

Analysis of biodiesel quality based on infrared spectroscopy and multivariate statistics

Hyesil Kim¹, Hyun-Woo Cho² and J. Jay Liu^{1*}

¹Department of Chemical Engineering, Pukyong National University

²Department of Industrial & Management Engineering, Daegu University

(Received February 20, 2012; Revised July 5, 2012; Accepted July 9, 2012)

적외선 분광분석과 다변량 통계에 기반한 바이오디젤 품질분석

김혜실¹ · 조현우² · 유 준^{1*}

¹부경대학교 화학공학과, ²대구대학교 산업경영공학과
(2012. 2. 20. 접수, 2012. 7. 5. 수정, 2012. 7. 9. 승인)

Abstract: ASTM (American Society for Testing and Materials) D6751-10 suggests analytical methods as well as specifications for biodiesel quality. However, it is expensive and time-consuming to follow the ASTM testing methods to analyze biodiesel and various impurities. This paper develops a quantitative analysis system for biodiesel and impurities based on Infrared spectroscopy and a multivariate statistical method, PLS (partial least squares). In addition, four different pre-processing techniques were compared for spectrum correction and noise reduction. Savitzky-Golay pre-processing showed the best performance.

요 약: ASTM (American Society for Testing and Materials) D6751-10은 바이오디젤의 품질 규격 뿐 아니라 분석방법 또한 제시하고 있다. 하지만 ASTM 표준에 따른 바이오디젤 및 포함된 여러 불순물의 품질 분석은 경제적, 시간적으로 부담이 크다. 본 연구는 적외선 분광분석법(infrared spectroscopy)과 다변량 통계분석법 중 하나인 PLS (partial least square method)를 이용하여 1회 측정만으로 바이오 디젤 및 불순물들의 농도를 분석하는 시스템을 개발하고자 하였다. 특히, 적외선을 이용한 분석에서 생기는 각 물질의 스펙트럼에 대한 산란 보정, 노이즈 감소 등을 위해 SNV, MSC, OSC, Savitzky-Golay 등의 4가지 전처리 방법의 성능을 비교하였다. 품질 분석에 필요한 바이오 디젤 검량 모델을 PLS로 모델링 결과, Savitzky-Golay 전처리를 하였을 때 정확도가 가장 우수함을 알았다.

Key words: biodiesel, multivariate calibration, infrared spectroscopy, spectrum pre-processing, Savitzky-Golay method

★ Corresponding author

Phone : +82-(0)51-629-6453 Fax : +82-(0)51-629-6429

E-mail : jayliu@pknu.ac.kr

1. 서 론

최근 고유가와 기후 변화 협약이 국제 사회의 주요 이슈로 떠오름에 따라 대체 연료에 대한 관심이 높아지고 있다. 식물로 대표되는 바이오메스로부터 생산 가능한 바이오 연료는 직접 차량 연료로 활용 가능하며 CO₂ 및 공해 물질 절감 효과가 우수하며 지원 고갈의 우려가 없어 전 세계적으로 점차 사용량이 증가하고 있다. 이러한 바이오 연료의 대표적인 것으로는 휘발유 대체 연료인 바이오 에탄올과 경유 대체 연료인 바이오 디젤이 있다.

이 중 바이오 디젤은 동물성 지방과 식물성 기름 그리고 폐식용유와 같은 재생 가능한 자원을 산 또는 염의 촉매 존재 하에 알코올과 반응시켜 생성하는 에스터화 오일을 말하는데 최대 장점은 생분해성이며 저독성이라는 것이다. 이러한 바이오 디젤유의 품질확인을 위해 물리적 성질을 ASTM에 따른 표준 시험 방법에 따라 측정하여 왔다.¹ 하지만 ASTM에서 제시한 방법으로는 시간적, 경제적 손실이 클 뿐 아니라 상대적으로 많은 양의 시료를 사용해야 한다는 단점

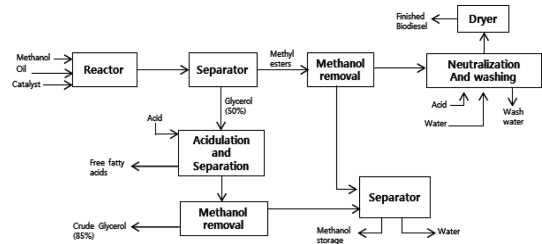


Fig. 1. A schematic of biodiesel production process.

이 있다. 따라서 이를 보완하는 방법으로서 적외선 분광분석법을 이용해 바이오 디젤의 생산을 위한 실시간 품질 분석시스템을 개발하여 국내 바이오디젤 대량생산 시스템 및 관련 산업에 도움이 되고자 한다.

