

과실의 비파괴 당도 예측 모델의 성능향상을 위한 투과스펙트럼의 전처리

Preprocessing of Transmitted Spectrum Data for Development of a Robust Non-destructive Sugar Prediction Model of Intact Fruits

노상하*[†], 류동수**
 Sang Ha Noh*[†] and Dong Soo Ryu**

초 록 본 연구는 초당 2개의 속도로 이송되는 사과를 대상으로 측정된 투과 에너지 스펙트럼 데이터를 이용하여 사과 당도예측 모델을 개발하기 위해 각종 전처리가 당도 예측 모델의 정밀도에 미치는 영향을 구명하고, 신뢰성이 높은 당도 예측 회귀 모델을 개발하기 위해 수행되었다. 스펙트럼의 산란 보정, 노이즈 감소 등을 위해 1차미분, MSC, SNV, OSC 및 이들 조합으로 구성된 전처리 알고리즘을 프로그래밍하고, 이들 전처리를 스펙트럼데이터에 적용한 결과 특히 MSC와 SNV에 의해 각 파장에서의 투과에너지와 당도와의 상관관계가 전처리를 하지 않은 경우에 비해 현저히 증가하였다. 각종 전처리를 수행한 후 당도 예측 회귀 모델을 개발하고, 검증한 결과, 전처리 방법에 따라 예측모델의 SEP가 최대 1.265 %brix에서 최소 0.507 %brix로 큰 차이를 나타내었다. 이는 SEP를 최소화하기 위해 주어진 스펙트럼 데이터의 특성에 알맞는 전처리 방법이 개발 또는 선택되어야 함을 의미한다. MSC와 SNV는 예측 정밀도와 밀접한 관계가 있으며, OSC는 PLS의 factor 수와 관계되는 것으로 판단되었다. 1차미분은 오히려 모델의 예측 성능을 저하시키는 것으로 나타났다. 이는 실시간으로 측정된 투과스펙트럼에 상대적으로 노이즈 성분이 많이 포함되어 이들 성분이 미분에 의해 강조된 것으로 판단되었다. 본 연구에 사용된 스펙트럼 데이터의 경우 MSC와 OSC 전처리를 수행한 당도예측모델이 $R^2=0.8823$, $SEP=0.5071$ %brix, $bias=0.0327$ 로 가장 우수하였다.

주요용어: 과실, 당도, 예측모델, 전처리, 온라인

Abstract The aim of this study was to investigate the effect of preprocessing the transmitted energy spectrum data on development of a robust model to predict the sugar content in intact apples. The spectrum data were measured from 120 Fuji apple samples conveying at the speed of 2 apples per second. Computer algorithms of preprocessing methods such as MSC, SNV, first derivative, OSC and their combinations were developed and applied to the raw spectrum data set. The results indicated that correlation coefficients between the transmitted energy values at each wavelength and sugar contents of apples were significantly improved by the preprocessing of MSC and SNV in particular as compared with those of no-preprocessing. SEPs of the prediction models showed great difference depending on the preprocessing method of the raw spectrum data, the largest of 1.265 %brix and the smallest of 0.507 % brix. Such a result means that an appropriate preprocessing method corresponding to the characteristics of the spectrum data set should be found or developed for minimizing the prediction errors. It was observed that MSC and SNV are closely related to prediction accuracy, OSC is to number of PLS factors and the first derivative resulted in

decrease of the prediction accuracy. A robust calibration model could be developed by the combined preprocessing of MSC and OSC, which showed that SEP = 0.507 %brix, bias = 0.0327 and R2 = 0.8823.

Keywords: fruit, sugar content, prediction model, preprocessing, on-line

1. 서 론

과실류의 상품가치는 크기, 색상, 모양 등 외관과 당도, 산도, 속도, 경도, 내부 부패 등 내부품질에 의해 결정되지만 내부품질 판별기술이 개발되지 않아 외관중심으로 선별되어 거래되어 왔다. 최근 농산물과 같은 생체의 성분을 비교적 신속하고, 정확하게 측정할 수 있는 근적외(NIR) 분광분석법이 개발됨에 따라 이를 이용한 과실류의 내부 품질(당도, 산도 등)을 비파괴적인 방법으로 측정하기 위한 많은 연구가 수행되어 왔다[1-6].

