

Development of Measuring Technique for Milk Composition by Using Visible-Near Infrared Spectroscopy

Chang-Hyun Choi¹, Hyun-Woong Yun¹ and Yong-Joo Kim^{2*}

¹*School of Life Science and Biotechnology, Sungkyunkwan University, Suwon 440-746, Korea*

²*LS Mtron LTD, Anyang 431-080, Korea*

가시광선-근적외선 분광법을 이용한 유성분 측정 기술 개발

최창현¹ · 윤현웅¹ · 김용주^{2*}

¹성균관대학교 생명공학부, ²(주)LS엠트론

Abstract

The objective of this study was to develop models for the predict of the milk properties (fat, protein, SNF, lactose, MUN) of unhomogenized milk using the visible and near-infrared (NIR) spectroscopic technique. A total of 180 milk samples were collected from dairy farms. To determine optimal measurement temperature, the temperatures of the milk samples were kept at three levels (5°C, 20°C, and 40°C). A spectrophotometer was used to measure the reflectance spectra of the milk samples. Multilinear-regression (MLR) models with stepwise method were developed for the selection of the optimal wavelength. The preprocessing methods were used to minimize the spectroscopic noise, and the partial-least-square (PLS) models were developed to prediction of the milk properties of the unhomogenized milk. The PLS results showed that there was a good correlation between the predicted and measured milk properties of the samples at 40°C and at 400~2,500 nm. The optimal-wavelength range of fat and protein were 1,600~1,800 nm, and normalization improved the prediction performance. The SNF and lactose were optimized at 1,600~1,900 nm, and the MUN at 600~800 nm. The best preprocessing method for SNF, lactose, and MUN turned out to be smoothing, MSC, and second derivative. The Correlation coefficients between the predicted and measured fat, protein, SNF, lactose, and MUN were 0.98, 0.90, 0.82, 0.75, and 0.61, respectively. The study results indicate that the models can be used to assess milk quality.

Key words : Milk properties, Visible-NIR spectroscopy, PLS, MLR

서 론

국내 낙농 산업은 우유의 생산량과 소비량이 증가하는 반면, 유제품 수입의 자유화로 값싸고 품질이 우수한 수입 제품이 증가하여 우유의 자급률이 하락하는 수급 불균형 현상이 나타나고 있다(1). 그러므로 우유의 수급률을 조절하고 국내 우유의 국제 경쟁력을 높이기 위해서는 고품질 우유의 생산이 필요한 실정이다.

고품질 우유란 풍미가 좋고 단백질, 지방 등의 유성분 함량이 풍부하며 세균수와 체세포수가 적은 신선한 우유를 말하며, 우유의 품질에 영향을 미치는 유성분에는 지방(fat),

단백질(protein), 유당(lactose), 무지고형분(solid not fat, SNF) 등이 있다(2). 또한 요소태질소(milk urea nitrogen, MUN)는 사료 중 단백질과 당, 전분 등 비구조성 탄수화물과의 균형 상태를 반영하는 요소로 젖소의 건강상태를 나타내는 중요한 요소이다(3).

현재 우유의 품질판정 방법에는 크게 전문가에 의한 이화학적 분석법과 자동화된 장치를 이용한 분석법(Combifoss, FT6000, Denmark)이 있다. 이화학적 분석법은 지방의 경우 Gerber법, 단백질은 Kjeldahl법, 유당은 HPLC법, 체세포수는 직접현미경법 등의 서로 다른 방법을 사용하며(4-5), 전 처리가 복잡하여 오랜 시간이 소요되므로 온도 변화 등에 민감한 우유 시료를 일반 축산 농가에서 분석하기에는 어려운 실정이다. 자동화된 장치를 이용한 분석법은 이화학적

*Corresponding author. E-mail : babina@lsmtron.com
Phone : 82-31-688-5383, Fax : 82-31-290-7830

분석방법의 문제를 해결할 수 있으나 장비의 가격이 비싸 일부 유가공 업체에서만 사용이 가능하고 규모가 크므로 이동 중에 사용하거나 휴대할 수 없어 우유 생산현장에서 실시간 품질측정이 불가능한 단점이 있으므로 이를 해결할 수 있는 시스템이 필요한 실정이다(6). 이에 반해 근적외선 분광분석법은 화학적 전처리 없이 현장에서 실시간으로 비파괴적 검사가 가능하고 비교적 사용방법이 간단하여 많은 연구가 이루어지고 있다(7).

가시광선-근적외선 분광분석법을 이용한 유성분 예측에 관한 연구는 활발히 진행되고 있으며, Tsenkova 등(8)은 1,100~2,400 nm 근적외선 영역에서 지방, 단백질, 유당의 예측모델을 개발한 결과 결정계수(r^2)는 각각 0.98, 0.71, 0.74로 유성분의 예측이 가능함을 확인하였으며, Kawasaki 등(9)은 600~1,050 nm에서 요소태질소를 측정하여 0.53의 결정계수를 얻었다고 보고하였다. Choi 등(10)은 무지고형분의 예측모델을 개발한 결과, 결정계수는 0.84로 400~2,500 nm 영역의 반사 스펙트럼을 이용하여 무지고형분의 측정이 가능하다고 하였으며, 가시광선-근적외선 스펙트럼은 화학적 성분 정보 외에도 여러 가지 물리적 인자와 관련된 정보 및 노이즈를 포함하고 있으므로 예측모델의 신뢰성을 높이기 위해 다양한 수학적 전처리가 필요하다고 보고하였다. Pravdova 등(11)은 수학적 전처리를 이용하여 원유의 체세포수 예측모델의 성능을 결정계수 0.79에서 0.87로 향상시켜 그 필요성을 입증하였으며, 동시에 원유의 단백질 및 지방은 온도에 따라 민감하게 변화하므로 유성분의 예측성능을 향상시키기 위해서는 스펙트럼 측정을 위한 원유 시료의 최적온도 선정이 필요하다고 보고하였다.

