

폐액 중 프로필 글리콜 모노메틸 에테르 아세테이트(PGMEA) 회수하는 증류공정에서 회귀분석을 이용한 공정 최적화

최용석 · 변현수[†]

전남대학교 화공생명공학과
550-749 전남 여수시 대학로 50
(2014년 5월 28일 접수, 2014년 6월 16일 수정본 접수, 2014년 6월 20일 채택)

Process Optimization Using Regression Analysis of Distillation Processes for the Recovery of Propylene Glycol Monomethyl Ether Acetate (PGMEA) Containing Waste Organic Solvent

Yong-Seok Choi and Hun-Soo Byun[†]

Department of Chemical and Biomolecular Engineering, Chonnam National University, 50 Daehak-ro, Yeosu, Jeonnam 550-749, Korea
(Received 28 May 2014; Received in revised form 16 June 2014; accepted 20 June 2014)

요 약

본 연구는 Liquid Crystal Display (LCD) 세척 후 발생하는 Propylene Glycol Monomethyl Ether Acetate (PGMEA) 폐액 재활용을 위한 2기 증류탑을 사용하는 공정의 최적 조건을 산출하였다. 공정조건 최적화 기법으로 다중회귀분석을 이용하여 1차 증류 시 Bottom 온도(BTM 온도), Reflux 량, Feed 량 및 Feed 온도에 따라 PGMEA 함량에 대한 최적 조건과 2차 증류 시 BTM 온도, Reflux 량, Feed 량에 따른 PGMEA 함량에 대한 최적 조건과 공정인자를 산출하였다. 1차 증류탑의 공정인자 중 Reflux 량, Feed 온도 및 Feed 량이 중요한 인자로 산출되었다. 본 연구의 공정조건 범위에서는 BTM 온도범위가 PGMEA 함량에 크게 영향을 주지 못하였다. 따라서 최적 공정 조건은 Feed 량 5,700 ℓ, Reflux 량 2,500 ℓ, BTM 온도 165 °C 및 Feed 온도 130 °C이며 이때 예측된 PGMEA 함량은 92.12~94.62%로 산출되었다. 2차 증류탑에서는 Reflux 량이 함량에 많은 영향을 미치고 있으며, Feed 량과 BTM 온도도 영향을 미치는 인자로 산출되었다. 다중공선성(Multicollinearity)이 Reflux 량과 BTM 온도 간에 강한 양의 상관관계가 있어, 두 인자 중 다중회귀식에 영향이 지배적인 인자 하나를 선택하였으며, 최적조건은 BTM 온도 199 °C 기준에서 최적 공정 조건은 Feed 량 4,275 ℓ 및 Reflux 량 6,200 ℓ이며, 이때 예측 PGMEA 함량은 99.0~99.5%로 산출되었다.

Abstract – The aim of this study is to obtain optimum process condition for using two tower distribution to recycle the waste Propylene Glycol Monomethyl Ether Acetate (PGMEA) that is formed after washing LCD. The optimum process condition for the content of PGMEA, which is dependent variable, at 1st distillation was calculated according to Bottom temperature (BTM temperature), Reflux amount, Feed amount, Feed temperatures, and the optimum process conditions and optimum factors for the content of PGMEA at 2nd distillation according to Bottom temperature (BTM temperature), Reflux amount, Feed amount, Feed temperatures. At 1st distillation, Reflux amount, Feed amount, and Feed temperature are significant variables. However, it is found that the BTM temperature range is not significant in the range of process condition used in this study. The optimum process conditions are based on 5700 ℓ of Feed amount, 2500 ℓ of Reflux amount, 165 °C of BTM temperature, and 130 °C of Feed temperature. For the this condition, the predicted content of PGMEA was calculated as 92.12~94.62%. Significant factors at 2nd distillation are Reflux amount, Feed amount, and BTM temperature. Multicollinearity is between Reflux amount and BTM temperature. BTM was omitted in the multiple regression equation because there is a strong positive correlation between Reflux amount and BTM temperature. Base on 199 °C of BTM temperature, The optimum process conditions are based on 4275 ℓ of Feed amount, 6200 ℓ of Reflux amount and 130 °C of Feed temperature. In this condition, the predicted content of PGMEA was calculated as 99.0~99.5%.

Key words: Multiple Regression Analysis, Propylene Glycol Monomethyl Ether Acetate, Distillation Processes, Optimum Process Conditions, Waste Organic Solvent

[†]To whom correspondence should be addressed.

E-mail: hsbyun@jnu.ac.kr

*이 논문은 한양대학교 배성열 교수님의 정년을 기념하여 투고되었습니다.

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. 서 론

Propylene Glycol Monomethyl Ether Acetate(PGMEA)는 TFT-LCD(Thin Film Transistor Liquid Crystal Display(LCD))에 사용하고 있는 화학 전자재료 용제중 대표적인 것으로 일본과 중국으로부터 수입하여 사용한다. PGMEA는 알콜계 용제에 비하여 휘발성이 매우 적으며, 인화점도 높고 유리나 박막재료에 손상을 주지 않기 때문에 LCD뿐만 아니라 유기 박막트랜지스터의 절연막에 용매로 사용하고 있으며 세척용 용매로도 사용되고 있다. 인체에 누적되지 않는 친환경 용제로 페인트나 잉크의 희석제로도 이용된다[1]. PGMEA를 비롯한 화학 전자재료 제품은 수입에 의존하는 고가 제품이라 국내 수급의 어려움과 고가의 제품 원가로 인하여 2중고를 국내 기업은 겪고 있다. 이에, 2006년부터 화학 전자재료에 대한 재활용 연구가 활발히 이루어지고 있으며 특히, 종류에 의한 재활용 기술의 발달로 상용화가 이루어져 원가절감에 기여하고 있다. 무한 경쟁시대에는 기술 상용화를 뛰어넘어 최적화를 통하여 원가절감을 하지 않으면 생존하기 어려운 시대이다[2]. 필연적인 과제로 초기 조건을 찾기 위한 손실과 작업자에 의한 경험적 운전 조건 설정은 보이지 않는 손실이며 무한 경쟁시대에서는 극복하여야 할 문제이다. 이에, 품질은 유지하면서 수율 향상을 시키기 위해서는 공정 조건의 최적화라는 통계적 문제로 귀착된다. 본 연구에서는 최적화 기법 중 다중회귀분석을 적용하였다[3,4].

재활용 증류 공정은 폐액의 종류와 상태, 설비 차이로 다양한 공정 조건으로 운전되고 있다. 경험적 요소와 공학적 고유기술로는 단시간에 미지의 물질에 대한 공정조건을 찾는다는 것은 어렵다. 산업현장에는 6M(Man, Material, Machine, Method, Measurement, Mother Nature)으로 인하여 고유기술을 기반으로 통계적으로 접근하지 않으면 보이지 않는 손실이 발생하고 경영에 어려움을 발생시킨다.

본 연구의 목적은 현장에서 적용한 공정 자료로부터 최적의 조건을 산출하고 공정 변경시 응용 가능한 다중회귀식을 산출하는데 있다. 어떤 원인인자가 결과값에 영향을 미치는지를 확인할 수 있고 다중회귀식의 시뮬레이션을 통하여 새로운 공정 조건을 찾을 수 있을 것이다. 따라서 본 연구는 LCD TFT 공정에서 세척용으로 사용되는 PGMEA 폐액을 2기 증류탑을 사용하여 순도 99% 이상 PGMEA를 회수하는 공정이다. 즉, 처음 1기의 증류탑은 수분 등의 불순물을 Overhead로 분류해 내고 약 93% 이상의 PGMEA를 Bottom으로 회수하고 다음 증류탑은 Overhead로 99% 이상의 PGMEA를 회수하여 제품화하는 공정으로 한다. 자료의 분석기법으로는 회귀 분석을 사용하였으며 원인인자는 BTM 온도, Feed 량, Feed 온도, Reflux 량을 선택하여 결과값인 PGMEA 함량에 영향을 미치는 인자(유의 인자)를 산출하고 최적화 다중회귀식을 나타내었으며, 이를 통하여 경제성을 고려한 최적의 조건들을 얻었다. 이에 따라 순성분 PGMEA의 화학적 물성은 Table 1과 같다.

2. 회귀분석의 이론적 배경

예측되는 변수인 종속변수(dependent variable)와 예측에 사용하는 변수인 독립변수(independent variable)의 관계를 설명하기 위해

Table 1. Chemical properties of PGMEA

CAS NO.	Flash point	Boiling Point	Auto ignition	Chemical Formula
108-65-6	42 °C	145~147 °C	315 °C	C ₆ H ₁₂ O ₃

회귀분석(regression analysis)이라는 통계적 방법을 활용하였다. 하나의 독립변수 x 와 하나의 종속변수 y 가 있고 두변수 사이의 관계는 직선에 가까운 가장 단순한 형태의 회귀를 단순선형회귀모형(simple linear regression model)이라고 한다. 독립변수가 둘 이상인 회귀를 다중선형회귀모형(multiple linear regression model)이라고 한다[5].