1980년대에는 오프라인상태에서 과실류, 채소류 등 농산물을 대상으로 반사스펙트럼을 측정하여 주요 성분을 예측하고자 하는 연구가 대부분이었으나[1,4,5], 1990년대 중반에 일본의 마끼제작소에서 투과광을 이용한 과실의 온라인 당도선별장치가 개발된 것으로 보고되고 있다. 국내에서는 최근 황(2000)등에 의해 투과스펙트럼을 이용한 복숭아의 당도 측정기술이 발표된 바 있으며, 이 연구를 바탕으로 온라인 상태에서 과실의 투과스펙트럼을 측정할 수 있는 장치를 구축하고, 과실을 대상으로 당도 및 산도를 보다 정확하게 측정할 수 있는 스펙트럼 분석방법에 대한 연구가 수행된 바 있다[2,3,6].

$$A = -\log_{10} T = \log_{10} \frac{I_0}{I} = \epsilon cd \quad (1)$$

투과 또는 반사 스펙트럼을 이용하여 과실과 같은 생체의 주요 성분을 측정하는 원리는 Beer-Lambert의 법칙에 근거한다. 즉, 입사되는 광량을 I_0 , 물질을 투과한 후 센서에 의해 측정된 광량을 I 라고 하면 물질의 흡광도 A (투과도 T)는 다음과 같이 표시된다.

여기서 ϵ 은 물질의 흡광계수, c 는 성분의 농도 그리고 d 는 투과거리이다.

이 식에 의하면 특정 성분의 농도뿐만 아니라 광의 투과 거리는 투과율 또는 흡광도와 밀접한 관계가 있다. 과실을 대상으로 투과스펙트럼을 측정하여 특정 성분의 함량

을 예측할 경우, 성분의 농도 이외에도 과실마다 크기의 차이로 인한 투과거리의 차이, 대상체 표면의 굴곡, 거칠기 등에 의한 확산반사의 차이, 과육의 경도, 결 등에 따른 산란의 차이 등이 투과스펙트럼의 강도 및 형상에 영향을 미치기 때문에 이들 물리적 인자에 의한 영향이 보정되어야 한다. 이밖에도 실시간으로 과실의 스펙트럼을 측정할 경우 광센서 자체의 노이즈나 측정환경의 변화 등에 의한 외란의 영향을 받게 된다. 따라서 이러한 구조적인 외란을 제거하기 위하여, 평활화, 산란보정, 경로보정 등을 위한 다양한 방법이 제시되어 있다.

본 연구에서는 소프트웨어적인 관점에서 과실의 당도 예측 모델 개발 및 예측성능을 향상시키기 위한 투과 에너지 스펙트럼 데이터의 전처리방법을 개발하고, 이들 전처리 효과를 구명하며, 최종적으로 신뢰성이 높은 당도 예측 모델을 개발하는데 그 목적이 있다.

2. 재료 및 방법

2.1. 투과스펙트럼 측정

과실을 대상으로 실시간 투과 스펙트럼을 측정하기 위하여 Fig. 1 및 2와 같은 실험장치를 구성하였다. 이 장치는 크게 과실 시료를 담은 접시, 접시를 이송하는 이송장치, 과실에 광을 조사하기 위한 광원, 과실을 투과한 광을 수집하기 위한 프로브, 그리고 각 파장별 투과에너지를 측정하기 위한 분광광도계로 구성된다.

과실에 입사된 광이 과실을 투과하여 프로브에 도달하도록 하기 위해서는 높은 방사에너지를 갖는 광원이 필요하기 때문에 본 연구에서는 텅스텐-할로겐등(500W) 4개를 조합하여 광원으로 사용하였다. 또한 투과된 광은 상대적으로 미약하기 때문에 이를 측정하기 위해 파장범위 550~1050nm의 고감도 분광광도계(S2000, Ocean Optics, USA)를 사용하여 투과 스펙트럼 측정장치를 구성하였다.