따라서 본 연구에서는 원유의 실시간 휴대용품질측정 시스템 개발을 위한 기초 연구로써 원유 시료의 온도에 따른 가시광선-근적외선 스펙트럼을 측정하여 원유 시료의 최적온도를 선정하였으며, 다양한 수학적 전처리 방법을 적용하여 유성분 예측모델을 개발하였다.

재료 및 방법

원유 시료

원유 시료는 한국식품연구원과 협조하여 대부분 2008년 10월~12월경에 낙농 농가에서 수집한 시료를 중심으로 각각의 유성분 함량이 균등하게 분포하도록 총 180개의 시료를 선별하였다. 원유 시료는 미생물 등의 번식에 따른 환경변화의 오차를 최소화하기 위해 방부제를 처리한 후 아이스박스에 넣어 냉장상태(5°C 이하)로 보관하여 이동한 후 1일 이내에 스펙트럼 측정을 완료하였고, 측정이 끝난 시료는 식품공전에 고시된 방법으로 이화학적 분석을 실시하였다(12).

유지방 측정

유지방은 Gerber법에 따라 원유 시료 11 mL를 취해 황산으로 단백질, 유당을 분해시킨 후 이소아밀알콜 1 mL를 가해주어 지방의 분리를 촉진시키고 50~60°C 수욕조에 15분간 담근 후 1,100 rpm 에서 원심분리하여 그 용량을 측정하였다(12).

단백질 측정

원유의 단백질은 Kjeldahl법 중에서 semi-micro법을 이용하여 측정하였다(12). 농황산 7 mL, 분해 촉매($\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ 각 1g, K_2SO_4 9 g) 2 g, 작은 유리알을 넣은 Kjeldahl 분해관에 시료 2 g을 넣고 처음 10분은 약한불에서 가열하고 다음 5시간은 고온에서 가열하여 용액의 색이 무색투명하게 된 후 냉각하였다. 그 후, 4% 붕산액 10 mL에 혼합지시약(methyl red & methylene blue)과 50% NaOH액 10 mL를 가하고 분해시료를 주입 한 후 N/10 HCL액으로 적정하여 적자색이 될 때까지의 소요량을 계산하는 방법으로 단백질량을 산출하였다.

유당 측정

유당은 HPLC (high performance liquid chromatography, Waters, USA)를 이용하여 분석하였다(12). 시험용액은 시료 5 g을 정밀히 칭량하여 물과 acetonitrile로 용해시킨 후 45 μm 의 membrane filter로 여과하여 사용하였다. 물과 acetonitrile을 17:83 (v/v)으로 혼합하여 1.0 mL/min으로 흘러주었고 주입량은 10 μL 로 하여 측정하였다.

무지고형분 측정

무지고형분은 원유를 밀지름이 5 cm인 칭량관에 정제하여 15 g과 작은 유리봉을 넣고 100°C의 건조기에서 항량이 될 때까지 건조한 다음, 시료 약 5 g을 정밀히 달아 앞의 칭량관에 넣고 수욕 상에서 내용물을 때때로 저어 섞으면서 가열하였다. 대부분의 수분을 증발시킨 후 앞의 건조기에 옮겨 항량이 될 때까지 건조시켜 건조물의 질량을 구하고 이 건조물질의 함량과 지방의 함량의 차이를 무지고형분의 함량으로 계산하였다(12).

요소태질소 측정

요소태질소는 원유 샘플 내의 요소함량에 따라서 반응되는 효소의 pH 변화를 감지하는 국제낙농연합회(International Dairy Federation, IDF)에 공인된 요소 분석 장비(CL-10 Micro, Eurochem, Italy)를 이용하여 측정하였다.

원유 유성분 분포

표준화된 이화학적 분석법을 이용하여 유성분을 측정할 결과는 Table 1과 같다. 지방 함량의 범위는 2.44~6.42%, 평균 4.05%로, 국내 유대지급 기준치인 3.4%보다 높게 나타났다. 단백질 함량의 범위는 2.44~4.28%, 평균 3.35%로

나타났으며, 국내 유대지급 기준에는 없으나 영국 등의 유대지급 기준인 3.3%보다 높게 나타났다. 무지고형분 함량의 범위는 7.85~9.57%, 평균 8.76%로 권장 범위인 8.0~8.6%보다 다소 높게 관찰되었다. 유당 함량의 범위는 3.93~5.24%, 평균 4.74%로 노르웨이의 유대지급 기준 4.85%보다는 다소 낮게 나타났다. 요소태질소의 함량의 범위는 4.6~15.1 mg/dl, 평균 10.27 mg/dl로 덴마크의 권장기준인 8.4~14.0 mg/dl에 대부분 만족하지만 범위를 벗어나는 경우도 있어 일부 젖소가 단백질 불균형 상태임을 알 수 있다. 유성분 예측모델에 사용된 원유 시료의 일부는 국내의 기준을 벗어나는 경우도 있었으나 대부분 권장 기준을 만족하였다.

Table 1. Chemical properties of raw milk tested

(N=180)				
	Avg.	Max.	Min.	Std.
Fat (%)	4.05	6.42	2.44	0.82
Protein (%)	3.35	4.28	2.44	0.41
SNF (%)	8.76	9.57	7.85	0.36
Lactose (%)	4.74	5.24	3.93	0.24
MUN (mg/dl)	10.27	15.1	4.6	2.46

실험 방법

원유의 유성분 예측모델 개발을 위한 전체적인 실험 방법은 Fig. 1과 같다. 원유 시료는 낮은 온도에서는 지방이 응고하고 높은 온도에서는 단백질이 변성되는 등의 온도에 따라 성분이 민감한 변화를 보이기 때문에(13) 원유 온도에 따른 스펙트럼의 유성분 예측 성능을 구명하기 위해 동일 시료에 대해 온도 조건을 냉장 온도인 5°C, 상온인 20°C, 착유 온도인 40°C의 3 수준으로 설정하여 스펙트럼을 측정하였다. 이때 원유 시료의 온도는 향온 수조(HB-205WS, Hanbaek, Korea)를 이용하여 유지하였다. 최적 온도는 넓은 파장영역을 이용하여 예측모델을 개발하므로 화학적 조성이 복잡한 액체 시료의 예측모델 개발에 용이한(14) 부분최소자승법(Partial Least Squares, PLS)을 이용하여 전체 파장영역에서 시료의 온도에 따른 예측성능을 이용하여 선정하였다. 최종 유성분의 예측모델은 선정된 최적 온도

에서의 스펙트럼에 대해 다양한 수학적 전처리 방법을 이용하여 스펙트럼의 산란 및 노이즈를 보정한 후, 부분최소자승법을 이용하여 개발하였다. 이때 각각의 유성분에 따른 파장영역의 선택은 다중회귀분석법(Multi Linear Regression, MLR)의 stepwise 방법을 이용하였다.