2-1. 단순선형회귀모형

변수 x 와 y 간에 단순선형회귀모형을 적합 시킬 경우에는 다음과 같은 가정이 전제조건을 이루고 있다.

첫째, y 와 x 사이에 존재하는 관련성은 주어진 x 의 값에서 y 의 기댓값을 $E(y)$ 라고 할 때

$$E(y) = \beta_0 + \beta_1 x \quad (1)$$

와 같은 선형식으로 표현될 수 있다.

둘째, 주어진 x 의 값에서 y 는 정규분포(normal distribution)를 하며, 평균은 $\beta_0 + \beta_1 x$ 로 x 에 따라서 변하나 분산은 x 값에 관계없이 일정하다.

셋째, 독립변수 x 는 오차 없이 측정할 수 있는 수학변수(mathematical variable)이며, 종속변수 y 는 측정오차를 수반하는 확률변수(random variable)이다. 또 변수 y 의 측정오차들은 서로 독립이다. 위의 가정에서 단순선형회귀모형을 표현하면 다음과 같다.

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad (2)$$

여기서 y_i 는 i 번째 측정된 y 의 값, β_0 와 β_1 은 회귀 모형의 모수이며, x_i 는 i 번째 측정된 x 의 값이며, ε_i 는 i 번째 측정된 y 의 잔차항으로 확률분포는 $N(0, \sigma^2)$ 와 y 사이의 선형관계로 설명될 수 없는 변동성(variability)이다. 잔차항은 4가지 원칙을 따른다. ① 불편성으로 $E(\varepsilon)=0$ 이며 ② 정규성으로 $\varepsilon \sim N(0, \sigma_\varepsilon^2)$ 이며 ③ 독립성으로 ε 값들은 독립이고 ④ 등분산성으로 $V(\varepsilon') = V(\varepsilon'') = \sigma_\varepsilon^2$ 이다.

모형 (2)의 표현에서 첨자 i 를 빼고 종속변수의 값을 일반형으로 나타내면 단순선형회귀모형은 다음과 같다.

$$y = \beta_0 + \beta_1 x + \varepsilon \quad (3)$$

단순선형회귀식 (1)의 그래프는 직선이다. β_0 는 회귀식의 y 절편이고 β_1 은 기울기이다. $E(y)$ 는 주어진 x 값에 대한 y 의 기댓값, 평균이다.

모집단의 모수인 β_0 와 β_1 의 값이 알려져 있다면, 주어진 x 값에 대한 y 의 평균값을 계산하는데 식 (1)을 사용하면 되나 현실적으로 모수값은 미지이므로 표본 자료를 이용하여 추정한다. 표본 통계량 b_0 와 b_1 은 모집단의 모수인 β_0 와 β_1 의 추정값이다. 회귀식에 있는 β_0 와 β_1 에 표본 통계량의 값인 b_0 와 b_1 을 대입하면, 추정 회귀식(estimated regression equation)이 얻어진다.

위에서 설명한 단순선형모형에서 단순추정회귀식을 유도하는 과정을 나타낸 흐름도를 Fig. 1에 나타내었다.

2-2. 단순선형회귀식의 추정

단순선형회귀모형을 위한 추정 회귀식은 다음과 같다.

$$\hat{y} = b_0 + b_1 x \quad (4)$$

여기서, b_0 는 추정 회귀식의 \hat{y} 절편이고 b_1 은 기울기이다.

y 절편과 b_1 기울기를 추정하기 위한 기법으로 최소제곱법을 활용한다. 표본자료를 이용하여 y_1 의 관측치와 종속변수 추정치 \hat{y}_1 사

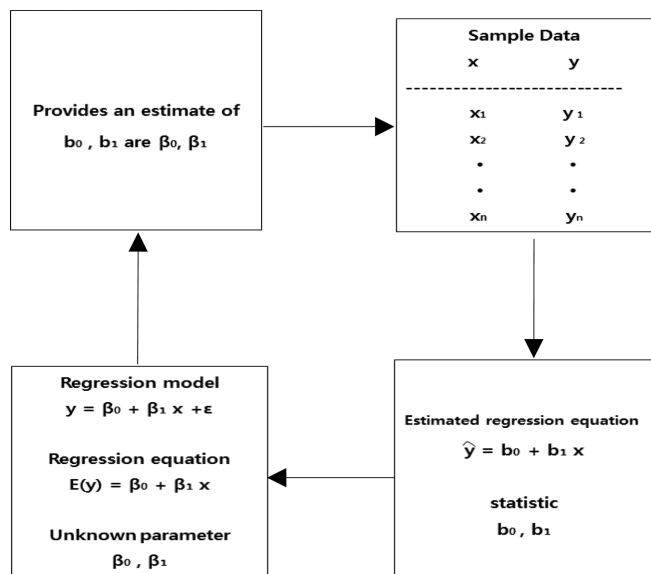


Fig. 1. The estimation process of simple linear regression.

이의 편차 제곱합을 최소화 시켜주는 b_0 와 b_1 을 찾아내는 방법이다.

모든 점 y_1, y_2, \dots, y_n 과 추정값 $\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n$ 사이의 편차 제곱합의 식은 다음과 같다.

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - b_0 - b_1 x_i)^2 \quad (5)$$

최소제곱법은 편차 제곱합이 최소가 되도록 b_0 와 b_1 을 정한다. 식 (5)은 최소값이 되기 위해서는 미지수 b_0 와 b_1 의 편미분 값이 0이 되어야 한다. b_0 와 b_1 에 대한 편차 제곱합의 편미분은 다음과 같다.

$$b_0 \text{에 대한 편미분} = \frac{d\left(\sum_{i=1}^n \epsilon_i^2\right)}{db_0} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 x_i) = 0 \quad (6)$$

$$b_1 \text{에 대한 편미분} = \frac{d\left(\sum_{i=1}^n \epsilon_i^2\right)}{db_1} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 x_i) x_i = 0 \quad (7)$$

식 (6)과 (7)를 b_0 와 b_1 에 대해 정리하면 다음과 같다.

$$기울기 \ b_1 = \frac{\sum_{i=1}^n x_i y_i - \left(\sum_{i=1}^n x_i\right) \left(\sum_{i=1}^n y_i\right) / n}{\sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2 / n} = \frac{S_{xy}}{S_{xx}} \quad (8)$$

$$y \text{ 절편 } b_0 = \frac{\sum_{i=1}^n y_i - b_1 \sum_{i=1}^n x_i}{n} = \bar{y} - b_1 \bar{x} \quad (9)$$

Fig. 2는 단순선형회귀모형을 나타내었고 x_1 에서 y_1 의 추정값은 $b_0 + b_1 x_1$ 이며 추정값과 실측값의 편차는 ϵ_1 이다. x_2 에서 y_2 의 추정

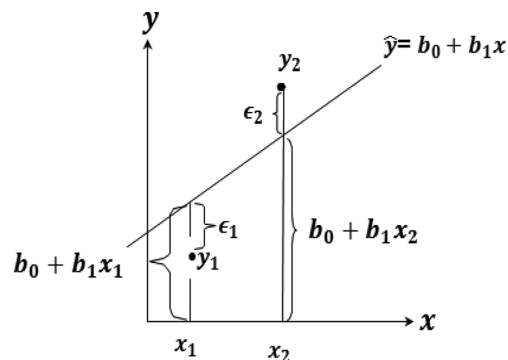


Fig. 2. Simple linear regression model.

값은 $b_0 + b_1 x_2$ 이며 추정값과 실측값의 편차는 ϵ_2 이다. 임의의 x 에서 y 의 추정값은 $b_0 + b_1 x$ 이다. 편차 제곱합인 $\epsilon_1^2 + \epsilon_2^2 + \dots + \epsilon_n^2$ 가 최소가 될 때 회귀식은 실측값을 가장 잘 나타낸다.

2-3. 결정계수(Coefficient of determination: R^2)

결정계수 R^2 은 추정 회귀식의 적합도를 평가하는 계수로 $0 \leq R^2 \leq 1$ 의 값을 가지며 1에 근접할수록 적합도가 좋다. 현실적인 문제로 사회과학의 전형적인 자료에 대해서는 R^2 의 값이 0.25보다 낮아도 유용하다고 간주한다. 물리학이나 생명과학에서는 R^2 의 값이 0.6 이상인 경우도 많다.

편차를 분해하면 식 (10)과 같다.

$$(y_i - \bar{y}) = (y_i - \hat{y}_i) + (\hat{y}_i - \bar{y}) \quad (10)$$

식 (10)의 양변을 제곱하여 정리하면 식 (11)이 산출된다.