구성된 투과 스펙트럼 측정장치를 이용하여 초당 2개

의 속도로 후지 사과를 이송시키면서 시료(120개)의 투과 스펙트럼을 측정하고, 이를 당도 예측 모델 개발에 이용하였다. 시료 사과의 당도는 디지털 굴절당도계(Model DBX-55, ATAGO, Japan)를 이용하여 측정하였다. 실험에 사용된 시료의 당도 범위는 8.0~14.9 %brix이었으며, 평균 12.4 %brix였다.

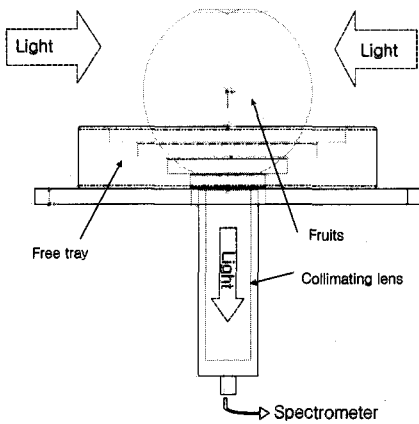


Fig. 1 Schematic diagram for measuring the transmittance spectrum of fruits



Fig. 2 View of the apparatus for measuring the transmittance spectrum

2.2. 스펙트럼 전처리

이송되는 과실을 대상으로 실시간으로 측정된 투과 스펙트럼은 이미 설명한 바와 같이 수분이나 당과 같은 화

학적 성분정보 이외에 여러 가지 물리적 인자와 관련된 정보 및 노이즈를 포함하고 있다. 스펙트럼 전처리는 성분 예측 모델 개발과정에서 가장 기본적인 단계이며, 스펙트럼에 포함된 외란을 제거하여 보다 안정적이고 정확도가 높은 예측 모델을 만들기 위하여 이용된다.

본 연구에서 검토된 각종 전처리의 종류 및 특징은 다음과 같다.

2.2.1. 평활화(Smoothing)

평활화의 목적은 스펙트럼 측정센서에 기인되는 노이즈를 감소시키는데 있다. 평활화 알고리즘은 단순 이동평균, 가중 이동평균, 다항함수법 등이 있으나 이 연구에서는 Savitzky Golay 다항함수법을 채택하였다[11].

2.2.2. 미분(derivatives)

스펙트럼의 미분은 광원의 전압 변동 등으로 인한 스펙트럼의 baseline의 이동을 제거할 뿐만아니라 스펙트럼의 미세변화 즉 미세 정보를 강조하기 위해 이용된다. 미분은 1차미분, 2차미분 등이 주로 이용되며, 미분간격($\Delta \lambda$)에 따라 미분스펙트럼의 특성이 달라지기 때문에 일반적으로 시행착오과정을 거쳐 적절한 미분간격을 구해야한다.

2.2.3. MSC(Multiplicative Scattering Correction)

투과스펙트럼은 시료의 크기, 경도 등 물리적인 특성차이로 인한 투과광의 광경로 길이 및 산란정도에 큰 영향을 받는다. MSC는 산란의 영향을 보정하기 위해 개발된 것으로 기본개념은 측정된 각 스펙트럼을 이상적인 스펙트럼을 이용하여 보정하는 것이다. 실제로는 이상적인 스펙트럼을 얻을 수 없으므로, 측정된 전체 스펙트럼의 각 파장에서 평균을 이상적인 스펙트럼으로 간주하며, 다음과 같이 보정된 스펙트럼을 구한다.

$$\bar{x}_j = \frac{\sum_{i=1}^m x_{i,j}}{m} \tag{2}$$

여기서, x = 스펙트럼 ($m \times n$), \bar{x}_j = j 번째 파장의 평균 흡광도 값($j=1, n$)이다. 이들 평균값과 각 시료 스펙트럼의 각 파장에서 흡광도 값을 이용하여 다음과 같은 일차 회귀식을 구한다.

$$x_i = \bar{a}_i x + b_i \quad (3)$$

여기서, $x_i =$ 각 파장에서 예측된 값이며, $a_i, b_i =$ 회귀계수이다. 선형회귀를 통하여 구한 a_i 및 b_i 값을 이용하여 각 시료 스펙트럼의 각 파장에서의 보정된 흡광도 값을 다음과 같이 구한다[8, 9].