일반적으로 바이오디젤을 생산하기 위한 주요 공정은 소·중규모의 회분식 전환공정이 대부분이다(Fig. 1). 이는 연속식공정과는 달리 강한 비선형성을 띄기 때문에 세심한 제어 및 모니터링, EPD(end-point determination)와 같은 대량 생산기술이 더욱 요구된다. 특히, 회분식 공정의 대량생산 기술의 개발에는 많은 노력을 필요로 하는데 이는 강한 비선형성 외에도, 작업의 복

Table 1. Specification and measurement methods for pure biodiesel given by ASTM D6751¹

Property	Limit	Test method
Calcium & magnesium, combined, ppm, max	5	EN 14538
Flash point (closed cup), °C, min	93	D 93
Alcohol control (one of the following must be met)		
1. Methanol content, vol %, max	0.2	EN 14110
2. Flash point, °C, min	130	D 93
Water & sediment, vol %, max	0.05	D 2709
Kinematic viscosity, 40 °C, mm ² /sec	1.9 - 6.0	D 445
Sulfated ash, % mass (ppm), max	0.02	D 874
Sulfur		
S 15 grade, % mass (ppm), max	0.0015	D 5453
S 500 grade, % mass (ppm), max	0.05	D 5453
Copper strip corrosion, max	No. 3	D 130
Cetane, min	47	D 613
Colud point, °C	report	D 2500
Carbon residue 100% sample, % mass, max	0.05	D 4530
Acid number, mg KOH/g, max	0.50	D 664
Free glycerin, % mass, max	0.020	D 6584
Total glycerin, % mass, max	0.240	D 6584
Phosphorus content, % mass, max	0.001	D 4951
Distillation, T90 AET, °C, max	360	D 1160
Sodium/potassium, combined, ppm, max	5	EN 14538
Oxidation stability, hours, min	3	EN 14112
Cold soak filtration, seconds, max	360	Annex to D 6751
For use in temperatures below -12 °C, seconds, max	200	Annex to D 6751

잡성과 실행에 있어서 다소 복잡한 공정 절차를 필요로 하기 때문이다. 또한 바이오디젤 대량 생산기술이 성공하기 위해서는 공정 중 또는 최종제품의 품질을 측정하여 피드백 할 수 있는 측정 시스템이 요구된다.

바이오디젤 생산 공정에서 공장단위의 제어 및 최적화와 같은 대량생산 기술을 개발하기 위해서는 품질 분석 시 소요되는 시간을 단축하는 것과 저가격화에 중점을 뒀다. 즉 제조공정의 최적화와 제어를 통해 최종 제품의 품질을 확보하기 위해서는 무엇보다 최종 생산물인 바이오디젤에 미량 포함되어 있는 불순물들의 정량 측정이 실시간으로 가능하게 해야 한다. 이처럼 바이오디젤의 경우에는 바이오디젤 생산 과정에서 생성되는 여러 불순물들로 인해 바이오디젤 품질이 저하 될 수 있다. 따라서 Table 1과 같이 ASTM D6751-10과 같은 생산 및 품질관리를 위한 기준과 측정 방법이 제시되고 있다.¹

바이오 디젤은 염기 촉매 하 지방의 에스테르화 반응 공정으로 수율이 매우 높지만 전체 제조공정이 최적화 되지 않을 경우에는 글리세린(glycerine), 글리세라이드(glyceride), 물, 메탄올, 그리고 자유지방산(free fatty acid, FFA) 등의 소량의 불순물이 생성된다. ASTM 표준에 따르면 이 불순물들은 바이오 디젤 최종 품질에서 가장 중요한 불순물이며, 이에 대한 분석방법 또한 표준에서 제시하고 있다. 하지만 이 표준에 따른 각각의 불순물에 대한 정량측정방법은 각기 다른 고가의 분석 장비와 시약을 요구할 뿐 아니라 운영 및 유지보수에 있어 상당한 물적, 인적자원이 요구되며 오랜 측정시간을 필요로 한다. 따라서 공정의 제어 및 최적화를 위한 효율적인 개선책 마련이 시급하다. 즉 시간과 경제적인 측면을 모두 고려한 실시간 분광분석 시스템 개발이 필요하다고 할 수 있다.

이를 위해 본 연구에서는 적외선 분광분석법을 사용하였다. 적외선 분광분석법은 반복 및 재현성이 우수하고 기기가 매우 안정적이므로 검량 모델(calibration model)을 완벽하게 만든다면 오랜 기간 사용이 가능하다. 뿐만 아니라 기기의 유지 보수에 대한 부담도 적으며 안정적으로 사용이 가능하고 특히 동시 다 성분 측정도 가능하다는 것이 적외선 분광분석법의 가장 큰 장점이라고 할 수 있다.

하지만 바이오디젤과 같이 여러 종류의 복잡한 화합물을 가지는 경우에는 각 화합물 고유의 적외선 스펙트럼간의 중첩이 심해 데이터를 분석하는데 있어 어려움이 있다. 따라서 이를 보완하여 정확한 분석을 위한 우수한 예측력과 잡음에 강건한 검량 모델 개발

이 필요하다. 본 연구에서는 다변량 통계기법 중 하나인 부분최소자승법(partial least squares method, PLS)과 여러 스펙트럼 전처리법을 평가하여, 적외선 분광 데이터에 기반을 둔 바이오디젤의 불순물 조성 검량 방법을 제시하였다.