스펙트럼 측정

스펙트럼은 원유 시료에 대하여 Fig. 2와 같은 고성능 분광분석기(Model 6500, NIR System, USA)의 horizontal 모듈과 샘플 셀 및 반사판을 이용하여 400~2,498 nm의 파장영역에서 2 nm 간격으로 측정하였다. 반사 스펙트럼은 분광분석기를 이용하여 기준(reference) 스펙트럼을 측정하고 원유 시료를 샘플 셀에 부은 후, 반사판을 넣고 시료가 완전히 정지하면 반사 스펙트럼을 측정하는 방법으로 5°C, 20°C, 40°C의 온도 순서로 모두 540점의 시료를 측정하였다. 기준 스펙트럼은 항상 일정한 흡광도를 나타내는 세라믹판을 이용하였으며, 반사 스펙트럼은 시료에 광선을 32회 주사하여 평균한 스펙트럼을 흡광도(log(1/R))로 변환하여 저장하였다. 샘플 셀은 지름 50 mm, 높이 30 mm의 크기로 바닥은 두께 3 mm의 수정으로 제작하였고 반사판은 지름 35 mm의 알루미늄으로 평평하게 제작한 후 반사율을 높이기 위해 금으로 도금하였다. 이때 반사판의 높이는 원유 시료가 다른 액체 시료보다 투과도가 낮은 것을 고려하여 광투과거리(path-length)를 1 mm로 제작하였다(15).

부분최소자승법

원유 시료의 온도에 따른 유성분 예측모델을 개발하기 위해 전체 스펙트럼 개수의 50%를 교정부로, 50%를 검증부로 분류하였다. 스펙트럼은 각각의 유성분에 대해 크기 순서로 정렬한 후에 순차적으로 교정부와 검증부로 분류하여 교정부와 검증부가 유사한 크기와 범위를 갖도록 조정하였다. 예측모델은 다양한 화학적 성분조성을 갖는 시료 분석에 적합한 부분최소자승법을 이용하여 원유 시료의 스펙트럼과 유성분과의 상관관계를 분석하였으며, 분석은 상용 프로그램인 Unscrambler (version 9.7, Camo, Norway)를 사용하였다. 개발된 모델의 신뢰성을 높이기 위해 교차검증(cross validation)을 이용하였으며, 주성분(principal component,

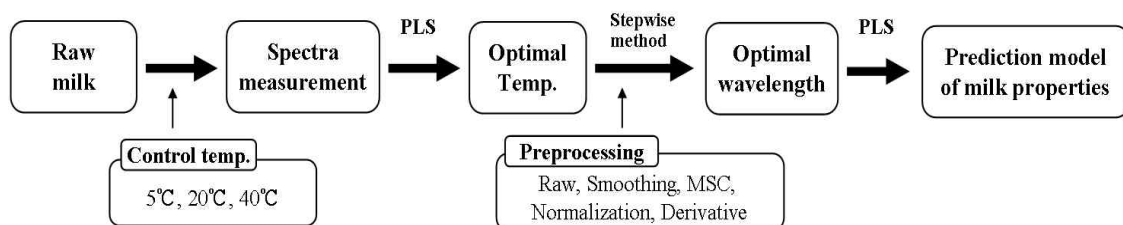


Fig. 1. Experiment method for development model of milk properties using spectroscopic measuring technique.



Fig. 2. Sample cell, reflectance plate(left) and horizontal type spectrophotometer(right) for reflectance spectra measurements.

PC) 수가 클수록 과적합(overfitting)이 발생하여 예측모델의 신뢰성이 떨어지므로 최대 주성분은 20으로 설정하고 (10) 각각의 주성분에 대한 결정계수(r^2)와 교정부 오차(standard error of calibration, SEC)를 비교하여 최적 모델을 선정하였다. 이때, 개발된 모델의 예측성능을 검증하기 위해 검증부 결정계수, 검증부 오차(standard error of prediction, SEP), RPD (ratio of standard deviation to SEP)를 이용하였는데, 검증부 결정계수는 0.66 미만, 0.66 ~ 0.81, 0.82~0.90, 0.91 이상의 범위에 대해 각각 상관관계가 낮음(poor), 유의함(approximate), 높음(good), 우수함(excellent)을 의미하며(16), 예측모델의 정밀도를 의미하는 RPD는 2.5~5 이면 높음, 5 이상이면 우수하다고 판단한다(17).

수학적 전처리

수학적 전처리 방법은 일반적으로 액체 시료에 많이 사용되는 평활화(smoothing), 정규화(normalization), 다분산보정(multiplicative scatter correction, MSC), 미분(derivative)을 이용하였다(18).

평활화는 측정 장치의 분해능에 따라 발생하는 스펙트럼의 불연속성을 연속적으로 나타내기 위한 보간법으로 기기 자체의 측정 오차를 보정하는데 사용된다. 정규화는 스펙트럼 측정 시 외부 온도 및 광원의 미세한 변화에 따라 스펙트럼의 기준선이 상하로 이동하는 현상의 보정이 가능하며, 다분산보정은 액체 시료가 미세한 기포에 의해 빛이 산란되는 현상을 보정하기 위해 사용된다. 미분은 흡광도의 크기보다는 파장영역간의 기울기에 중점을 두므로 스펙트럼의 기준선 변화에 따른 오차를 최소화 할 수 있으며, 상대적으로 흡광도가 작은 스펙트럼의 강조가 가능하여 가시광선-근적외선 스펙트럼의 중첩현상을 줄일 수 있는 장점이 있다(11).