$$x_{i_{MSC}} = \frac{(x_i - b_i)}{a_i} \quad (4)$$

여기서, $x_{i_{MSC}}$: MSC 보정된 값

2.2.4. SNV(Standard Normal Variate)

이것은 MSC와 마찬가지로 광산란 보정을 위한 방법으로, 목적은 동일하지만 수학적 방법은 다르다. SNV는 이상적인 스펙트럼이 필요하지 않으며, 대신에 각 스펙트럼을 전체 스펙트럼의 표준편차로 정규화하여 산란의 영향을 제거하는 방법이다. SNV보정된 스펙트럼은 무차원이 된다[7].

$$\bar{a}_i = \frac{\sum_{j=1}^n A_{i,j}}{n} \quad (5)$$

$$x_{i_{SNV}} = \frac{(x_i - \bar{a}_i)}{\sqrt{\frac{\sum_{j=1}^n (x_{i,j} - \bar{a}_i)^2}{(n-1)}}} \quad (6)$$

여기서, \bar{a}_i : i 번째 스펙트럼의 모든 파장에 대한 평균 ($n \times 1$)

$x_{i_{SNV}}$: SNV 보정된 스펙트럼 ($1 \times n$)

2.2.5. OSC(Orthogonal Signal Correction)

전처리를 하는 이유는 농도 정보와 상관이 없는 변이만을 제거하기 위한 것이지만 성분과 관련된 정보도 동시에 제거되는 경향이 있다. Wold(1998) 등이 스펙트럼(\mathbf{x})에서 성분(\mathbf{y})과 상관이 없는 변이만 제거하면 성분과 상관이 높은 스펙트럼 성분만 남는다는 개념을 제안하고, PLS(Partial Least Square)에서 계산되는 스코어(\mathbf{t})를 성분 \mathbf{y} 와 직교시켜 성분에 직교하는 스펙트럼 성

분을 계산한 다음, 이를 원래의 스펙트럼에서 제거하면 성분과 상관이 매우 높은 스펙트럼 성분만을 찾을 수 있다고 보고하였다[13]. 즉 성분 \mathbf{Y} 와 스펙트럼 \mathbf{X} 가 서로 직교하기 위해서는 PCA(Principle Component Analysis)의 기본식 및 PLS의 NIPALS 알고리즘에서,

$$\mathbf{x} = \mathbf{t}\mathbf{p}' + \mathbf{e} \quad (7)$$

여기서 \mathbf{x} = 스펙트럼 행렬($m \times n$), \mathbf{y} = 성분 벡터($m \times 1$), \mathbf{t} = 스코어벡터($m \times 1$), \mathbf{p}' = 로딩벡터($n \times 1$), \mathbf{e} = 잔차행렬($m \times n$), m = 샘플의 수, n = 파장의 수이다.

\mathbf{t} 와 \mathbf{y} 의 내적이 0($\mathbf{t}' \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{t}' = 0$)이 되면 직교하게 된다. 즉,

$$\mathbf{t}^* = (1 - \mathbf{y}(\mathbf{y}'\mathbf{y})^{-1}\mathbf{y}')\mathbf{t} \quad (8)$$

따라서, 이 \mathbf{t}^* 를 이용하여 성분 \mathbf{y} 와 상관이 없는 스펙트럼 성분($\mathbf{t}^* \times \mathbf{p}'$, OSC component)을 계산할 수 있다. 이를 원 스펙트럼에서 제거하면 상관이 높은 스펙트럼 성분을 얻을 수 있다.

$$\mathbf{x} = \mathbf{x} - \mathbf{t}^* \times \mathbf{p}' \quad (9)$$

OSC Component 수는 성분과 직교하는 스펙트럼 성분($\mathbf{t}^* \times \mathbf{p}'$)을 몇번 제거할 것인가를 나타낸다.

2.2.6. 전처리 조합

각 전처리의 효과를 알아보기 위해 Table 1과 같이 전처리 조합을 구성하여 스펙트럼 데이터 셋에 적용하였다. 여기서 원스펙트럼에 대하여 기본적으로 스펙트럼의 평활화, 미분 및 MC (Mean Centering)를 먼저 행하였다. 평활화 및 미분은 Savitzky-Golay법을 이용하였으며, 파장간격은 27nm, 다항함수의 차수는 1차로 하였다.