본 연구는 다음과 같은 순서로 진행된다. 먼저 적외선 분광분석에 대한 기초 이론 및 바이오디젤에 포함되어있는 여러 작용기들의 흡수 파장을 나타낸다. 이어서 띠나비가 넓고 중첩이 심한 적외선 스펙트럼 영역에서 원하는 조성 및 물성의 검량을 가능하게 해주는 몇 가지 다변량 통계기법 중 가장 적합한 PLS법과 스펙트럼 전처리 법을 서술한다. 그 다음 분석결과를 요약하고, 끝으로 결론에서는 바이오디젤 품질 분석 시스템 개발의 응용에 대해 기술한다.

2. 이 론

2.1. 적외선 분광분석(infrared spectroscopy, IR)

적외선이란 가시광선과 마이크로파 사이에 존재하는 빛으로, 파장에 따라 근적외선(800~2,500 nm), 중적외선(2,500~5,000 nm), 원적외선(5,000~10,000 nm)로 나눌 수 있다. 이론적으로 에너지는 파장이 짧을수록 높기 때문에 중적외선 같은 경우 가시광선과 근적외선보다는 에너지가 낮고 원적외선보다는 에너지가 높다.

일반적으로 공유결합을 형성하고 있는 화합물들은 대개 스펙트럼의 IR영역에서 작용기의 분자의 회전과 진동에 의해서 피크가 나타나게 된다. 이처럼 IR 분광분석기는 전통적인 CEMS (continuous emission monitoring system)와는 달리 수백-수천의 가스성분을 동시에 측정할 수 있고 측정 속도와 정확도 뿐 아니라 신뢰성 등에서도 우위에 있어 여러 분야에서 화합물에 대한 실시간 측정용 분석법으로 적용되고 있다.^{2,3}

또한 적외선 분광분석법은 고체, 분체, 액체, 가스에 이르기까지 다 성분 흐름상을 동시에 분석할 수 있는 이점을 가지고 있어 공정의 최적화를 위해 필수적이라 할 수 있다. 따라서 본 연구에서는 바이오디젤 및 불순물에 대한 실시간 품질 분석을 위해 중적외선(2,500~5,000 nm)영역을 이용한 분광분석법을 이용하였다. Table 2에 적외선에서의 몇 가지 작용기들의 흡수 파장을 정리하였다.

2.2. 다변량 통계기법과 화학계량학(Chemometrics)

중적외선 분광학(Mid-IR spectroscopy)에서는 2,500

Table 2. IR absorption of some functional groups⁴

Functional group	Type of vibration	Characteristic absorptions (cm ⁻¹)	Intensity
Alcohol			
O-H	Stretch, H-bonded	3,200 ~ 3,600	Strong, broad
O-H	Stretch, free	3,500 ~ 3,700	Strong, sharp
C-O	Stretch	1,050 ~ 1,150	Strong
Alkane			
C-H	Stretch	2,850 ~ 3,000	Strong
-C-H	Bending	1,350 ~ 1,480	Variable
Carbonyl			
C=O	Stretch	1,670 ~ 1,820	Strong
Ether			
C-O	Stretch	2,210 ~ 2,260	Medium
Acid			
C=O	Stretch	1,700 ~ 1,725	Strong
C-O	Stretch	1,210 ~ 1,320	Strong
Ester			
C=O	Stretch	1,735 ~ 1,750	Strong
C-O	Stretch	1,000 ~ 1,300	Two bands or more

~5,000 nm 또는 4000~400 cm⁻¹의 파장을 갖는 적외선을 시료에 조사하여 흡수 스펙트럼을 측정하는데, 이 대역은 주로 분자들의 기본 밴드가 나타나는 영역으로 각 분자의 특정한 흡수밴드가 비교적 깨끗하게 분리되어 있어 특정 화합물의 존재 유무를 확인하는데 쓰여 왔다. 그러나 일반적으로 시료에는 한 가지 화합물만이 존재하는 것이 아니고 여러 가지 화합물이 섞여 있는 경우가 많다. 특히 관심 있는 분자의 스펙트럼이 혼합물을 구성하는 다른 성분의 스펙트럼과 겹치게 되는 경우에는 화합물의 유무를 확인하기가 매우 어렵게 된다.

이와 같이 여러 가지 화합물이 섞여있는 경우에는 복잡한 적외선스펙트럼을 분석하기 위해서 예측력이 우수하고 잡음 등에 강건한 성능을 가진 검량 모델을 개발하는 것이 화학계량학(chemometrics)의 핵심이며, 신속하고 정확한 화학적·물리적 성분의 예측을 가능하게 해주는 중요한 역할을 수행한다. 다변량 통계기법은 중첩이 심한 적외선 스펙트럼에서도 원하는 조성 및 물성의 검량을 가능하게 한다. 즉 어떤 입력 변수 자체의 수가 너무 많거나 이론적인 해석이 쉽지 않아 기존의 일반 회귀분석으로는 상관관계를 찾기 어려울 때, 우리가 알고 있는 출력변수를 이용해 입력 변수와 통계적인 상관관계를 만드는 방법이다. 이는 실험적인 검량기법으로 이 연구의 분광분석에서 입력 변수는 시료의 스펙트럼에 해당되며 출력변수는 시료