유성분 예측모델 개발

부분최소자승법은 파장 선택에 관한 정보를 제공하지 않아 최적 파장영역을 선정하기 위해서는 모든 파장영역에 따라 모델을 개발하여 비교해야 하기 때문에 많은 시간과

노력이 소모되는 단점이 있다. 그러므로 본 연구에서는 최적 파장 영역의 선정을 위하여 다중회귀분석의 변수선택법 중 stepwise 방법을 이용하였다. 유성분 예측모델은 stepwise 방법에 따라 선정된 파장을 기준으로 가급적 연속적인 파장영역에서 높은 상관관계를 보이는 파장을 선택하여 최적 파장영역을 선정하였으며, 예측모델의 개발 및 성능분석은 최적온도 선정과 동일한 방법으로 수행하였다.

결과 및 고찰

원유 시료의 온도에 따른 스펙트럼 변화

원유 시료의 온도 변화에 따른 스펙트럼은 Fig. 3과 같이 가시광선 영역을 포함한 500~1,200 nm의 영역에서는 차이가 없었으나 1,400~2,500 nm의 범위에서는 큰 차이를 보여 주었다. 특히, 1,400~1,600 nm와 1,900~2,200 nm는 단백질과 유당의 주요 흡수 영역으로써(19) 온도가 상승함에 따라 흡광도가 높아져 온도에 민감하게 변화함을 알 수 있다.

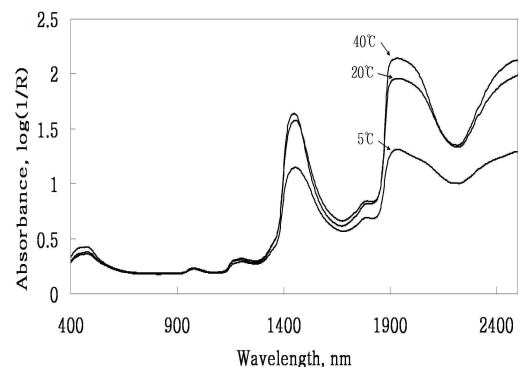


Fig. 3. Absorbance spectra of raw milk at different temperature.

최적 온도 선정

원유 시료의 온도에 따른 유성분 예측모델을 400~2,500 nm의 영역에서 개발한 결과는 Table 2와 같이 나타났다.

원유 시료의 온도에 따른 지방의 예측성능은 검증부 결정계수가 0.92~0.93, 오차는 0.15~0.30%로 거의 유사한 범위를 보여주어 원유 시료의 온도변화가 유지방의 예측성능과 무관함을 알 수 있다. 단백질의 예측성능은 원유 시료의 온도에 따라 크게 변하였으며, 검증부 결정 계수와 오차는 온도 40℃에서 각각 0.92, 0.12%, 온도 20℃에서 각각 0.71, 0.20%, 온도 5℃에서 각각 0.41, 0.38%로 예측성능이 급격히 감소함을 알 수 있다. 무지고형분의 예측성능은 온도변화에 따라 검증부 결정계수가 0.80~0.85, 오차는 0.15~0.20%로 지방의 결과보다는 다소 차이는 있으나 유사한 범위를 보여주었다. 원유 시료의 온도에 따른 유당의 예측성능은 검증부 결정 계수와 오차가 온도 40℃에서 각각 0.75, 0.11%, 온도 20℃에서 각각 0.61, 0.22%, 온도 5℃에서 각각 0.32, 0.09%로 급격히 감소하였다. 또한 젖소의 단백질 균형과 밀접한 관련이 있는 요소태질소의 검증부 결정계수와 오차는 단백질의 결과와 마찬가지로 온도 40℃에서 각각 0.61, 1.56 mg/dl, 온도 20℃에서 각각 0.23, 2.16 mg/dl, 온도 5℃에서 각각 0.13, 2.35 mg/dl로 나타났다. 이것은 온도 변화에 따라 유단백질과 유당이 민감하게 변화하기 때문에 스펙트럼을 이용한 유성분 예측모델의 신뢰성을 확보하기 위해서는 원유 시료의 온도를 40℃로 유지하는 것이 필요함을 알 수 있다.

Table 2. PLS results to select optimal temperature of raw milk in wavelength 400~2,500 nm

(N=180)							
Temp.	Milk properties	Calibration			Prediction		
		PC ¹⁾	r ²	SEC ²⁾	r ²	SEP ³⁾	RPD ⁴⁾
5℃	Fat (%)	8	0.93	0.26	0.92	0.30	2.73
	Protein (%)	2	0.42	0.40	0.41	0.38	1.08
	SNF (%)	6	0.84	0.53	0.85	0.18	2.00
	Lactose (%)	3	0.43	0.09	0.32	0.09	2.67
	MUN (mg/dl)	4	0.30	2.04	0.13	2.35	1.05
20℃	Fat (%)	8	0.90	0.28	0.92	0.15	5.47
	Protein (%)	6	0.72	0.17	0.71	0.20	2.05
	SNF (%)	6	0.83	0.56	0.80	0.20	1.80
	Lactose (%)	6	0.73	0.17	0.61	0.22	1.09
	MUN(mg/dl)	11	0.97	0.40	0.23	2.16	1.14
40℃	Fat (%)	6	0.94	0.17	0.93	0.20	4.10
	Protein (%)	15	0.96	0.05	0.92	0.12	3.42
	SNF (%)	11	0.88	0.10	0.82	0.15	2.40
	Lactose (%)	16	0.89	0.06	0.75	0.11	2.18
	MUN (mg/dl)	15	0.99	0.09	0.61	1.56	1.58