2.3. 정량 모델 개발 및 성능 평가

원시 스펙트럼데이터에 각종 전처리를 수행한 다음 이를 이용한 당도 측정회귀 모델은 PLSR (Partial Least Square Regression)을 이용하여 개발하였다. 이 모델의 특징은 PCR(Principal Component Regression)모델과는 달리, 농도정보를 이용함으로써, 고농도의 스펙트럼은 저농도의 스펙트럼보다 큰 가중치를 갖게 하는 방법이다.

Table 1 Combination of Preprocessing methods to correct the original spectrum data

Abbreviation	Spectrum correction method
N	No Preprocessing
S	SNV
S1	1 st Deriv. + SNV
M	MSC
M1	1 st Deriv. + MSC
O	OSC
O1	1 st Deriv. + OSC
MS	MSC+SNV
MS1	1 st Deriv + MSC + SNV
MO	MSC + OSC
MO1	1 st Deriv + MSC + OSC
SO	SNV + OSC
SO1	1 st Deriv + SNV + OSC

또한 최적 PLS factor의 결정방법은 cross - validation (leave-one out)에 의한 최소 PRESS (prediction residual error sum of squares)법으로 하였다[9].

OSC 전처리가 들어있는 모델의 경우는 적정 OSC component의 수를 다음과 같이 구하였다. 각 OSC component를 수행한 후, 제거된 스펙트럼의 분산과 원스펙트럼의 분산의 비를 계산하였으며, 이 값이 일정하게 될 때(≤ 0.01)의 OSC component수를 적정 components 수로 결정하였다[12].

$$C_i^2 = \frac{var(x_i)}{var(x)} \quad (10)$$

개발된 모델의 평가를 위해 SEC(Standard Error of Calibration), SEP(Standard Error of Prediction) 및 bias를 다음과 같이 계산하였다[10].

$$SEC = \sqrt{\frac{\sum_{i=1}^m (y_i - \hat{y}_i)^2}{m - f - 1}} \quad (11)$$

$$bias = \frac{\sum_{i=1}^m (y_i - \hat{y}_i)}{m} \quad (12)$$

$$SEP = \sqrt{\frac{\sum_{i=1}^m ((y_i - \hat{y}_i) - bias)^2}{m - 1}} \quad (13)$$

또한 본 연구에서는 각 전처리 조합에 의해 개발된 모델의 성능 평가를 위하여 다음과 같은 기준을 설정하고 평가에 고려하였다[3,6].

- ① SEC 및 SEP가 작을 것
- ② 예측모델의 bias가 작을 것
- ③ 개발된 모델의 R²와 검정시료의 R²가 클 것
- ④ 모델에 포함된 PLS factor의 수가 작을 것

상기의 스펙트럼 데이터의 전처리 알고리즘 및 정량분석 모델 개발을 위하여 Matlab (ver. 5.3)을 이용하였다.

3. 결과 및 고찰

3.1. 전처리별 스펙트럼 형태의 변화

Fig. 3은 사과 시료를 대상으로 측정된 투과에너지 스펙트럼 데이터의 개형으로 엽록소의 흡수파장인 675nm, 물의 흡수파장인 765nm 등의 대역에서 투과에너지가 상대적으로 약함을 알 수 있다. 이들 스펙트럼을 대상으로 MSC와 SNV 전처리를 수행한 결과 Fig. 4 및 5에서와 같이 y축방향의 투과에너지 변화폭이 좁아지고, 또한 그림에 확대 표시된 바와 같이 시료의 당도에 따라 투과에너지 스펙트럼의 순서가 정렬됨을 알 수 있다. 이 결과는 이들 전처리에 의해 산란에 기인되는 투과에너지의 차이가 감소됨을 의미하는 것으로 판단된다.