의 조성에 해당된다. 화학계량학에서 주로 사용되는 다변량 통계기법은 주성분분석(principal component analysis), 부분최소자승법(partial least square) 등이 있다.⁵

2.1. 부분최소자승법(partial least square method, PLS)

부분최소자승법은 다수의 입력변수와 다수의 응답변수 사이의 관계를 모델링하는 투영법으로, 입력 또는 출력변수 간에 상관관계가 심할 때 발생하는 다중선형회귀법(multiple linear regression, MLR)의 문제점을 극복할 수 있는 방법이다. N번의 측정으로 얻어진 입출력 변수들의 값을 출력변수 행렬 Y 및 입력변수 행렬 X라 할 때, MLR을 행렬 방정식으로 표현하면 다음과 같다.

$$Y = XB + E \quad (1)$$

여기서 B는 입력변수 행렬 X(N×K) 및 응답변수 행렬 Y(N×M)의 선형 관계식을 나타내는 계수로 이루어진 행렬이며, E는 오차로 이루어진 행렬이다. 검량 모델링의 경우 N은 시료의 개수, K는 스펙트럼 내의 파장 수, 그리고 M은 시료내 화학적 성분의 개수이다. 행렬 B는 입출력 변수가 각각 한 개씩인 단변량 데이터에 적용하는 단순 최소 자승법(simple least square method)을 입출력 변수가 각각 여러 개인 경우로 확장하면

$B = (X^T X)^{-1} X^T Y$ 와 같이 구할 수 있다. 하지만 근적외선 스펙트럼을 사용한 검량 모델링에 MLR을 사용할 경우엔 몇 가지 문제점이 발생한다고 널리 알려져 있다.⁶

한편, PLS에서 X, Y 두 행렬은 투영에 의해 다음과 같이 분해된다.

$$X \cong \sum_{a=1}^A t_a p_a^T = TP^T \tag{2}$$

$$Y \cong \sum_{a=1}^A u_a c_a^T = UC^T \tag{3}$$

여기서 T와 U는 각각 입력변수와 출력변수의 score 행렬, P와 C는 각각 입력변수와 출력변수의 loading 행렬이며 잠재변수의 개수 A는 주로 교차검증(ross-validation)을 이용하여 결정된다.⁷ 입력변수 스코어 t_a 는 아래의 식과 같이 X 행렬(a=1 일 때) 또는 X의 잔차 행렬(a>1 일 때)의 선형 조합이 된다.

$$t_a = X_{a-1} W_a, X_a = X_{a-1} - t_a P_a^T \tag{4}$$

w_a 는 a번째 잠재변수의 가중치 벡터로 $p_i = w_i$ 인 관계를 갖는다. 위와 같은 관계식은 t_a 와 u_a 사이의 공분산이 최대가 될 때 성립하며, 결과적으로 $u_a \cong t_a$ 의 관계가 성립한다.⁸ Fig. 2에 PLS의 개념을 도식적으로 나타내었다.

2.3. 스펙트럼 전처리

동일한 분광기라 할지라도 광원, 시료링 장치 그리고 검출기의 열화 등에 의해 시간에 따른 스펙트럼의 불안정성이나 비선형성이 야기될 수 있다.⁹ 이러한 기기 조건의 변화에 의한, 시료의 화학적 조성의 변화와는 상관없는 스펙트럼의 변화가 발생할 때 마다 검량 모델링을 반복하는 것보다는, 검량 모델링 전에 이러한 변화를 스펙트럼으로부터 제거 또는 보완하면 검량 모델링의 유지보수에 드는 노력을 줄일 수 있음은 물론, 보다 안정적이고 정확도가 높은 검량 모델을 얻을 수 있다. 이 때 사용되는 방법 중 널리 알려진 방법으로는 배수산란보정법(multiplicative scatter correction,

MSC), 표준정규변량법(standard normal variate, SNV), Savitzky-Golay법, 직교신호보정법(orthogonal signal correction, OSC) 등이 있다.¹⁰

2.3.1. 배수산란보정(multiplicative scatter correction, MSC)

빛의 산란현상의 결과로 스펙트럼의 바탕선이 이동하거나 파장에 따라 스펙트럼 왜곡의 정도가 다를 수 있는데, 이러한 광학 간섭현상이 배수산란보정법을 이용한 선형화 작업에 의해 제거될 수 있음이 보고되었다¹¹. 배수산란보정법에서는 이상적인 스펙트럼을 이용하여 각 스펙트럼의 파장별 흡광도들과 평균 스펙트럼 흡광도와의 상관관계, 즉 기울기와 절편을 구한다. 이상적인 스펙트럼으로 평균 스펙트럼을 이용할 경우 계산은 다음과 같다.

$$x_r = a_r \bar{x} + b_r \mathbf{1} + e_r \tag{5}$$

여기서, \bar{x} 는 각 파장에서의 시료의 평균 스펙트럼(1×K), x_r 는 시료 r의 스펙트럼이다. 선형회귀를 하여 구한 기울기 a_r 및 절편 b_r 값을 이용하여 다음과 같이 보정된 스펙트럼 \hat{x}_r 을 구한다.