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 \quad (11)$$

식 (11)에서 $\sum_{i=1}^n (y_i - \bar{y})^2$ 을 총 제곱합(total sum of squares)이라고 부르고 SST로 표기하며 $\sum_{i=1}^n (y_i - \hat{y}_i)^2$ 을 잔차 제곱합(sum of squares due to error)이라고 부르고 SSE로 표기하고 $\sum_{i=1}^n (\hat{y}_i - \bar{y})^2$ 은 회귀 제곱합(sum of squares due to regression)이라고 부르며 SSR로 표기한다. 식 (11)을 다음과 같이 표기한다.

$$SST = SSR + SSE \quad (12)$$

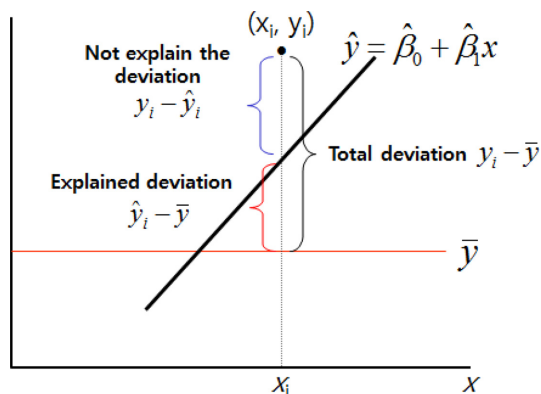


Fig. 3. Description of the deviation and the deviation of the total unexplained variation decomposes.

식 (12)은 통계학에서 가장 중요한 결과 중의 하나이다.

Fig. 3은 총 편차 ($y_i - \bar{y}$)를 설명할 수 있는 편차 ($\hat{y} - \bar{y}$)와 설명할 수 없는 편차 ($y_i - \hat{y}$)로 분해하여 나타내었다. 편차의 합은 0 이므로 각각을 제곱하여 합한다. 설명할 수 있는 편차 제곱합은 회귀 제곱합이며 설명할 수 없는 편차 제곱합은 잔차 제곱합이다. 총 제곱합은 회귀 제곱합과 편차 제곱합의 더한 것이다.

이제 SST, SSR, SSE이 추정 회귀식의 적합성의 척도로 이용된다. 만약 종속변수 y_i 의 모든 값이 추정 회귀선 위에 위치한다면 추정 회귀식은 완벽한 적합을 이루게 된다. 즉, SSR/SST의 비율은 1이 된다. 적합이 가장 나쁜 SSR=0 인 경우 SSR/SST의 비율은 0이 된다. 따라서 0과 1사이의 값을 갖는 SSR/SST 비율은 추정 회귀식의 적합도를 평가하는데 이용된다. 이 비율을 결정계수(coefficient of determination)라고 부르며 R^2 로 표기한다.

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} \quad (13)$$

2-4. 상관계수(Coefficient of correlation: R)[6]

두 변수 x, y 사이의 상관관계를 설명하는 데 결정계수가 쓰이기도 하지만, 결정계수는 x 와 y 의 관계가 음의 상관관계인지 양의 상관관계인지를 구별하지 못하는 단점을 가지고 있다. 이와 같은 단점을 보완하기 위하여 두 변량간의 상호관계를 측정하는 척도로서 상관계수라고 하는 것이 있다. 단순회귀분석에서는 상관계수 R 을 다음과 같이 구할 수 있다.

$$R = \pm\sqrt{R^2} \quad (13)$$

식 (13)은 단순회귀분석에서는 x 와 y 의 함수관계가 식 (3)에 의해서 표현되고, 이모형의 기본 가정의 하나로서 y 는 확률변수이나 x 는 수확변수로 가정한 후에 유도된 상관계수이다.

x 도 확률변수이고 x 와 y 가 어떤 이변량분포(bivariate distribution)를 하고 있다면 x 와 y 의 모집단상관계수(population coefficient of correlation)는 다음에 의해 정의한다.

$$\rho_{xy} = \frac{COV(x, y)}{\sqrt{Var(x)Var(y)}} \quad (14)$$

변수 x, y 의 분산 $Var(x), Var(y)$ 을 각각 σ_x^2, σ_y^2 으로 표시하고 x 와 y 의 공분산을 σ_{xy} 로 표시하면 식 (14)의 모집단의 상관계수는

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \cdot \sigma_y} \quad (15)$$

로 나타내진다.

모집단으로부터 크기가 n 인 표본을 샘플링 하였을 때 n 개의 자료 $(x_1, y_1), (x_2, y_2) \dots (x_n, y_n)$ 이 얻어졌다면 이 두 변수 사이의 표본상관계수(Sample coefficient of correlation)는

$$R_{xy} = \frac{S_{(xy)}}{\sqrt{S_{(xx)} \cdot S_{(yy)}}} \quad (16)$$

으로 정의된다. 상관계수는 두 변수간의 선형관계(Linear relationship)가 어느 정도인가를 재는 척도이지 함수관계를 알아보는 척도는 아니다.

Fig. 4는 x 와 y 의 선형관계를 나타낸 산점도이다. (a)는 $R=1$ 인 양의 완전한 상관관계이고 (b)는 $R=-1$ 인 음의 완전 상관이다. (c)는 $0 < R < 1$ 인 양의 상관관계이고 (d)는 $-1 < R < 0$ 인 음의 상관관계이

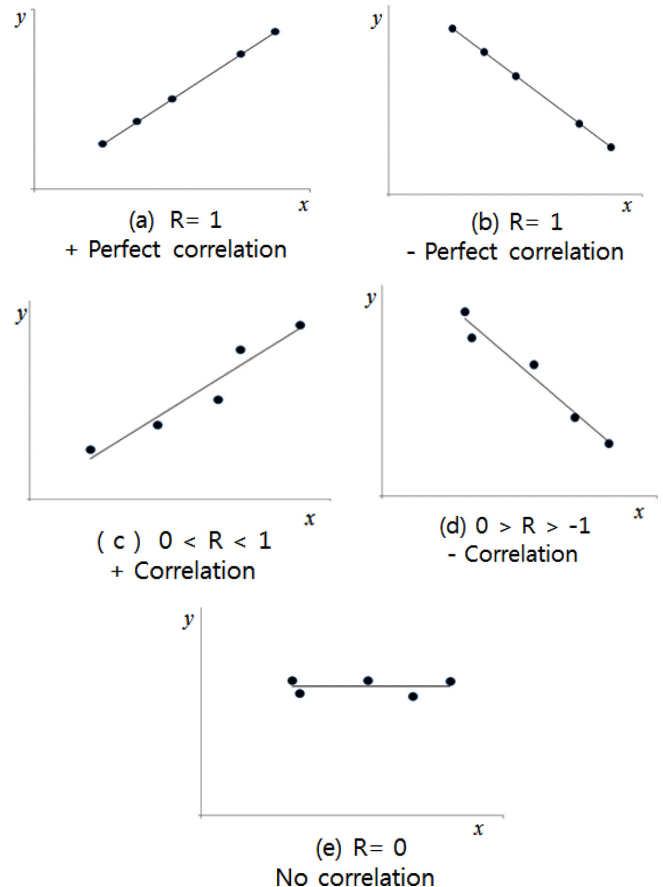


Fig. 4. The scatter plot for the relationship between x and y .

다. (d)는 $R=0$ 인 무 상관관계를 나타내고 있다.

2-5. 분산분석(ANOVA: Analysis of variance)

회귀식이 얻어지면 그 식이 통계적으로 의미가 있는지 어떤지를 검토할 필요가 있다. 회귀분석에 있어서 식 (11)에서 본 바와 같이 총 제곱합을 회귀선에 의하여 설명되는 부분(회귀 제곱합)과 안 되는 부분(잔차 제곱합)으로 나누어 고찰하는 것은 매우 의미 있는 일이다. 회귀선에 의하여 설명되는 제곱합을 회귀에 의한 제곱합이라 부르며, 이는 식 (11)을 이용하여 표현하면,

$$SSR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = \frac{[S_{(xy)}]^2}{S_{(xx)}} \quad (17)$$

그리고 회귀선에 의하여 설명 안 되는 제곱합을 잔차(residual)로 인한 제곱합이라고 부르며, 이 제곱합의 간편한 계산은 아래와 같다.

$$SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - \bar{y})^2 - \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = SST - SSR \quad (18)$$

주어진 자료를 적합시키는데 있어서 회귀선이 유의(significant)한가 하는 것은 SSR이 상대적으로 SSE보다 어느 정도 큰가로 알 수 있다.

