + e_h, \quad (h = 1, \dots, a) \quad (3)$$

식 (3)에서와 같이, PLS 방법은 입력 데이터 score 행렬과 출력 데이터 score 행렬의 선형적인 관계이다. PLS 방법은 위의 식을 이용한 NIPALS algorithm[2]에 의해 수행된다.

### 2.3. Quadratic partial least squares(QPLS) 방법

QPLS 방법은 PLS 방법을 수정하여 입출력 데이터의 비선형성을 모델에 반영한 방법이다[3]. PLS 방법과 같이 입출력 데이터를 통해 잠재 변수를 찾아내고, 입출력 score간의 관계를 식 (3)과 같이 선형으로 표현하는 것이 아니라, 식 (4)와 같이 2차 식으로 나타낸다.

$$u_h = c_{0,h} + c_{1,h} t_h + c_{2,h} t_h^2 + e_h \quad (h = 1, \dots, a) \quad (4)$$

QPLS 방법은 score간의 2차식을 이용해 NIPALS algorithm을 수정하여 수행된다[3]. PLS와 QPLS 방법은 실험 데이터 자료의 잠재 변수간의 관계에 따라 모델이 구성되고 실제 실험값과 근사한 값을 예측하게 된다. 그리고 실험하지 않은 새로운 데이터에 대해서도 뛰어난 예측성능을 보이는 장점을 갖고 있다.

## III. 2성분계 혼합물의 인화점 예측

2성분계 혼합물의 인화점 예측을 위해 M.E.K. (methylethylketone) + toluene, n-propanol + n-propionic acid의 조성 변화에 따른 인화점의 실제 실험데이터[7,9]를 이용하였다. 실험데이터 가운데 일부는 PLS, QPLS 방법을 이용한 회귀모델에 사용하고, 나머지 데이터는 모델을 이용한 예측값과의 비교에 사용하였다.

2성분계 혼합물의 증기압이 폭발하한계의 농도와 같은 때의 온도가 인화점이므로, 혼합물의 조성과 순수물질의 증기압, 폭발하한계 등이 인화점 예측에 중요한 자료들이다. 본 연구에서는 PLS, QPLS 방법에 사용할 입력 데이터 블록에 이러한 혼합물의 조성, 증기압과 관련된 Antoine constants, 폭발하한계, 연소열등을 사용하였고, 출력데이터에 인화점의 실제 실험값을 사용하였다. M.E.K. + toluene의 인화점 예측에는 PLS, QPLS 방법에 사용될 잠재변수를 2개 선택하였고, n-propanol + n-propionic acid의 인화점 예측에는 잠재변수를 3개 선택하였다.

**Table 1.** Comparison of flash point [°C] for M.E.K.(X1) + toluene(X2) system.

	Sample	Mole fraction		Ex.	Raoult	Van Laar	PLS	QPLS
		X1	X2					
Regression	1	0.0	1.0	5.0	1.98	1.98	5.65	5.43
	2	0.1	0.9	6.0	0.79	0.01	4.56	4.88
	3	0.2	0.8	3.0	-0.33	-1.60	3.48	3.60
	4	0.4	0.6	1.0	-2.40	-3.97	1.32	1.19
	5	0.6	0.4	-1.0	-4.26	-5.61	-0.84	-1.05
	6	0.8	0.2	-3.0	-5.95	-6.74	-3.01	-3.10
	7	1.0	0.0	-5.0	-7.49	-7.49	-5.17	-4.97
	Average of root square error					<b>3.38</b>	<b>4.20</b>	<b>0.46</b>
Prediction	8	0.3	0.7	2.0	-1.39	-2.90	2.40	2.38
	9	0.5	0.5	0.0	-3.35	-4.86	0.24	0.05
	10	0.7	0.3	-2.0	-5.13	-6.23	-1.92	-2.10
	11	0.9	0.1	-4.0	-6.74	-7.17	-4.09	-4.06
	Average of root square error					<b>3.15</b>	<b>4.29</b>	<b>0.20</b>

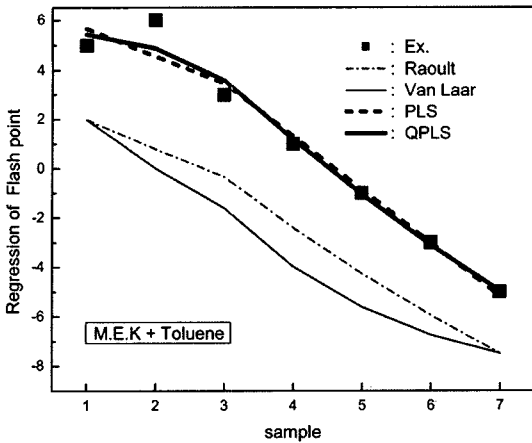
**Table 2.** Comparison of flash point [°C] for n-propanol(X1) + n-propionic acid(X2) system.