¹⁾PC : principle component

²⁾SEC : standard error of calibration

³⁾SEP : standard error of prediction

⁴⁾RPD : ratio of standard deviation to SEP

최적 파장영역 선택

다중회귀분석법의 stepwise 방법으로 유성분과의 최적 파장영역을 선택한 결과는 Table 3과 같이 나타났다. 지방은 1,674, 1,704, 1,876, 500, 1,633 nm의 파장을 이용하여 예측모델을 개발한 결과 결정계수는 0.96 오차는 0.21%로 나타났으며 1,600 nm 이상의 근적외선 영역에서 큰 가중치를 나타냈다. 단백질과 스펙트럼의 상관도는 1,636, 1,726, 1,846, 1,734, 480 nm의 순서로 높은 가중치를 보여주었으며, 결정계수와 오차는 각각 0.91, 0.11%로 나타났다. 무지고형분의 예측모델은 1,782, 1,892, 1,674 nm에서 결정계수 0.87의 우수한 상관관계를 보여주어 이 영역을 이용하면 원유의 무지고형분을 오차 0.30% 이내에서 측정 가능함을 알 수 있다. 유당의 예측모델은 1,618, 1,860, 1,912 nm 영역을 조합한 경우가 가장 우수하였으나 결정계수와 오차는 0.67, 0.16%로 상관관계가 다소 낮아 다중회귀분석만으로는 정확한 예측이 어려움을 확인하였다. 요소태질소의 경우 600, 736, 816의 다른 성분들과는 달리 가시광선 영역에서 가장 우수하나 결정계수는 0.41로 매우 낮았고 오차도 1.88로 높은 것으로 나타났다.

Table 3. Results of multi-linear regression with selected wavelength for milk properties using stepwise method

(N=180)				
Milk properties	Wavelength (nm)	r ²	SEC ¹⁾	
Fat (%)	1,674, 1,704, 1,876, 500, 1,633	0.96	0.21	
Protein (%)	1,636, 1,726, 1,846, 1,734, 480	0.91	0.11	
SNF (%)	1,782, 1,892, 1,674	0.87	0.30	
Lactose (%)	1,618, 1,860, 1,912	0.67	0.16	
MUN (mg/dl)	600, 736, 816	0.41	1.88	

¹⁾SEC : standard error of calibration

수학적 전처리 결과

40℃의 원유 스펙트럼에 대해 신호 보정 및 노이즈 제거를 위해 수학적 전처리를 수행한 결과는 Fig. 4와 같이 나타났다. 전처리를 하지 않은 원유의 스펙트럼은 Fig. 4(a)와 같이 수분의 주요 흡수 영역인 1,400~1,600 nm와 1,900~2,200 nm영역에서 높은 흡광도를 보여주었다(20). 원시 스펙트럼의 신호 보정을 위해 평활화를 수행한 결과는 Fig. 4(b)와 같이 파장 사이 불연속점의 기울기를 완만하게 하여 측정 장치 자체의 노이즈가 감소됨을 알 수 있다. 정규화를 수행한 결과는 Fig. 4(c)와 같이 전체적인 스펙트럼의 기준선이 일치하였으며, MSC를 수행한 결과 Fig. 4(d)와 같이 빛의 산란에 의한 영향을 제거하여 1,400~1,600 nm사이에서 스펙트럼 간의 차이가 감소하였음을 알 수 있다. 1차 및 2차 미분을 수행한 결과는 Fig. 4(e), (f)와 같이 스펙트럼의 변화만을 나타내어 기준선이 일치되었으며, 특히 기존 스펙트럼에서 나타나지 않았던 파장영역에 대한 분석이 가능함을 알 수 있다.