Table 2는 분산분석표로 회귀선이 유의한지를 확인하기 위하여 작성한다. 유의성 검정을 위해서는 우선 가설설정을 하여야 한다. 귀무가설은 “회귀모형이 유의하지 않다”로 하고 대립가설은 “회귀모형이 유의하다”로 설정한다. 다음으로 유의수준(Level of significance) α 를 결정하여야 한다. 여기서 유의수준 α 는 진실은 참인데 거짓으로

Table 2. Simple regression ANOVA

Factor	Sum of square	Degree of freedom	Mean square	F_0	$F(\alpha)$
Regression	SSR	1	MSR=SSR	MSR/MSE	$F(1, n-2 : 1-\alpha)$
Residual	SSE	n-2	MSE=SSE/n-2		
Total	SST	n-1			

Table 3. F-distribution

n_2-1	$1-\alpha$	n_1-1											
		1	2	3	4	5	6	7	8	9	10	11	12
1	0.90	39.9	49.5	53.6	55.8	57.2	58.2	58.9	59.4	59.9	60.2	60.5	60.7
	0.95	161	200	216	225	230	234	237	239	241	242	243	244
	0.975	648	800	864	900	922	937	948	957	963	969	973	977
	0.99	4052	5000	5403	5625	5764	5859	5928	5981	6022	6056	6082	6106
2	0.90	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.40	9.41
	0.95	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4
	0.975	38.5	39.0	39.2	39.3	39.3	39.3	39.4	39.4	39.4	39.4	39.4	39.4
	0.99	88.5	99.0	99.2	99.2	99.3	99.3	99.4	99.4	99.4	99.4	99.4	99.4
3	0.90	5.54	5.46	6.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.22	5.22
	0.95	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.76	8.74
	0.975	17.4	16.0	15.4	15.1	14.9	14.7	14.6	14.5	14.5	14.4	14.4	14.3
	0.99	34.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.3	27.2	27.1	27.1
4	0.90	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.91	3.90
	0.95	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.94	5.91
	0.975	12.2	10.7	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.80	8.75
	0.99	21.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.7	14.5	14.4	14.4
5	0.90	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.28	3.27
	0.95	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.71	4.68
	0.975	10.0	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.57	6.52
	0.99	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2	10.1	9.96	9.89
6	0.90	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.92	2.90
	0.95	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00
	0.975	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.41	5.27
	0.99	13.7	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.79	7.72
7	0.90	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70	2.68	2.67
	0.95	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.60	3.57
	0.975	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.71	4.67
	0.99	12.2	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.54	6.47
8	0.90	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.52	2.50
	0.95	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.31	3.28
	0.975	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.25	4.20
	0.99	11.3	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.73	5.67
9	0.90	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.40	2.38
	0.95	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.10	3.07
	0.975	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.91	3.87
	0.99	10.6	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.18	5.11
10	0.90	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.30	2.28
	0.95	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.94	2.91
	0.975	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.67	3.62
	0.99	10.0	7.66	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.77	4.71
11	0.90	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27	2.25	2.23	2.21
	0.95	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.82	2.79
	0.975	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.48	3.43
	0.99	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.46	4.40
12	0.90	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19	2.17	2.15
	0.95	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.72	2.69
	0.975	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.32	3.28
	0.99	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.22	4.16

판단오류를 발생시킬 확률을 의미고 P-value는 귀무가설이 참일 확률을 의미한다. 유의수준은 통상 5%로 설정을 하고 있다. Table 2에서 $F_0 > F(1, n-2 : \alpha)$ 이거나 $P\text{-value} < \alpha$ 이면 귀무가설을 기각한다. 즉, “유의수준 α 에서 회귀선이 유의하다.”라고 검정한다.

Factor 항목은 회귀(Regression), 잔차(Residual) 및 합계로 구성된다. Sum of square 항목은 회귀 제곱합, 잔차 제곱합 및 총 제곱합으로 구성된다. Degree of freedom 항목은 자유도를 의미한다. 여기서 자유도란 자유롭게 선택할 수 있는 변수의 수를 의미하며 샘플일

경우 제곱합의 평균을 산출하기 위해 나누는 수이다. 회귀항의 자유도는 1이며 잔차항의 자유도는 $n-2$ 이다. 자유도가 아닌 객체수로 나눌 경우 샘플집단에서 추정된 분산이 모집단의 분산과 편이가 발생하여 모집단 분산 추정이 정확하지 않기 때문이다. Mean square 항목은 제곱합을 자유도로 나눈 표본 분산이며, F_0 항목은 회귀의 평균제곱 MSR과 잔차의 평균제곱 MSE와의 비율이다. 이 비율이 높으면 회귀에 의한 제곱합이 잔차의 그것보다 상대적으로 많이 커서 회귀선이 x 와 y 간의 관계를 설명하는데 유의하다는 의미이다. $F(\alpha)$ 항목은 기각치이다. 구하는 방법은 엑셀함수[=FINV(α , n_1-1 , n_2-1)] 확률 P , 자유도 1, 자유도 2에 따라 F 값을 구하거나 엑셀의 분석도구 기능을 이용하여 구한다. 혹은 Table 3을 이용하여 주어진 유의수준 α 와 자유도(1, $n-2$)를 사용하여 기각치 $F(1, n-2 : \alpha)$ 를 구한다.

판정은 $F_0 > F(\alpha)$ 인 경우 유의수준 α 에서 귀무가설 기각한다. 즉, “회귀모형은 유의하다.”로 판정한다.

Table 3은 F-분포표로 자유도 1, 자유도 2와 신뢰수준(90%, 95%, 99%)에 따라 확률변수 F 값을 나타낸 표이다. $F(\alpha)$ 의 조건에 따라 값을 Table 3에서 찾아 기입하여 F_0 의 값과 비교할 수 있게 한다. 읽는 방법은 n_1-1 은 n_1 의 자유도이며 n_2-1 은 n_2 의 자유도이다. 신뢰수준(confidence Level)을 90%, 95%, 99%를 결정하여 결과값을 읽는다. 신뢰수준이란 참값이 존재할 확률을 의미한다.

회귀분석을 시행할 Software는 엑셀 Ad-in Software인 분석도구를 활용하였다.

3. 적용장치 및 방법

3-1. 폐액 및 제품의 조성

본 폐액의 조성은 평균 조성이다. 질량비로 계산하였으며 원료의 평균 조성은 PGMEA 84.3%, 기타 13.7%, 수분 2%이며 1차 증류탑 통과 후 Bottom의 평균 조성은 PGMEA 93%, 기타 7%, 수분 0.02%이고 수율은 75%이다. Overhead 평균 조성은 PGMEA 93%,

기타 5%, 수분 2%이고 수율은 25%이다.

Fig. 5는 PGMEA를 재활용 하는 2기 모드의 증류 제조 공정도다. 1차 증류탑의 역할은 폐액 원료가 Feeding되면 Overhead는 수분을 분류하여 폐기시키며 Bottom으로 중간제품을 생성하여 2차 증류탑으로 Feeding시킨다. 2차 증류탑 통과 후 Overhead 평균 조성은 PGMEA 99% 이상, 기타 1% 미만, 수분 0.02%이고 수율은 85%이다. Bottom 평균 조성은 PGMEA 55~65%, 기타 35~45%, 수분은 측정하지 않았고 수율은 25%이다. 2차 증류탑의 기능을 분석하면 Overhead에서는 제품을 생산하고 Bottom은 불순물이 다량 함유되어 폐기 처분한다. 현재는 수율을 향상시키기 위해서 1차 증류탑의 Overhead 제품과 2차 증류탑의 Bottom 제품을 재활용하는 방법도 활발히 연구되고 있다.

PGMEA 재활용에 1기 증류탑만을 사용하지 않는 이유는 수분을 전처리하여야 하는 어려움이 있고 PGMEA 함량을 99% 이상 지속 유지하기가 어려움이 있기 때문이다. 또한, 다단 증류를 하지 않는 이유는 폐액의 조성을 항상 일정하게 유지하여야 하는데 생산현장의 여건과 폐액 배출업체 여건에 따라 조건이 일정하지 않은 어려움이 있기 때문이다.

3-2. 적용장치

1차 증류탑은 직경 약 950 mm, 높이 약 45 m이며 약 120 단으로 구성되었다. Tray는 Fig. 6과 같이 Sieve tray이다. Sieve tray는 Downcomer와 1~20 mm 직경의 hole이 일정한 간격으로 배치되어 있는 Plate로 구성되어 있으며 Perforated tray라고도 한다. Hole을 통하여 상층 Tray로 상승하는 증기는 Tray 위에 머물러 있는 액체가 구멍으로 흘러내리지 않도록 하면서 액체 층을 통과하며, 방출된 독은 Tray 위의 액체가 일정한 높이를 유지하는 역할을 한다.

2차 증류탑은 직경 약 1000 mm, 높이 약 35 m이며 약 70단으로 구성되었다. 1단은 패킹 층이며 다른 Tray는 Fig. 6과 같이 Sieve tray이다. 패킹은 탑 내에서 기-액 간의 물질이동 및 열 이동 효율을 증대시키기 위해 Tray와 같은 역할을 하는 것으로, 그 종류가 매우 다양하며 Steel, Ceramic, Plastic 재질 등으로 구성된다. 대표적인 형태는 Fig. 7과 같이 Random Packing과 Structured Packing 두 종류다. 패킹 타입에 따라 수율에 영향이 있기 때문에 적절한 패킹을 선택하여야 한다.