$$\hat{x}_r = \frac{(x_r - b)}{a_r} = \bar{x} + \frac{e_r}{a_r} \tag{6}$$

2.3.2. 표준정규변량법(standard normal variate, SNV)

고체 시료의 경우, 산란현상은 입자의 물리적인 성질에 의존하며 입자크기에 따라 분광경로 길이는 크게 변하므로, 결과적으로 스펙트럼 내에 화학적 변화와는 상관없는 변동을 일으킨다. 각기 다른 물리적 형상을 가진 자당(sucrose) 고체 시료의 스펙트럼을 개별 시료 스펙트럼의 평균과 표준편차로 다음과 같이 정규화하면 이러한 광산란 현상의 영향이 제거됨이 보고되었다.¹²

$$\hat{x}_r = \frac{(x_r - \bar{x}_r \mathbf{1})}{\sqrt{\frac{\sum_{j=1}^K (x_{r,j} - \bar{x}_r)^2}{K-1}}} \tag{7}$$

이때 \bar{x}_r 는 시료 r의 전 파장에서의 흡광도 평균, $x_{r,j}$ 는 시료 r의 파장 j에서의 흡광도, 그리고 K는 전체 파장 수를 나타낸다.

2.3.3. Savitzky-Golay 법

바탕선 이동 등의 노이즈를 감소시키는데 흔히 사

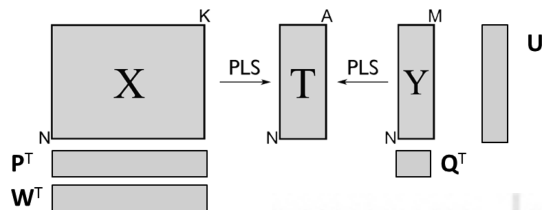


Fig. 2. PLS modeling.

용되는 미분법은 역필터(high-pass filter) 역할을 하므로 바탕선 같은 낮은 주파수대의 특징을 제거하지만 고주파대의 잡음도 동시에 강조되므로, Savitzky-Golay 법은 미분을 계산할 때 동시에 스펙트럼을 평활(smoothing)하므로 이러한 문제점을 해결한다.

Savitzky-Golay 법에서는 실제로 일정한 크기를 갖는 창 내의 스펙트럼 데이터에 다항식을 적합하고 이를 데이터 평활에 이용한다. 이 방법은 창의 크기(즉 데이터의 갯수), 다항식의 차수, 그리고 미분차수의 설정을 필요로 하는데, 일반적으로 창의 크기가 크고 다항식의 차수가 낮을수록 평활효과가 크다고 알려져 있다.¹⁰

2.3.4. 직교신호보정법(orthogonal signal correction, OSC)

직교신호보정법, OSC는 스펙트럼에서 시료의 농도 정보와 상관이 없는 스펙트럼 내의 구조적인 변화를 제거하기 위해 농도 정보와 직교하는 변화를 일련의 행렬연산을 통해 제거한다. 스펙트럼 행렬 X에서 첫 번째 주성분을 찾으면(식 (4)에서 $a=1$ 일 때), 해당 loading 벡터를 회전시켜 농도 정보 행렬 Y와 직교되는 주성분 t^* 를 구하는데 이는 다음과 같이 간단히 구할 수 있다.

$$t^* = (1 - Y(Y^T Y)^{-1} Y^T)t \quad (8)$$

회전된 loading 벡터 p^{*T} 는 스펙트럼 행렬 X 중 농도의 변화와 상관없는 특징들을 나타내므로, 식 (4)를 이용하여 이를 원래의 스펙트럼에서 제거하면 농도와 상관관계가 높은 스펙트럼 성분만 남게 된다.¹³

3. 결과 및 고찰

3.1. 기준 시료의 준비 및 분광측정

본 실험에서 IR 분광기는 Varian 660-IR을 사용하였다. 이 system에서 광원과 검출기는 각각 Duraglow™ lamp와 DLATGS를 사용하였고, beam splitter로는 Ge/KBr을 사용하였다. 시료링 액세스리로는 감쇠 전반사(attenuated total reflectance, ATR) 장치를 사용하였는데, 이는 시료의 표면과 접촉한 적외선이 반사되어 나오는 적외선 beam의 이동경로를 조정해주는 장치로, 시료표면의 유기화합물의 특성에 따라 변동된 흡수도를 측정한다. 측정 옵션은 파장범위를 4,000~400 cm^{-1} 로, 분해능을 4.0 cm^{-1} 로 하고 16번 측정한 후 이것을 평균하여 한 개의 스펙트럼으로 나타내도록 설정하였

다. 이 때 1개의 시료당 약 30 초의 측정시간이 소요되었다. 스펙트럼의 수집과 바탕선 수정 등은 Varian의 FT-IR Resolution Pro 프로그램을 이용하였다.