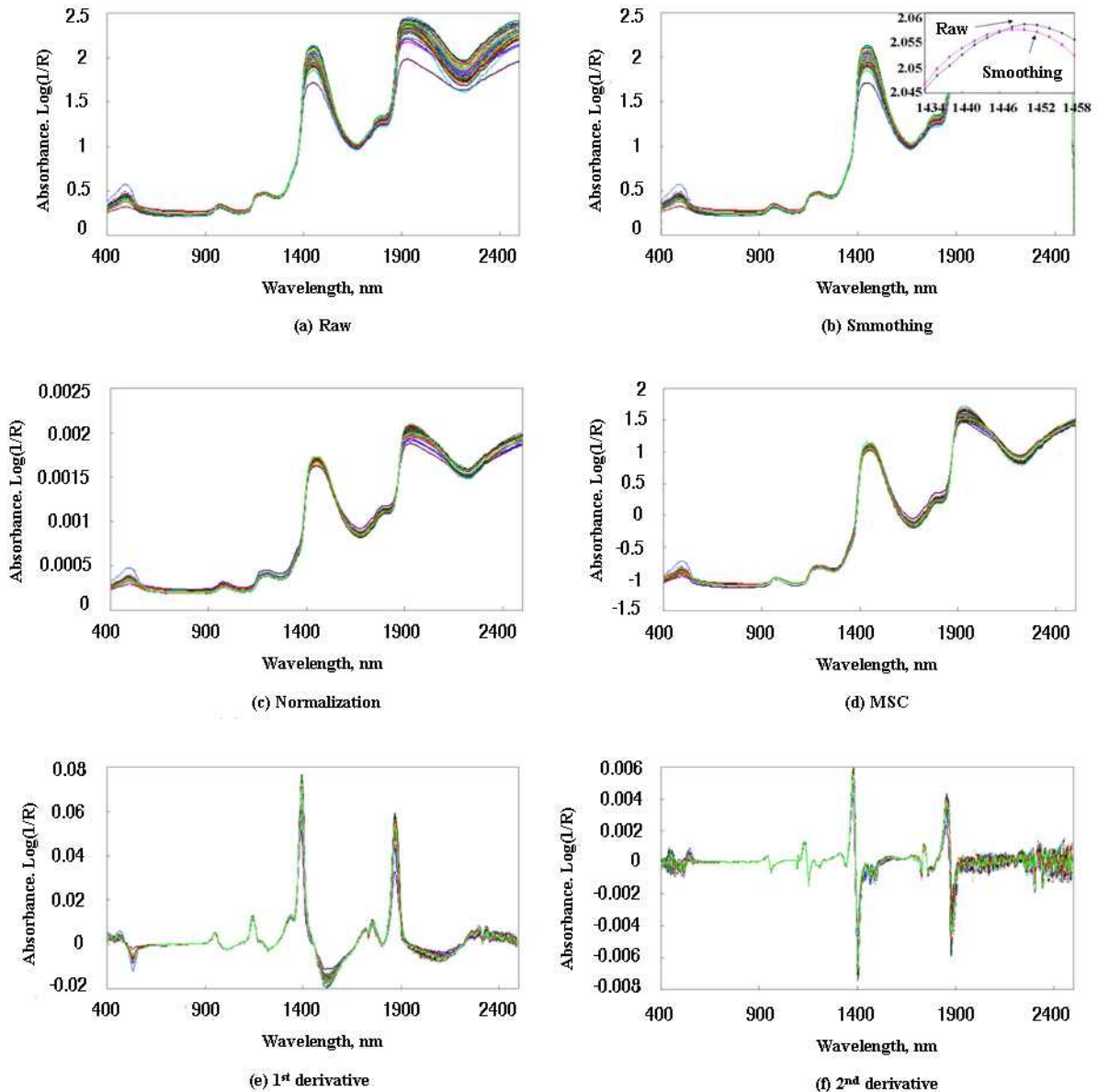


Fig. 4. Absorbance spectra of raw milk by mathematical preprocessing.

유성분 예측모델 개발

40℃의 원유 스펙트럼을 이용하여 선택된 파장영역에서 수학적 전처리 방법에 따라 유성분 예측모델을 Table 4와 같이 개발하였다. 전체적인 유성분 예측모델은 요소태질소를 제외하고 대부분 1,600~1,900 nm의 근적외선 영역에서 우수한 상관관계를 보여주었으며, 원시 스펙트럼 보다 수학적 전처리를 수행한 경우가 전반적으로 더 우수한 예측성을 보여주고 있어 사용된 수학적 전처리 방법을 이용하여 스펙트럼의 노이즈 제거 및 신호 보정이 가능함을 알 수 있다.

지방 예측모델의 검증부 결정계수와 오차는 전처리를 하지 않은 원시 스펙트럼의 경우 각각 0.93, 0.20%로 나타났

으며, 전처리 방법으로 정규화를 수행한 경우 예측성능이 전반적으로 향상되어 Fig. 5와 같이 각각 0.98, 0.15%로 가장 우수한 결과를 보여 주었으며 Tsenkova 등(8)의 연구 결과 0.98과 유사하였다. 지방의 RPD는 전처리 전과 후 각각 4.10, 5.47로 역시 가장 안정된 결과를 보여주었다. 단백질 예측모델의 검증부 결정계수와 오차는 원시 스펙트럼의 경우 각각 0.92, 0.12%, 정규화를 수행한 경우 Fig. 6과 같이 각각 0.92, 0.11%로 수학적 전처리에 따른 예측모델의 개선은 없었으나 Tsenkova 등(8)의 연구결과 0.71보다 매우 높아졌음을 확인할 수 있었고 RPD도 각각 3.41, 3.73으로 나타나 안정된 결과를 보여주었다. 무지고형분의 예측모델은 수학적 전처리 방법으로 평활화를 사용한 경우의

Table 4. PLS results with mathematical preprocessing at milk temperature of 40°C

(N=180)

Milk properties	Wavelength (nm)	Preprocessing	Calibration			Prediction		
			PC ¹⁾	r ²	SEC ²⁾	r ²	SEP ³⁾	RPD ⁴⁾
Fat (%)	1,630~1,880	Raw	6	0.94	0.17	0.93	0.20	4.10
		Normalization	6	0.99	0.16	0.98	0.15	5.47
Protein (%)	1,640~1,850	Raw	15	0.96	0.05	0.92	0.12	3.41
		Normalization	16	0.98	0.04	0.92	0.11	3.73
SNF (%)	1,680~1,900	Raw	11	0.88	0.10	0.82	0.15	2.40
		Smoothing	16	0.94	0.07	0.90	0.11	3.27
Lactose (%)	1,620~1,920	Raw	16	0.89	0.06	0.75	0.11	2.18
		MSC	16	0.89	0.07	0.80	0.11	2.18
MUN (mg/dl)	600~820	Raw	15	0.99	0.09	0.61	1.56	1.58
		2nd derivative	11	0.99	0.15	0.67	1.44	1.71

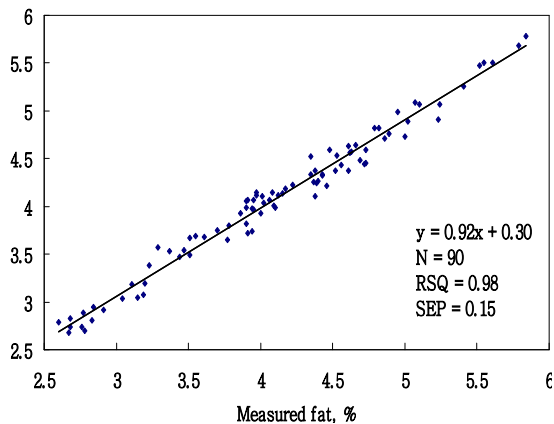
¹⁾PC : principle component²⁾SEC : standard error of calibration³⁾SEP : standard error of prediction⁴⁾RPD : ratio of standard deviation to SEP

Fig. 5. Prediction of fat in raw milk using wavelength 1,600~1,900 nm.

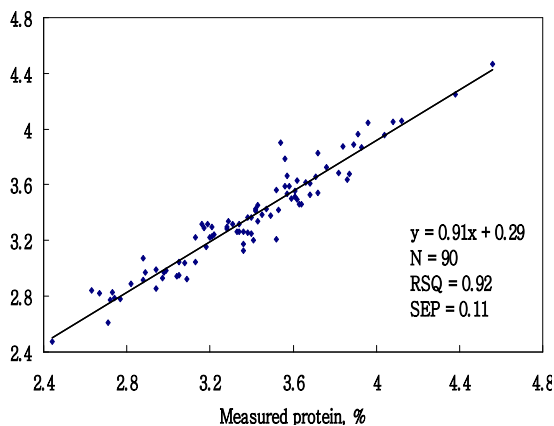


Fig. 6. Prediction of protein in raw milk using wavelength 1,600~1,800 nm.