3-3. 적용자료

본 연구에서 활용한 자료는 현장에서 적용한 공정자료이다. 1차 증류탑의 공정 조건 자료는 Feed 량, Reflux 량, BTM 온도, Feed 온

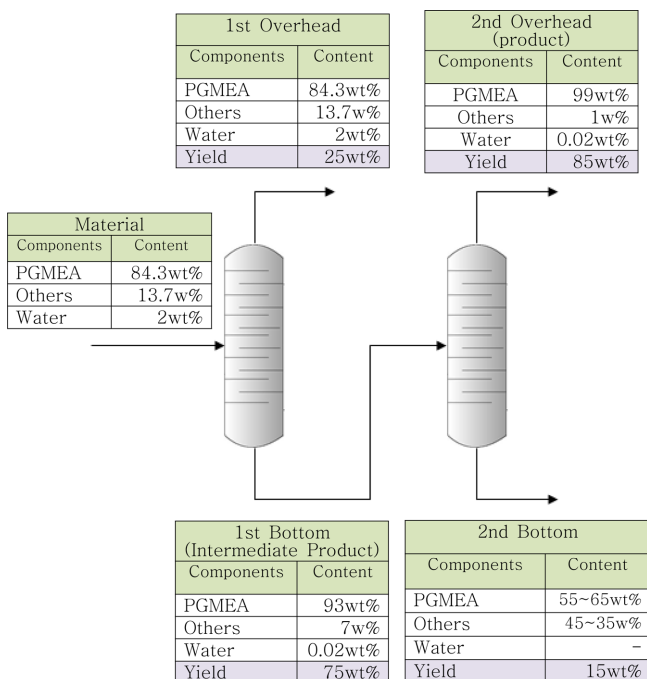


Fig. 5. Manufacturing process utilizes a two-stage distillation column.

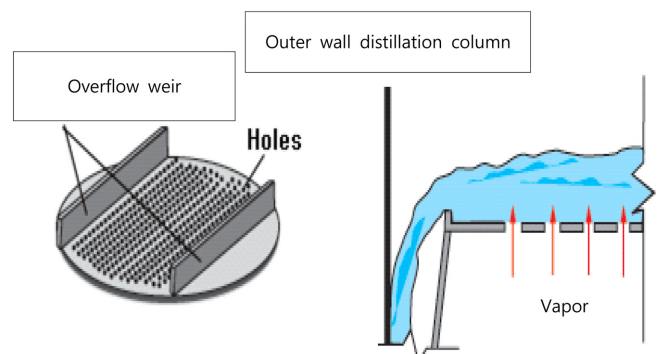


Fig. 6. Sieve tray.

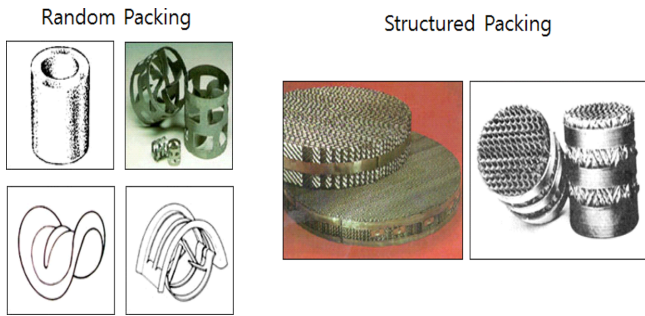


Fig. 7. Packing type.

도 조건에 따른 Bottom의 PGMEA 함량(%) 및 수분의 함량(ppm) 조성이며 이상값을 제거하고 분석에 활용한 자료와 2차 증류탑의 공정 조건인 Feed 량, Reflux 량, BTM 온도, Feed 온도 조건에 따른 Bottom의 PGMEA 함량(%) 및 수분의 함량(ppm) 조성이며 이상값을 제거하고 분석에 활용한 자료는 현장에 사용된 자료이다. 여기서 측정단위인 함량은 wt%로 GC로 측정한 PGMEA 함량, 기타와 K-fisher로 측정한 수분함량의 값을 합산하여 무게비로 하였다.

4. 결과 및 고찰

4-1. 1차 증류탑 PGMEA 함량에 미치는 인자 및 최적조건 산출

4-1-1. 인자별 상관분석

상관분석 전에 비정상적인 관측치를 4차에 걸쳐서 30개의 관측치를 제거하였으며 최종 85개의 관측치로 분석을 실시하였다.

Fig. 8는 PGMEA 함량과 각 인자간의 산점도이다. PGMEA 함량과 Reflux 량이 약한 양의 상관관계가 있다. PGMEA 함량과 BTM 온도는 약한 음의 상관관계가 있다. 또한, PGMEA 함량과 Feed 량은 무상관관계가 있고 PGMEA 함량과 Feed 온도는 무상관관계가 있다. 인자 간의 상관관계를 확인하면 Feed 량과 다른 원인인자와는 무상관관계가 있음을 알 수 있다. Reflux 량과 다른 원인인자와도 무상관관계가 있음을 알 수 있으며 Feed 온도와 다른 원인인자와도 무상관관계가 있다. Feed 온도와 다른 원인인자와도 상관관계가 없음을 알 수 있다.

산점도로 판단하기 어려운 상관관계를 수치화할 필요성 있다. Table 4는 인자별 상관계수를 수치화하여 나타내었다. PGMEA 함량과 Feed 량의 상관계수는 0.58이며 PGMEA 함량과 Reflux 량의 상관계수는 0.39 이므로 약한 양의 상관관계가 있다. PGMEA 함량과

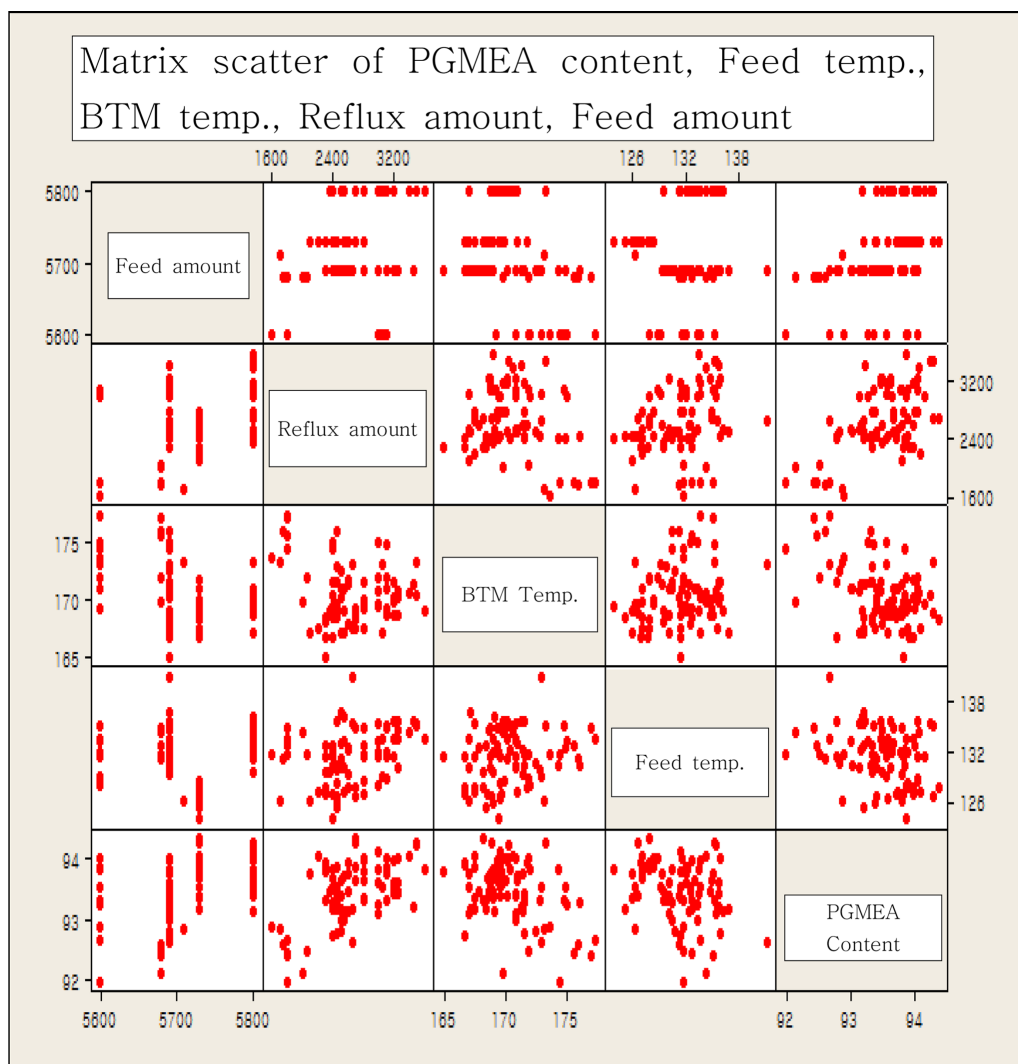


Fig. 8. The scatter plot of feed amount, reflux amount, BTM temp., feed temp. and PGMEA content at 1st tower.