	Sample	Mole fraction		Ex.	Raoult	Van Laar	PLS	QPLS
		X1	X2					
Regression	1	1.000	0.000	21.0	21.36	21.36	21.00	21.11
	2	0.910	0.090	21.0	22.53	22.61	19.52	20.90
	3	0.710	0.290	25.0	25.58	26.03	25.79	24.85
	4	0.301	0.699	36.0	35.46	37.34	38.62	36.96
	5	0.000	1.000	50.0	53.67	53.67	48.07	49.18
	Average of root square error					<b>1.34</b>	<b>1.60</b>	<b>1.36</b>
Prediction	6	0.510	0.490	30.0	29.54	30.64	32.07	30.11
	7	0.109	0.891	42.0	44.48	46.35	44.65	44.41
	8	0.045	0.955	44.0	49.24	50.38	46.66	47.18
	Average of root square error					<b>2.73</b>	<b>3.79</b>	<b>2.46</b>

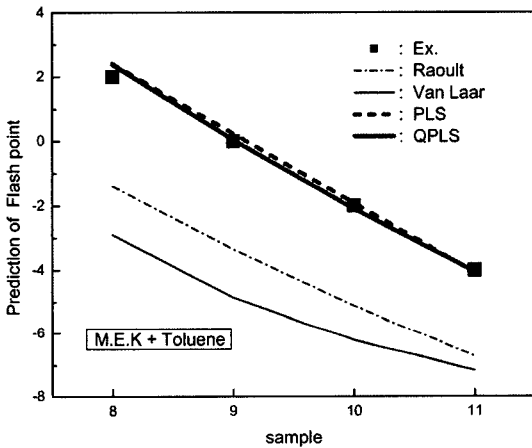
#### IV. 결과 및 고찰

다변량 통계 분석법(PLS, QPLS)을 이용한 회귀(regression)와 예측(prediction) 결과의 신뢰성을 확인하기 위해 Raoult의 법칙, Van Laar 식에 의한 계산값과 실제 인화점의 실험값과의 평균 제곱급 오차(root square error) 값을 이용하여 비교하였다. Table 1과 2에서는 각각 M.E.K. + toluene, n-propanol + n-propionic acid에 대한 회귀, 예측 결과들을 나타내었다.

M.E.K.+toluene에서는 기존의 Raoult의 법칙, Van Laar 식에 의한 계산값보다 PLS, QPLS 방법에 의한 회귀, 예측 결과가 뛰어난 것을 확인할 수 있다. n-propanol + n-propionic acid에서, PLS 방법의 결과는 회귀의 경우 Raoult의 법칙에 의한 계산값과 비슷한 결과를 나타내었으나, 회귀모형을 이용한 예측에서는 Raoult의 법칙에 의한 계산값보다 개선됨을 볼 수 있다. 이에 비해, QPLS 방법의 경우는 다른 방법들에 비해 뛰어난 예측, 회귀 결과를 보여주고 있다. 이는 Table 1과 2의 각각

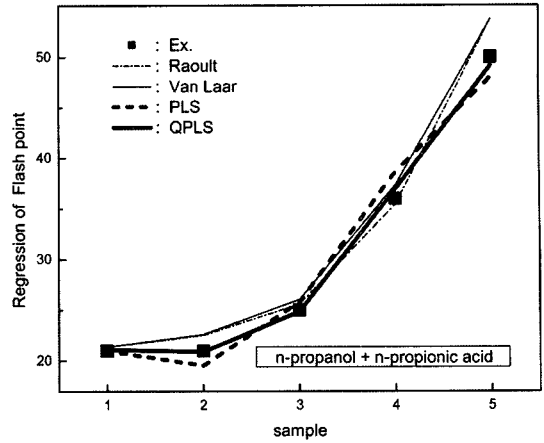


(a) Regression

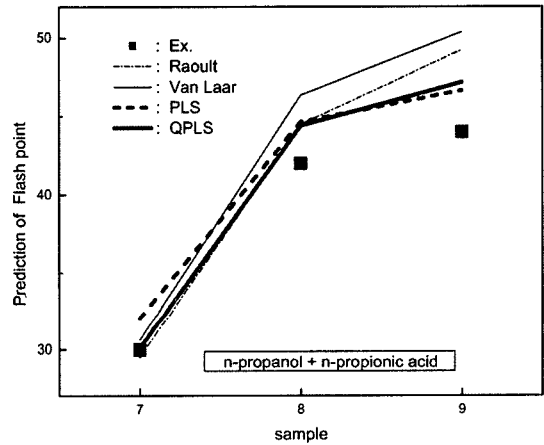


(b) Prediction

Fig. 1. Comparison of flash point for M.E.K. + toluene system.