농도를 알고 있는 바이오 디젤 혼합물을 만들기 위해 순수 바이오 디젤(B-100)에 시약 급의 불순물을 혼합하여 기준 시료를 제조하였다. 이때 사용된 시약은 99.8% (v/v) 메탄올, 99% (v/v) glycerol (free glycerin), 99%의 monolinolein, dilinolein, triolein을 일정비율로 섞은 total glycerin, 그리고 초순수 증류수를 사용하였다. 혼합물의 성분비는 통계 S/W MINITAB에서 혼합물 설계 중 심플렉스 중심 설계 유형을 이용하여 결정하였다. 이때 불순물 부피비(%)의 상한과 하한을 ASTM D6751에 근거하여 설정하였다. 총 33개의 sample의 혼합물의 성분 함량이 결정되었고 이에 따라 33가지의 sample을 제조하였다.

3.2. 결과 및 고찰

모든 기준 시료의 IR 스펙트럼을 Fig. 3에 겹쳐 나타내었다. 측정시 바탕선을 수정하였음에도 불구하고 바탕선의 이동이 심하며, 경우에 따라 피크와 바탕선을 구별하는 데 어려움이 있다. Fig. 3에서 보는 바와 같이 추가적인 스펙트럼 보정이 필요하며 앞에서 제시한 4가지 전처리 방법을 수행하고 그 결과를 Fig. 4-7에 나타내었다. 이 때 정성적인 비교를 위해 모든 스펙트럼을 1~100 사이의 값으로 스케일링한 후 모든 그림에 나타내었다. 또한, Table 2에 나타난 바와 같이 2,000~2,500 cm^{-1} 영역 대에서는 바이오디젤과 생산 중에 발생하는 불순물의 작용기와는 관계가 없기 때

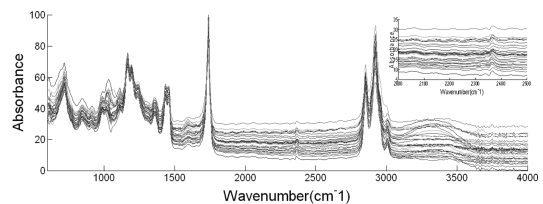


Fig. 3. Raw IR spectrum of reference samples.

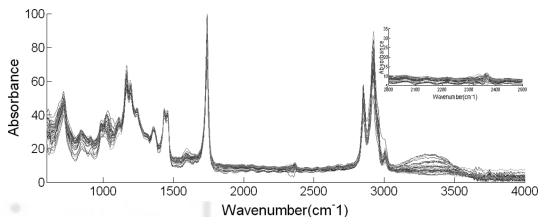


Fig. 4. IR spectrum corrected by MSC.

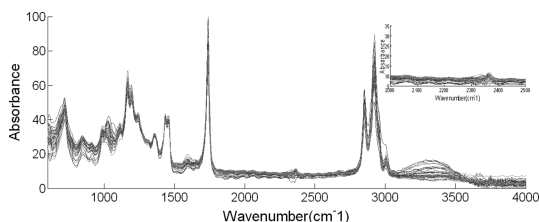


Fig. 5. IR spectrum corrected by SNV.

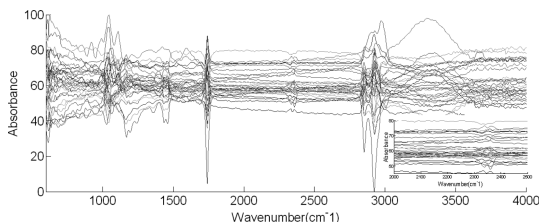


Fig. 6. IR spectrum corrected by OSC.

문에, 전처리에 따른 바탕선 보정효과를 정성적으로 살펴보기 위해 이 부분을 같은 스케일로 확대하여 그림에 나타내었다.

원래의 스펙트럼(Fig. 3)에서는 carbonyl기로 추정되는 $1,670\sim 1,820\text{ cm}^{-1}$ 대의 피크와 alkane기로 추정되는 $2,850\sim 3,000\text{ cm}^{-1}$ 대의 피크를 제외하고는 대부분의 신호가 증첩되어 있는 상태이다. 이를 MSC와 SNV로 전처리 한 스펙트럼(Fig. 4과 Fig. 5)과 비교하면 스펙트럼의 형태가 유사함을 알 수 있고 바탕선도 많이 개선되었음을 알 수 있다. Fig. 6의 OSC 전처리 결과의 경우, $1,670\sim 1,820\text{ cm}^{-1}$ 그리고 $2,850\sim 3,000\text{ cm}^{-1}$ 영역에서 확인할 수 있듯이 피크가 강조됨을 알 수 있으나, 각 파장에서의 스펙트럼 또한 상대적으로 강조되어 바이오 디젤 및 불순물의 성분 결정을 위한 데이터뿐만 아니라 불필요한 노이즈도 두드러짐을 알 수 있다. 하지만 Fig. 7의 Savitzky-Golay 법의 경우, 피크 영역의 스펙트럼만 강조됨과 동시에 바탕선 보정 효과는 다른 전처리법에 비해 월등히 뛰어난 것을 알 수 있다.