Table 4. Multiple correlation analysis

	Feed amount	Reflux amount	BTM temp.	Feed temp.	PGMEA content
Feed amount	1				
Reflux amount	0.199034	1			
BTM temp.	-0.40088	-0.2195	1		
Feed temp.	0.02942	0.318926	0.208357	1	
PGMEA content	0.386077	0.575837	-0.41327	-0.28127	1

Table 5. Regression analysis statistics

Regression Analysis	Statistics
Multiple correlation coefficient	0.801821
The coefficient of determination	0.642917
The adjusted coefficient of determination	0.625062
Number of observations	85

Table 6. Regression ANOVA

Factor	Degree of freedom	Sum of square	Mean square	F0	P-value
Regression	4	13.76	3.4	36.0	3.44E-17
Residual	80	7.65	0.095565		
Total	84	21.41			

BTM 온도의 상관계수는 -0.41이므로 약한 음의 상관관계가 있다. 또한, 원인인자 사이에 상관관계는 없어 회귀분석에 장애가 되는 다중공선성(Multicollinearity)은 없다.

4-1-2. 회귀분석

Table 5는 PGMEA 함량을 결과값으로 다중회귀분석을 실시한 결과이다. 결정계수 R^2 과 조정된 결정계수 $R^2\text{-adj} > 0.6$ 이상이므로 원인인자들의 변동은 다중회귀식으로 설명할 수 있다.

Table 6은 분산분석을 실시한 결과이다. 분산분석에서 귀무가설은 “회귀모형이 유의하지 않다.” 이고 대립가설은 “회귀모형이 유의하다.” 이다. 유의수준 5%에서 $P\text{-value} < 0.05$ 이므로 귀무가설을 기각한다. 즉, 다중회귀식은 유의하다. 다중회귀식은 PGMEA 함량에 대해 원인인자들로 설명할 수 있다. 여기서 유의수준(Level of significance : 有意水準)은 가설검정 시 귀무가설이 참일 때 거짓으로 잘못 판단할 확률을 의미한다. P-value는 귀무가설이 참일 확률을 의미한다.

Fig. 9은 잔차의 경향이 있는지 확인하는 그래프이다. 잔차는 기댓값이 0으로 특정한 경향이 없어야 정규성이 있다. 그래프 분석결과 특이 사항이 없음을 알 수 있다. 즉, 잔차항은 특이한 경향이 없음을 알 수 있다.

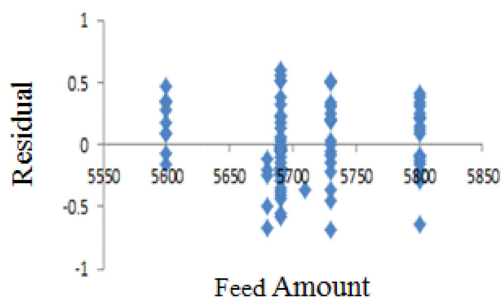


Fig. 9. Normality tested.

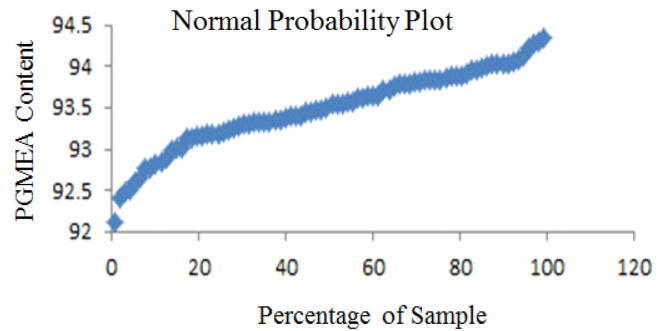


Fig. 10. Normality tested.

Fig. 10도 정규성에 대한 검토 결과이다. 직선을 나타내고 있으면 정규분포를 가정할 수 있다. 그래프 분석 결과 직선을 나타내고 있어 정규분포를 가정한다. 종합적으로 Table 5, Table 6, Fig. 9, Fig. 10로부터 아래와 같이 유도된 다중회귀식은 의미가 있음을 나타내는 결과이다.

Table 7은 다중회귀분석표이다. PGMEA 함량에 영향을 미치는 원인인자는 t 통계량이 크면 영향이 있는 것으로 확인되었다. Table 7의 공정 조건 중에 t 통계량은 Reflux 량, Feed 온도, Feed 량 및 BTM 온도 순으로 나타났다. 오히려 BTM 온도는 본 연구에서 적용한 온도 범위에서는 PGMEA 함량에 영향을 크게 미치지 않는 것으로 나타났다.

다중회귀식을 Table 7에서 산출하며 다음과 같다.

다중회귀식

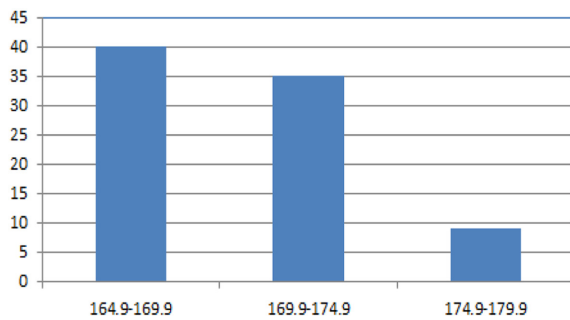
$$Y(\text{PGMEA content}) = 91.62749 + 0.002082 \times \text{Feed amount} + 0.000746 \times \text{Reflux amount} - 0.01331 \times \text{BTM temp.} - 0.07415 \times \text{Feed temp.}$$

Fig. 11은 공정에 적용된 각 원인인자의 분포도를 나타내었다. BTM 온도 분포는 164.9~169.9 °C에서 84회 운전 중 40회(47.6%) 운전하였으며 경제성을 고려하여 165 °C를 운전조건으로 설정하는 것이 타당하다. Reflux 량은 2500~2999 ℓ 운전의 빈도는 84회 중 25회

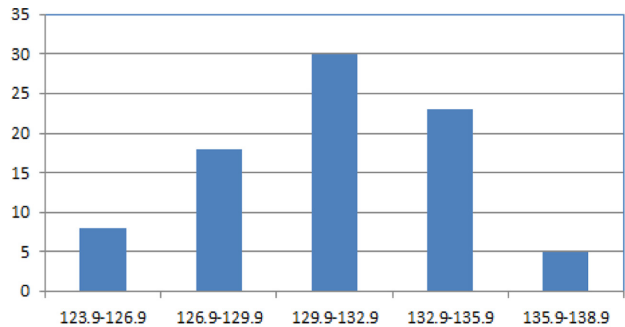


Table 7. Multiple regression analysis

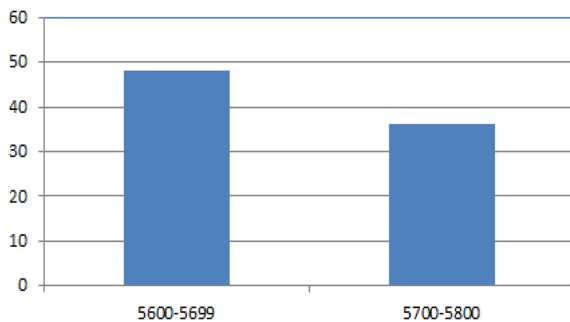
	Coefficient	Standard error	t statistic	P-value	Lower confidence interval 95%	Upper confidence interval 95%
Y-axis intercept	91.62749	5.0875433	18.01016	7.0505E-30	81.50295	101.752
Feed amount	0.002082	0.0006405	3.249941	0.00168949	0.00081	0.003356
Reflux amount	0.000746	8.30E-05	8.987605	9.2906E-14	0.00058	0.000911
BTM temp.	-0.01331	0.0149846	-0.88825	0.37706691	-0.04313	0.01651
Feed temp.	-0.07415	0.0112942	-6.56532	4.7903E-09	-0.09662	-0.05167



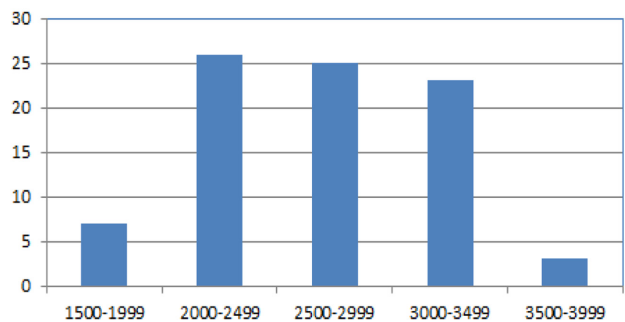
(a) Distribution of BTM temp.



(b) Distribution of Feed temp.



(c) Distribution of Feed amount



(d) Distribution of Reflux amount

Fig. 11. Distribution of each cause factor.