(a) Regression



(b) Prediction

Fig. 2. Comparison of flash point for n-propanol+n-propionic acid system.

의 평균 제곱근 오차 값에서 확인된다. Fig. 1과 2에서는 인화점의 실제 실험값, Raoult의 법칙과 Van Laar 식에 의한 계산값, PLS, QPLS 방법에 의한 예측값을 비교한 결과를 나타내었다.

이와 같이, 다변량 통계 분석법에 의한 인화점 예측 결과가 Raoult의 법칙과 Van Laar 식에 의한 계산값보다 개선되었음을 확인할 수 있고, 다변량 통계 분석법 중에서는 잠재 변수들 사이의 비선형적 관계의 영향으로 PLS 방법보다 QPLS 방법이 더 좋은 예측 성능을 보임을 확인할 수 있다.

## V. 결 론

2성분계 혼합물의 인화점에 대한 실험데이터를 이용

해 다변량 통계 분석법(PLS, QPLS)을 적용하여 인화점 예측을 수행하였다. PLS, QPLS 방법을 통해 몇 가지 경우의 조성에 대한 실험 데이터를 이용한 회귀모델을 만듦으로서 다양한 새로운 조성에 대한 인화점의 실험값을 예측할 수 있었다. 기존의 연구에서 발표된 수식 모델에 의한 인화점 예측에 비해 모델의 계산이 간편하고, 결과 또한 Fig. 1과 2, Table 1과 2에서와 같이 기존의 Raoult의 법칙, Van Laar 식에 의한 계산값보다 뛰어난 결과를 확인할 수 있었다.

## 감사의 글

본 논문은 한국과학재단(과제번호 R01-2003-000-10697-0)에서 지원하여 연구하였습니다.

### 참고문헌

- [1] Crowl, D.A. and J.F. Louvar, *Chemical Process Safety : Fundamentals with Applications*, Prentice Hall PTR, (2002)
- [2] Geladi, P. and B.R. Kowalski, "Partial Least Squares Regression : A Tutorial", *Analytica Chimica Acta*, **185**, 1-17, (1986)
- [3] Baffi, G., E.B. Martin, and A.J. Morris, "Non-linear Projection to Latent Structures Revisited : The Quadratic PLS Algorithm", *Computers and Chemical Engineerings*, **23**, 395-411, (1999)
- [4] Affens, W.A. and G.W. McLaren, "Flammability Properties of Hydrocarbon Solutions in Air", *J. of Chem. Eng. Data*, **17**(4), 482-488, (1972)
- [5] Thorne, P.F., "Flash Point of Mixtures of Flammable and Non-flammable Liquids", *Fire and Materials*, **1**, 134-140, (1977)
- [6] Gmehling, J. and P. Rasmussen, "Flash Point of Flammable Liquid Mixtures Using UNIFAC", *Industrial and Engineering Chemistry Fundamentals*, **21**(2), 186-188, (1982)
- [7] Ha, D.M., S.J. Lee, Y.C. Choi, and H.J. Oh, "Measurement of Flash Points of Binary Systems by Using Closed Cup Tester", *HWAHAK KONGHAK*, **41**(2), 186-191, (2003)
- [8] Liaw, H. and Y. Chiu, "A General Model for Predicting the Flash Point of Miscible Mixtures", *J. of Hazardous Materials*, **137**(1), 38-46, (2006)
- [9] Ha, D.M., Y.S. Mok, and J.W. Choi, "Flash Points of a Flammable Liquid Mixture of Binary System", *HWAHAK KONGHAK*, **37**(2), 146-150, (1999)
- [10] SFPE, *The SFPE Handbook of Fire Protection Engineering*, 2nd ed., Society of Fire Protection Engineers, Boston, (1995).
- [11] Kim, S.Y., B.S. Lee, C.B. Chung, and S.H. Choi, "Multivariate Statistical Analysis Approach to Predict the Reactor Properties and the Product Quality of a Direct Esterification Reactor for PET Synthesis", *J. of Control, Automation and Systems Engineering*, **11**(6), 550-557, (2005)