여러 물질로 이루어진 혼합물의 경우, IR 스펙트럼의 증첩에 의해 매우 복잡해서 각 성분의 작용기에 해당하는 특정 피크를 쉽게 선택할 수 없기 때문에, PLS 검량모델을 전처리한 스펙트럼으로부터 만들고 이를 이용하여 전처리 효과를 정량적으로 분석하였다. 일반적으로 검량 모델의 성능은 결정계수 R^2 의 값을 가지고 평가한다. ASTM 또는 KS (Korean Industrial Standard) 등의 표준에서는 그 값이 0.990 보다 같거나 클 것을 권장하므로 R^2 이 0.990 이상 일 경우의 PLS

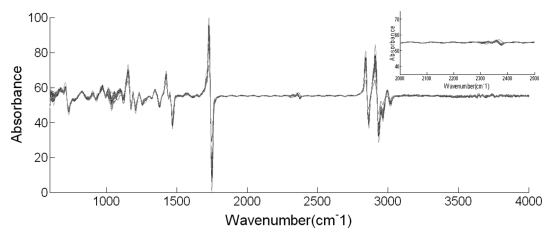


Fig. 7. IR spectrum corrected by Savitzky-Golay method.

의 인자 수를 이용하여 전처리 효과를 비교하였다. 전처리 효과가 작을수록 스펙트럼 중 농도와 상관관계가 없는 노이즈가 상대적으로 많이 포함되어 검량 모델에 보다 많은 PLS 인자가 필요할 것이기 때문이다.

PLS 검량 모델링은 Umetrics社의 Simca-P version 8.0을 사용하여 계산을 수행하였으며 이 때, Fig. 3-7의 스펙트럼의 흡광도 값을 x 변수로, 그리고 부피비를 y 변수로 각각 사용하였다. 각 전처리 방법의 검량 모델링 결과는 Table 3, 4와 같다. 표에서 원래 스펙트럼을 이용한 PLS 검량모델과 비교했을 때, SNV와 MSC는 인자 수가 약간 감소하며 OSC는 인자 수가 동등 수준이거나 오히려 약간 늘어나는 것을 볼 수 있는데, 이는 세 전처리 방법이 바탕선을 충분히 보정하지 못하거나 노이즈를 상대적으로 증폭시킨 영향인 것으로 보인다. 그에 비해 Savitzky-Golay 법은 PLS 인자 개수가 상대적으로 확실히 감소함을 보인다. 따라서 Savitzky-Golay 전처리가 가장 효과적인 스펙트럼 전처리 방법이라 할 수 있다.

Table 4에 나타난 것 처럼 주성분인 바이오 디젤 (fatty acid methyl ester, FAME)은 물론, 메탄올, 물, free glycerin, total glycerin 등 바이오 디젤 제조공정에서 발생하는 주요 불순물의 경우에도 모든 PLS 검량 모델은 정확한 농도 예측 값을 제공함을 알 수 있

Table 3. The number of PLS factors for each pre-treatment methods

Component	The number of PLS factors				
	Raw	MSC	SNV	OSC	Savitzky Golay
All component	23	21	21	23	13
FAME	14	12	12	14	6
Methanol	13	12	12	13	5
Free glycerin	14	12	13	14	5
Mono glyceride	13	12	13	14	5
Di glyceride	12	11	12	11	6
Tri glyceride	14	12	13	14	5

Table 4. Calibration results of pre-treatment methods

Component	R ² value				
	Raw	MSC	SNV	OSC	Savitzky-Golay
All component	0.991	0.992	0.991	0.992	0.990
FAME	0.993	0.993	0.993	0.990	0.996
Methanol	0.992	0.996	0.993	0.993	0.991
Free glycerin	0.991	0.992	0.993	0.99	0.991
Mono glyceride	0.992	0.992	0.995	0.992	0.992
Di glyceride	0.993	0.992	0.996	0.991	0.995
Tri glyceride	0.991	0.995	0.994	0.990	0.994

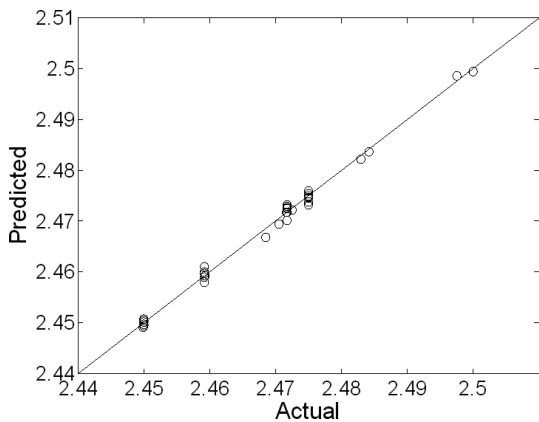


Fig. 8. Predicted versus actual FAME concentration in reference samples.