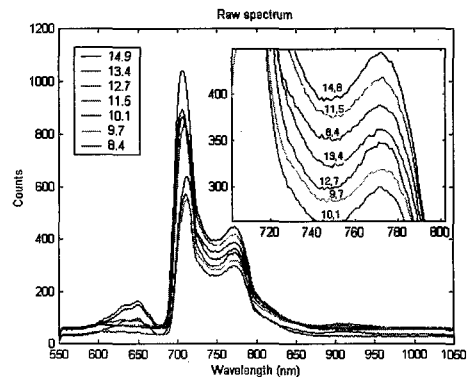


Fig. 3 Transmitted energy spectra of apple samples

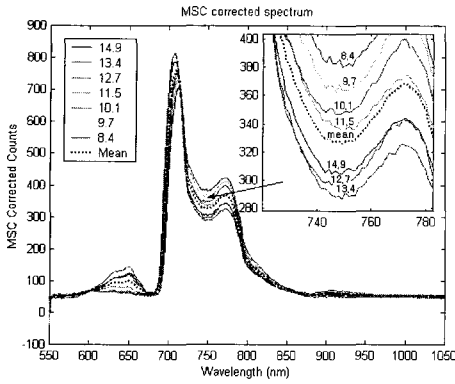


Fig. 4 Corrected spectra by MSC

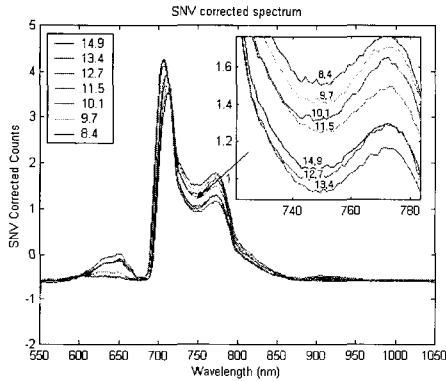


Fig. 5 Corrected spectra by SNV

Fig. 6의 OSC 전처리의 경우에도 투과에너지의 변화 폭이 감소될 뿐만아니라 시료의 당도별로 스펙트럼의 순서가 정렬됨을 알 수 있다.

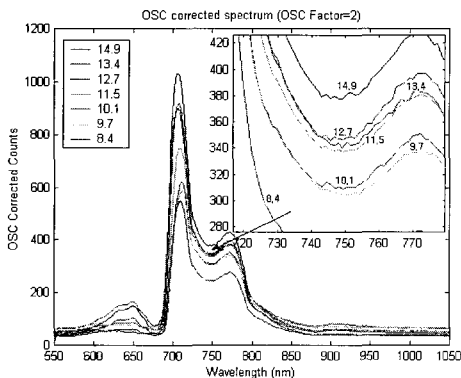


Fig. 6 Corrected spectra by OSC

3.2. 전처리에 따른 상관계수 변화

일차적으로 전처리 효과를 관찰하기 위해 측정된 원시 스펙트럼과 전처리를 수행한 후의 스펙트럼데이터를 이용하여 각 파장에서 투과에너지와 당도와의 상관계수를 구하였다. Fig.7에서 "Raw"로 표시된 것은 전처리를 하지 않은 데이터에 의한 상관계수 스펙트럼, "OSC", "MSC" 및 "SNV"로 표시된 것은 해당 전처리를 수행한 데이터에 의한 상관계수 스펙트럼이다. MSC나 SNV 전처리를 수행한 경우 전체 파장대역에 걸쳐 상관계수가 전반적으로 증가함을 알 수 있으며, 특히 750nm 근처에서 음의 상관관계가 현저히 증가함을 알 수 있다.

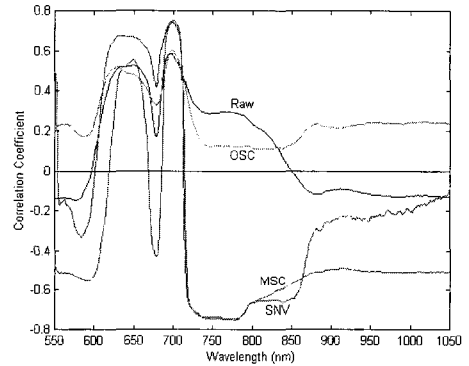


Fig. 7 Effect of spectrum-preprocessing on correlation coefficients between sugar contents and transmitted energy values at each wavelength

이는 Fig. 4 및 Fig. 5에서 보는 바와 같이 전처리에 의해 750nm부근에서 시료의 당도에 따라 스펙트럼의 순서가 정렬화되기 때문인 것으로 생각된다. 그러나 OSC 전처리의 경우에는 이 대역에서 오히려 상관계수가 저하되는 것으로 나타났다.