그림 7과 같이 검증부 결정계수와 오차가 각각 0.90, 0.11%로 Choi 등(10)의 연구결과 0.84보다 높게 나타났으며 RPD

도 2.40에서 3.27로 크게 향상되었다. 유당의 예측모델도 수학적 전처리 방법으로 MSC를 사용한 경우의 그림 8과 같이 검증부 결정계수와 오차가 각각 0.80, 0.11%로 원시 스펙트럼을 사용한 경우보다 향상되었으며 Tsenkova 등(8)의 연구결과 0.74보다 높게 나타났고 RPD는 전처리 전과 후 모두 2.18로 변화가 없었다.

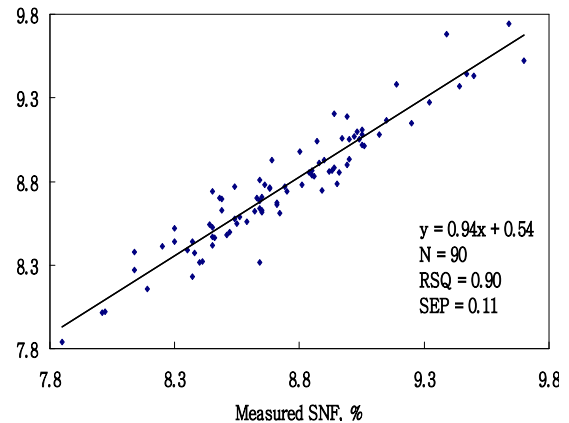


Fig. 7. Prediction of SNF in raw milk using wavelength 1,600~1,800 nm.

요소태질소의 예측모델은 다른 유성분과 달리 600~800 nm의 가시광선 영역에서 가장 우수한 상관관계를 보여주었으며, 원시 스펙트럼의 검증부 결정계수 0.61과 오차 1.56%를 개선하기 위하여 2차 미분을 사용하였으나 그림 9와 같이 검증부 결정계수와 오차는 각각 0.67, 1.44%로 Kawasaki 등(9)의 연구결과 0.53보다는 높으나 크게 개선되지 않았고 RPD 역시 1.58과 1.71로 모델의 정밀도가 다른 성분들에 비해 상대적으로 낮음을 알 수 있다. 이것은 미분을 사용하면 미세한 스펙트럼이 증폭되는 동시에 노이즈도 증가하기 때문으로 요소태질소의 신뢰성있는 모델을 개발

하기 위해서는 부분최소자승법 외에 다양한 알고리즘의 적용이 필요할 것으로 판단된다.

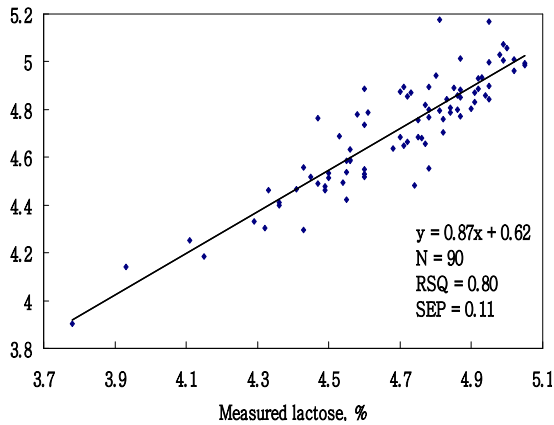


Fig. 8. Prediction of lactose in raw milk using wavelength 1,600~1,900 nm.

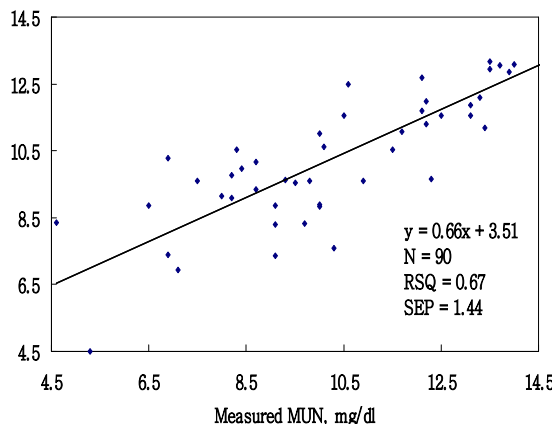


Fig. 9. Prediction of MUN in raw milk using wavelength 1,600~1,800 nm.

요 약

본 연구는 원유의 실시간 휴대용품질측정 시스템 개발을 위한 기초 연구로서 원유 시료의 온도에 따른 가시광선-근적외선 스펙트럼을 측정하였으며, 다양한 수학적 전처리방법을 적용하여 유성분 예측모델을 개발하였다. 스펙트럼 측정은 원유 시료 180개에 대해 스펙트럼의 수학적 전처리 방법으로 평활화, 정규화, MSC, 1차 및 2차 미분을 사용하였고 예측모델은 부분최소자승법을 이용하였다.

유성분을 분석한 결과 함량 범위와 평균은 지방이 각각 2.44~6.42%, 4.05%, 단백질은 각각 2.44~4.28%, 3.35%, 무지고형분은 각각 7.85~9.57%, 8.76%로 나타났다. 또한 유당의 함량 범위와 평균은 각각 3.93~5.24%, 4.74%였으며 요소태질소의 경우에는 각각 4.6~15.1 mg/dl, 10.27 mg/dl로 대부분 권장 기준을 만족하였다.

원유 시료의 온도에 따른 스펙트럼은 1,400~2,500 nm에서 큰 차이를 보였으며 온도가 상승함에 따라 흡광도가 높아지는 것을 알 수 있었다. 원유 시료의 온도에 따른 유성분 예측모델을 400~2,500 nm의 영역에서 개발하였으며 예측성능은 지방과 무지고형분의 경우 온도변화와 무관하였다. 단백질과 유당, 요소태질소의 예측성능은 온도가 낮을수록 급격히 감소하여 스펙트럼 측정 시 원유 시료의 온도를 40℃로 유지하는 것이 필요함을 알 수 있다.