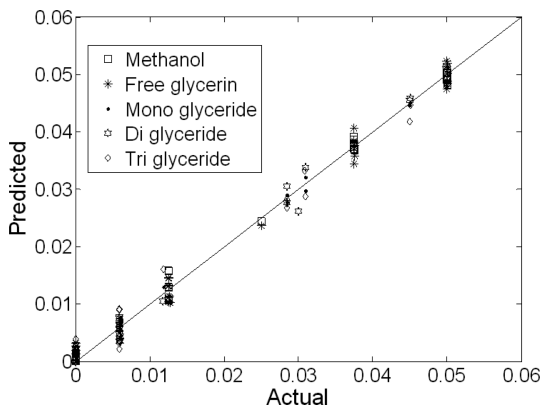


Fig. 9. Predicted versus actual impurities concentrations in reference samples.

다. 하지만 실시간 연속 및 회분식 공정인 바이오 디젤 공정에서 사용하려면 적은 수의 PLS 인자를 갖는 검량선이 선호되는데, 그 이유는 PLS 인자가 많을수록 모델링할 때는 물론 모델의 유지보수에 많은 노력이 필요하기 때문이다. 이러한 관점에서 보았을 때,

Savitzky-Golay 전처리법을 사용하여 만든 PLS 모델이 가장 선호된다고 할 수 있다. 이 모델의 상관계수 R²의 값이 0.99 이상으로 모든 성분에 대해 최종 제품의 품질관리를 위한 실험실 분석도 대체할 정도의 정확도를 보여주고 있다(Table 4). Table 4의 결과 중 Savitzky-Golay법의 결과를 정성적으로 나타내기 위해 개별 성분별 검량 모델을 이용하여 Fig. 8, 9에 바이오 디젤과 불순물들에 대한 실제 값 대비 예측 값의 그래프를 나타내었다. Table 4의 1에 가까운 R² 값들에서 알 수 있듯이 검량 모델로 예측한 값이 거의 y = x 직선에 매우 근접한 것을 확인할 수 있다.

4. 결 론

본 연구에서는 바이오 디젤 및 이에 포함된 methanol, glycerine 등 5가지 주요 불순물들에 대한 정량분석을 위해 적외선 분광 분석법과 화학계량학을 접목시켜 정확한 농도 예측이 가능한 시스템을 개발 하였다. 또한 바이오 디젤 시료의 스펙트럼을 이용하여 여러 전처리 방법의 스펙트럼 보완 및 노이즈 감소를 비교하였는데, 전처리 방법중에서 Savitzky-Golay 법이 가장 효과적임을 알 수 있었다.

이 연구에서 설계한 바이오디젤 실시간 모니터링 시스템은 ASTM 등의 표준에서 제시하는 분석방법보다 경제적 그리고 측정시간의 측면에서 우수하다. 즉, 한 가지 분석 장비로 하나의 성분만을 측정할 수 있는 기존의 분석방법과는 달리, 하나의 분석 시스템으로 여러 성분을 동시에 측정하는 것이 가능하기 때문에 분석에 관련된 제반 비용의 절감 또한 가능할 것으로 기대한다. 또한 광섬유 프로브 센서 같은 시료링 액세서리를 사용하면 분광분석 장치가 원거리에도 in-line으로 운영이 가능하기 때문에 예측 값을 공장 운전에도 피드백하는 것 또한 가능하다.

감사의 글

이 논문은 2011학년도 부경대학교의 지원을 받아 수행된 연구임(PK-2011-0760).

참고문헌

1. ASTM Standard D6751, 'Standard Specification for Biodiesel Fuel Blend Stock (B100) for Middle Distillate Fuels', ASTM International, West Conshohocken, PA (2011).
2. H. B. Sim and J. H. Lee, *Mon. Chem. Eng.*, **3**(11), 54-59 (1999).
3. H. B. Sim and J. H. Lee, *Mon. Chem. Eng.*, **3**(12), 90-93 (1999).
4. R. M. Silverstein, G. C. Bassler and T. C. Morrill, 'Spectrometric Identification of Organic Compounds', 4th ed., Wiley, New York (1981).
5. H. C. Chung and H. J. Kim, *Anal. Sci. Technol.*, **13**(1), 138-151 (2000).
6. S. Y. Kim, B. S. Lee, C. B. Chung and S. H. Choi, *J. Control Autom. Sys. Eng.*, **11**(6), 551-552 (2005).
7. S. Wold, *Technometrics*, **20**, 397-405 (1978).
8. A. Hoskuldsson, *J. Chemometr.*, **2**, 211-228 (1988).
9. R. N. Feudale, N. A. Woody, H. Tan, A. J. Myles, S. D. Brown and J. Ferre, *Chemo. Intell. Lab. Sys.*, **64**, 181-192 (2002).
10. B. M. Wise, N. B. Gallagher, R. Bro, J. M. Shaver, W. Windig and R. S. Koch, 'Chemometrics Tutorial for PLS Toolbox and Solo', Eigenvector Research Inc., (2006).
11. P. Geladi, D. MacDougall and H. Martens, *Appl. Spectrosc.*, **39**, 491-500 (1985).
12. R. J. Barnes, M. S. Dhanoa and S. J. Lister, *Appl. Spectrosc.*, **43**, 772-777 (1989).
13. S. Wold, H. Antti, F. Lindgren and J. Ohman, *Chemo. Intell. Lab. Sys.*, **44**(1-2), 175-185 (1998).