Table 8. Optimized process conditions at 1st tower

Estimated PGMEA content	Feed amount	Reflux amount	BTM temp.	Feed temp.
93.37%	5700 ℓ	2500 ℓ	165 °C	130 °C

29.7%로 Reflux 량이 적정하다. 이에, 2500 ℓ를 운전조건으로 설정하는 것이 타당하다. Feed 량은 생산량을 고려하여 5700 ℓ로 설정하였고 Feed 온도는 130 °C로 하였을 때 PGMEA 함량은 93.52%가 예상된다.

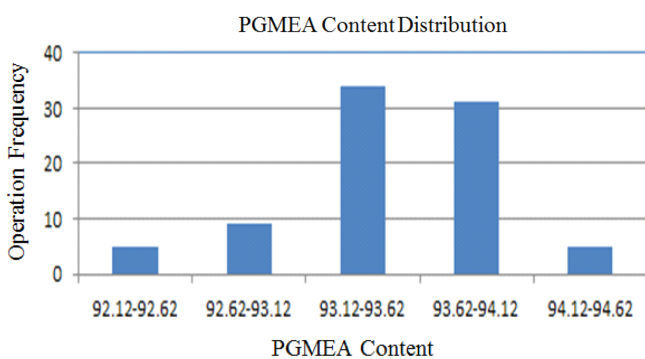


Fig. 12. Distribution of PGMEA content.

Table 8은 유도된 다중회귀식과 경제성을 고려하여 산출한 최적 조건을 정리하여 나타내었다. Feed 량 5700 ℓ, Reflux 량 2500 ℓ, BTM 온도 165 °C, Feed 온도 130 °C일 때 추정 PGMEA 함량은 평균 93.37%이다.

Fig. 12는 본 연구에서 적용한 공정조건에 따른 1차 증류탑의 PGMEA 함량의 분포를 나타내었다. PGMEA 함량은 92.12~94.62%까지 다양하게 분포되어 있으나 공정조건을 Table 8과 같이 표준화한다면 현재보다는 정밀하게 관리가 가능하다.

4-2. 2차 증류 PGMEA 함량에 미치는 인자 및 최적조건 산출

4-2-1. 인자별 상관분석

Fig. 13은 각 원인인자별 결과값과의 상관관계를 산점도로 나타낸 결과이다. PGMEA 함량과 Feed 량과는 양의 상관관계가 있고 PGMEA 함량과 Reflux 량도 양의 상관관계가 있다. PGMEA 함량과 BTM 온도도 양의 상관관계가 있음을 확인할 수 있다. 원인인자 간의 상관관계를 확인하여 보면 Feed 량과 다른 원인인자 간에는 무

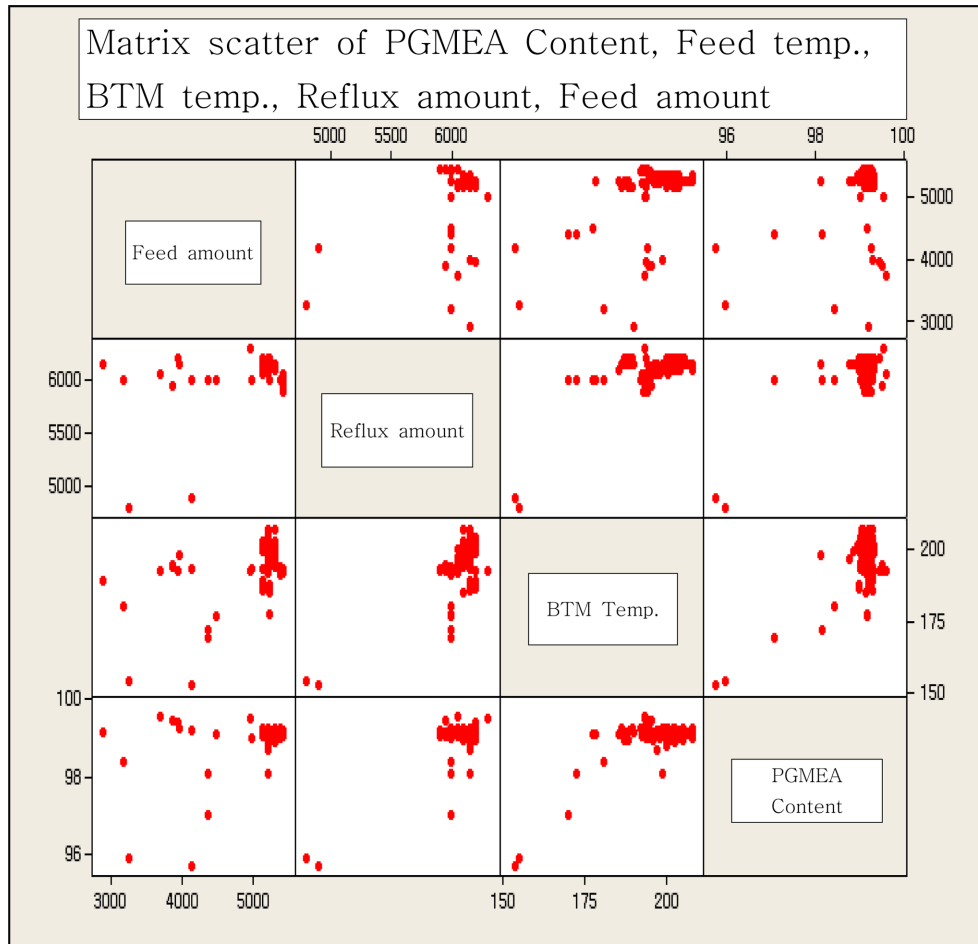


Fig. 13. The scatter plot of feed amount, reflux amount, BTM temp., feed temp. and PGMEA content at 2nd tower.

상관관계이며 Reflux 량과 BTM 온도는 양의 상관관계가 존재한다. 상관관계를 수치화하여 자세히 확인할 필요성이 있다. Table 9는 각 원인인자와 특성인자의 상관분석의 결과이다. PGMEA 함량과 Feed 량의 상관계수는 0.38로 약한 양의 상관관계가 있으며 PGMEA 함량과 Reflux 량의 상관계수는 0.75로 강한 양의 상관관계가 있다. 또한 PGMEA 함량과 BTM 온도는 0.64로 강한 양의 상관관계가 있음을 알 수 있다. 원인인자 간의 상관관계를 확인하여 보면 Feed 량과 BTM 온도의 상관계수가 0.42로 약한 양의 상관관계가 있음을 알 수 있으며 Reflux 량과 BTM 온도의 상관계수는 0.72로 강한 양의 상관관계가 있음을 알 수 있다. 즉, 회귀분석에 장애가 되는 다중공선성(Multicollinearity)이 Reflux 량과 BTM 온도에서 나타나고 있다. 즉, Reflux 량에 따라 BTM 온도가 변화하고 있다.

4-2-2. 회귀분석

다중회귀식이 의미가 있는지 다음과 같이 분석하였다. Table 10은 다중회귀분석을 실시한 결과이다. 결정계수인 R^2 과 조정된 결정계

Table 10. Regression analysis statistics

Regression Analysis	Statistics
Multiple correlation coefficient	0.77177
The coefficient of determination	0.59563
The adjusted coefficient of determination	0.590141
Number of observations	225

수 $R^2\text{-adj} > 0.59$ 이상이므로 원인인자들의 변동은 다중회귀식으로 설명할 수 있다. 즉, 다중회귀식은 PGMEA 함량에 대해 원인인자들로 설명된다.

Table 11은 분산분석의 결과이다. 귀무가설은 “회귀모형이 유의하지 않다”이고 대립가설은 “회귀모형이 유의하다”이다. 조건은 유의수준(Level of significance) 5%로 설정하였다. 분석 결과 유의수준 5%에서 P-value < 0.05 이므로 귀무가설을 기각한다. 즉 유도되는 “다중회귀식은 유의하다”이다.

Fig. 14는 잔차항의 경향을 분석한 그래프이다. 잔차는 기댓값이

Table 9. Multiple correlation analysis

	Feed amount	Reflux amount	BTM temp.	PGMEA content
Feed amount	1			
Reflux amount	0.328716	1		
BTM temp.	0.423379	0.722571394	1	
PGMEA content	0.377313	0.752976897	0.636619	1

Table 11. Regression ANOVA

Factor	Degree of freedom	Sum of square	Mean square	F ₀	P-value
Regression	3	18.91640346	6.305468	108.509587	3.26225E-43
Residual	221	12.84226054	0.05811		
Total	224	31.758664			

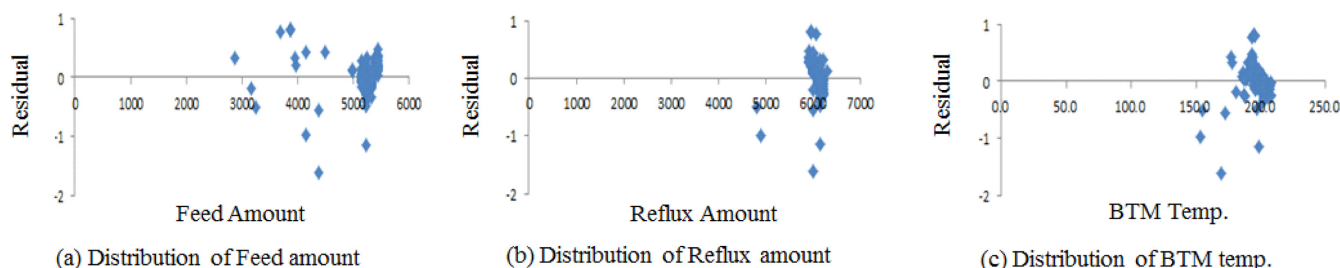


Fig. 14. Normality tested.