40℃의 원유 스펙트럼에 대해 수학적 전처리를 수행한 결과 평활화를 수행하여 측정 장치 자체의 노이즈를 감소시킬 수 있었고 정규화를 수행하여 기준선을 일치시킬 수 있었다. 또한 MSC를 수행하여 빛의 산란에 의한 영향을 제거하여 스펙트럼간의 차이를 감소시킬 수 있었고 1차 및 2차 미분을 수행한 결과 기준선 일치와 기존 스펙트럼에서 나타나지 않았던 과장영역에 대한 분석이 가능함을 알 수 있다.

다중회귀분석의 stepwise 방법을 이용하여 최적 과장영역을 선정하고 유성분 예측모델을 개발한 결과 요소태질소를 제외하고 대부분 근적외선 영역에서 우수한 상관관계를 보여주었다. 지방과 단백질은 원시 스펙트럼의 검증부 결정계수가 각각 0.93, 0.92에서 정규화를 수행한 결과 각각 0.98, 0.92로 원시 스펙트럼의 결과가 우수하여 큰 개선이 없었으나 RPD는 각각 4.10, 3.41에서 5.47, 3.73으로 높아져 정밀도가 향상됨을 알 수 있다. 무지고형분과 유당의 예측모델은 원시 스펙트럼의 경우 각각 0.82, 0.75로 예측모델로 사용하기에는 어려웠으나 각각 평활화와 MSC를 수행하였을 때 검증부 결정계수가 0.90, 0.80으로 크게 개선되어 유성분 예측모델의 신뢰성 향상에 기여할 수 있을 것으로 판단된다. 요소태질소의 경우 가시광선 영역에서 가장 우수한 상관관계를 보여주었으나 검증부 결정계수, 오차, RPD가 각각 0.61, 1.56%, 1.58로 다른 성분에 비해 매우 낮게 나타났다. 이를 개선하기 위해 수학적 전처리를 수행하였으나 크게 개선되지 않았으므로 요소태질소의 신뢰성 있는 모델을 개발하기 위해서는 부분최소자승법 외에 다양한 알고리즘의 적용이 필요할 것으로 판단된다.

참고문헌

1. National Veterinary Research & Quarantine Service (2004) Somatic cell control plan according to changing standards of milk selling price. Seoul, Korea, p 1-2
2. Kim KS, Noh HW, Lim SD, Choi CH, Kim YJ (2008) Development of rapid somatic cell counting method by dye adding nir spectroscopy. Korean J Food Sci An, 28, 63-68
3. Roseler DK, Ferguson JD, Sniffen CJ, Herrema J (1993)

- Dietary protein degradability effect on plasma and milk urea nitrogen and milk nonprotein nitrogen in Holstein cows. *J Dairy Sci*, 76, 525-534
4. Jung HO (2003) Food and nutrition. MJ media, Seoul, Korea, p 128-180
 5. Song JC (2006) The newest food technology. Yu Lim Mun Hwa Sa, Seoul, Korea, p 64-68
 6. Berglund I, Pettersson G, Sjaunja S (2002) Automatic milking: effects on somatic cell count and teat end-quality. *Livest Prod Sci*, 78, 115-124
 7. Shenk JS, Westerhaus MO (1993) Analysis of agriculture and food products by near-infrared reflectance spectroscopy. Infrasoftware International, Port Matilda, Pennsylvania, USA, p 63-66
 8. Tsenkova R, Atanassova S, Toyoda K, Ozaki Y, Itoh K, Fearn T (1999) Near-infrared spectroscopy for dairy management: measurement of unhomogenized milk composition. *J Dairy Sci*, 82, 2344-2351
 9. Kawasaki M, Kawamura S, Tsukahara M, Morita S, Komiya M, Natsuga M (2008) Near-infrared spectroscopic sensing system for on-line milk quality assessment in a milking robot. *Comput Electron Agr*, 63, 22-27
 10. Choi CH, Kim YJ, Kim KS, Choi TH (2008) Development of Measuring Technique for Somatic Cell Count in Raw Milk by Spectroscopy. *J of Biosystems Eng*, 33, 210-215
 11. Pravdova V, Walczak B, Massart DL, Kawano S, Toyoda K, Tsenkova R (2001) Calibration of somatic cell count in milk based on near-infrared spectroscopy. *Anal Chim Acta*, 450, 131-141
 12. KFDA (2011) Food code. Korea Food and Drug Administration, Cheongwan, Korea, p 341-369
 13. Tsenkova R, Atanassova S, Ozaki Y, Toyoda K, Itoh K (2001) Near-infrared spectroscopy for biomonitoring influence of somatic cell count on cow's milk composition analysis. *Int Dairy J*, 11, 779-783
 14. Martens H, Naes T (1987) Multivariate calibration by data compression. In: *Near-infrared Technology in the Agricultural and Food Industries*. Williams P and Norris KH(Editor), American Association of Cereal Chemists, Saint Paul, Minnesota, USA
 15. Chen JY, Iyo C, Kawano S (1999) Development of calibration with sample cell compensation for determining the fat content of unhomogenized raw milk by a simple near infrared transmittance method. *J Near Infrared Sepc*, 7, 265-273
 16. Karoui R, Mouazen AM, Dufour E, Pillonel L, Picque D, Bosset JO, De Baerdemaeker J (2006) Mid-infrared spectrometry: A tool for the determination of chemical parameters in emmental cheeses produced during winter. *Lait*, 86, 83-97
 17. Williams PC (1996) Recent advances in near-infrared applications for the agriculture and food industries. Presented at the International Japanese Conference on Near Infrared Spectroscopy, November 20-21, Tsukuba, Japan
 18. Noh SH, Ryu DS (2002) Preprocessing of Transmitted Spectrum Data for Development of a Robust Non-destructive Sugar Prediction Model of Intact Fruits, *J Korean Soc Nondestruct Test*, 22, 361-368
 19. Purnomoadi A, Keshab K, Batojoo, Ueda K, Terada F (1999) Influence of feed source on determination of fat and protein in milk by near-infrared spectroscopy. *Int Dairy J*, 9, 447-452
 20. Osborne BG, Fearn T (2000) *Near infrared spectroscopy in food analysis*. Longman Scientific & Technical, Newyork, USA, p 35-40

(접수 2011년 8월 12일 수정 2011년 12월 14일 채택 2011년 12월 23일)