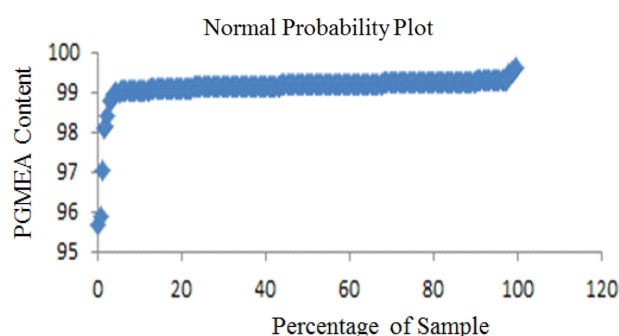


Fig. 15. Normality tested.

0으로 특정한 경향이 없어야 정규성이 있다. 그래프 분석결과 Feed 량, BTM 온도, Reflux 량은 중심선을 기점으로 랜덤하게 분포하고 있으며 특이 사항이 없음을 알 수 있다. 즉, 잔차항이 경향이 없음을 알 수 있다.

Fig. 15는 PGMEA 함량의 Plot를 나타낸 그래프로 정규성에 대한 검토 결과이다. 직선을 나타내고 있으면 정규분포를 가정할 수 있다.

그래프 분석결과 직선을 나타내고 있어 정규분포를 가정한다. 종합적으로 Table 10, Table 11, Fig. 14, Fig. 15로부터 아래와 같이 유도된 다중회귀식은 의미가 있다.

Table 12는 다중회귀분석표이다. PGMEA 함량에 영향을 미치는 원인은 t 통계량이 크면 영향이 있다. 2차 증류탑에서 PGMEA 함량에 영향을 미치는 원인인자는 t 통계량이 9.81인 Reflux 량이 핵심 인자이며 다음으로 BTM 온도와 Feed 량이다. 또한, BTM 온도와 Reflux 량과는 강한 상호작용, 즉 다중공선성이 있다. 따라서 다중회귀식에서는 BTM 온도를 고정하여 Reflux 량을 조절하여 최적조건을 얻을 수 있다. Table 12로부터 유도된 다중회귀식은 아래와 같다.

다중회귀식

$$Y(\text{PGMEA content}) = 86.82254 + 0.00012 \times \text{Feed amount} + 0.001669 \times \text{Reflux amount} + 0.007705 \times 199$$

Fig. 16은 각 원인인자들의 분포도를 나타내었다. BTM 온도 분포는 198.4~203.4℃에서 225회 운전 중 86회(38.0%) 운전하였으며 경제성을 고려하여 199℃를 운전조건으로 설정하는 것이 타당하며 Reflux 량은 6000~6199 ℓ 운전의 빈도는 225회 중 198회(88.0%)로 운전하였으

Table 12. Multiple regression analysis

	Coefficient	Standard error	t statistic	P-value	Lower confidence interval 95%	Upper confidence interval 95%
Y-axis intercept	86.82254	0.7256551	119.65	6.868E-20	85.3924459	88.25262
Feed amount	0.00012	4.946E-05	2.4223	0.0162290	2.2333E-05	0.000217
Reflux amount	0.001669	0.0001701	9.8130	4.243E-19	0.00133384	0.002004
BTM temp..	0.007705	0.0033348	2.3104	0.0217872	0.00113271	0.014277

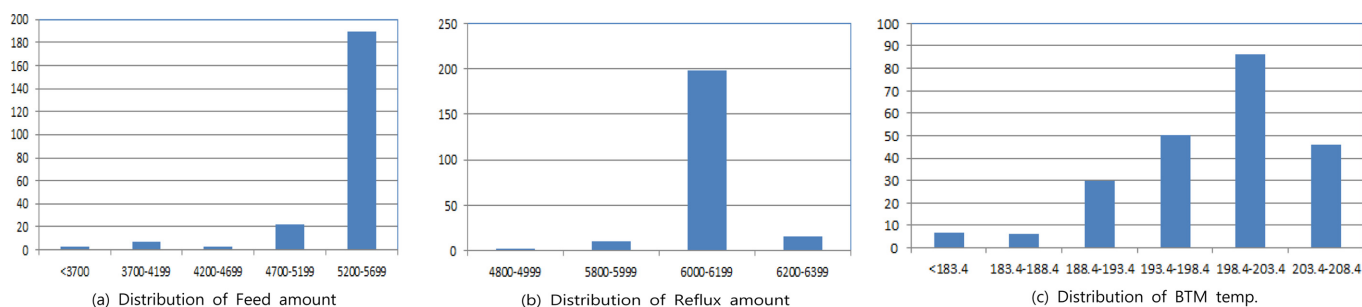


Fig. 16. Distribution of each cause factor.

Table 13. Optimized process conditions at 2nd tower

PGMEA content	Feed amount	Reflux amount	BTM temp.
99.25%	4,275 ℓ	6,200 ℓ	199 °C

며 Reflux 량은 6200 ℓ를 운전조건으로 설정하는 것이 타당하다. Feed 량은 1차 증류탑에서 수율을 75% 적용하여 4,275 ℓ로 설정하였을 경우 PGMEA 함량은 99.0~99.5%가 추산되었다. 따라서 Table 13은 2차 증류탑의 최적조건을 나타내었다. 최적조건은 Feed 량 4,275 ℓ, BTM 온도 199 °C, Reflux 량 6200 ℓ으로 운전하는 경우 PGMEA 함량은 평균 99.25%로 예측되었다.

5. 결 론

1차 증류탑의 최적 조건은 Feed 량 5,700 ℓ, Reflux 량 2,500 ℓ, BTM 온도 165 °C일 때 PGMEA 최적 함량범위는 92.12~94.62%가 예측되며 Reflux 량이 PGMEA 함량에 영향을 미치고 있고 Feed 량과 Feed 온도도 영향을 미치고 있으나, BTM 온도는 본 공정 운전조건에서 온도에 따라 PGMEA 함량에 크게 영향을 주지 못하고 있는 것으로 확인되었다.

2차 증류탑의 최적 조건은 Feed 량 4,275 ℓ, Reflux 량 6,200 ℓ, BTM 온도 199 °C일 때 PGMEA 최적 함량은 평균 99.25%가 예측되며 Reflux 량이 PGMEA 함량에 영향을 크게 미치고 있고, Feed 량과 BTM 온도도 영향을 미치고 있으나 Reflux 량과 BTM 온도는 다중공선성이 존재하므로 Reflux 량에 따라 지배적인 영향을 받으므로 BTM 온도는 고정하여 예측하였다.

1차 증류탑의 PGMEA 함량의 최적 조건을 산출하는 다중회귀식은

$$Y(\text{PGMEA content}) = 91.62749 + 0.002082 \times \text{Feed amount} + 0.000746 \times \text{Reflux amount} - 0.01331 \times \text{BTM temp.}$$

$$- 0.07415 \times \text{Feed temp.}$$

와 같이 산출된다.

2차 증류탑의 PGMEA 함량의 최적 조건을 산출하는 다중회귀식은

$$Y(\text{PGMEA content}) = 86.82254 + 0.00012 \times \text{Feed amount} + 0.001669 \times \text{Reflux amount} + 0.007705 \times 199$$

와 같이 산출된다.

감 사

이 논문은 2012년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구입니다(No. 2012R1A2A2A01045769).

References

1. <http://www.labpump.co.kr/thar/supercritical.htm>.
2. Lee, M. Y. and Kim, Y. H., "Design of Thermally Coupled Distillation Process Utilizing Existing Columns," *Korean Chem. Eng. Res.*, **46**(5), 1017-1022(2008).
3. Woo, K. S., Lee, C. J., Han, K. H., Ko, J. W. and Yoon, E. S., "Fault Detection & SPC of Batch Process using Multi-way Regression Method," *Korean Chem. Eng. Res.*, **45**(1), 32-38(2007).
4. Noh, H. J., *Statistical Quality Management by Excel 2007*, Hanol Publisher, Seoul(2008).
5. Anderson, D. R., Sweeney, D. J. and Williams, T. A., *Essentials of Modern Business Statistics with Microsoft Excel*, 5th ed., South-Western (Cengage Learning) Publishers, Canada(2012).
6. Park, S. H., *Regression Analysis*, 3th ed., Minyoungsa, Seoul (2007).