3.3. 각 전처리별 당도예측

실시간 투과 에너지 스펙트럼 데이터를 이용하여 각 전처리별 당도 예측 회귀모델(또는 검량식이라고도 함)을 개발하고 검증하였으며, 전처리 효과를 분석하였다(Table 2). Table 2에서 전처리를 하지 않은 경우(N)의 당도예측 표준오차(SEP)와 bias는 각각 0.948 %brix 및 0.0985 이며, 전처리 방법에 따라 이들 값이 다르게 나타남을 알

수 있다. SEP가 가장 작은 전처리 조합은 M 및 MO의 경우이며, 이때의 SEP는 0.507 %brix, bias는 각각 -0.0319와 -0.0327로서 전처리 효과가 현저히 있음을 말해주고 있다. 분석 결과 MSC와 SNV의 경우 둘다 산란보정효과가 있는 것으로 판단되며, MSC가 SNV에 비해 상대적으로 유리한 것으로 관찰되었다. 또한 MSC와 SNV의 조합의 경우(MS)에는 SNV만의 결과와 동일한 결과를 보였다. 이는 산란보정의 효과를 좀더 높이기 위해 MSC를 처리한 후 SNV를 다시 처리하더라도 아무런 추가적인 효과는 없음을 말해준다.

Table 2 Calibration and validation results of the sugar contents in intact apples depending on various preprocess methods

Prepro.	No. of factor	Calibration		Validation		
		R ²	SEC	R ²	SEP	bias
N	15	0.769	0.783	0.627	0.948	0.0985
S	12	0.893	0.495	0.879	0.515	-0.0229
S1	10	0.957	0.319	0.814	0.641	-0.0510
M	12	0.909	0.469	0.882	0.507	-0.0319
M1	7	0.909	0.462	0.819	0.628	-0.0430
O	10	0.789	0.708	0.706	0.806	-0.0155
O1	7	0.822	0.658	0.407	1.265	0.1134
MS	12	0.898	0.495	0.879	0.515	-0.0229
MS1	10	0.957	0.319	0.814	0.641	-0.0510
MO	11	0.909	0.468	0.882	0.507	-0.0327
MO1	2	0.912	0.449	0.820	0.627	-0.0418
SO	10	0.896	0.497	0.873	0.527	-0.0454
SO1	3	0.923	0.419	0.824	0.621	-0.0588

스펙트럼 베이스라인의 보정, 미세정보의 강조 등에 사용되는 1차미분이 각 전처리조합에 포함되는 경우(S와 S1, M과 M1, O와 O1 등) 오히려 모델의 예측성능이 크게 저하하는 것으로 나타났다. 이는 상대적으로 노이즈성분을 많이 포함하고 있는 실시간 스펙트럼의 경우 미분에 의해 스펙트럼의 노이즈 성분이 오히려 강조되어 주 성분분석에 있어 불리하게 작용되는 것으로 판단된다. 미분의 경우 팩터수를 크게 감소시키는 경향은 있으나 본 스펙트럼 데이터 셋의 경우에 있어서는 오히려 SEP를 크게 증가시켜 모델 예측성능을 저하시켰다.

Fig. 8은 전처리를 수행하지 않은 데이터를 이용하여 개발된 당도 예측 모델에 의한 예측값과 실제 측정된 당도값과의 관계를, 그리고 Fig. 9는 앞서 분석된 최적 전처리(MO)를 수행한 데이터를 이용하여 개발된 모델에 의한 예측값과 실측값과의 관계를 그림으로 표시한 것이다.

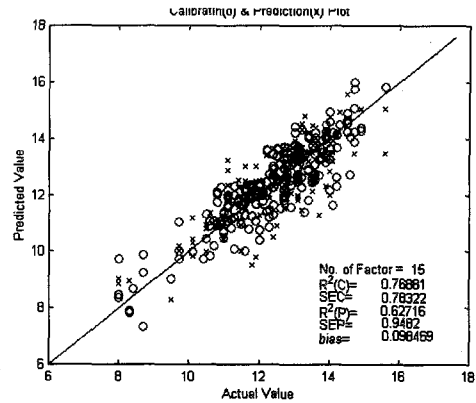


Fig. 8 Plot of calibration and validation results (no preprocess on spectral data)

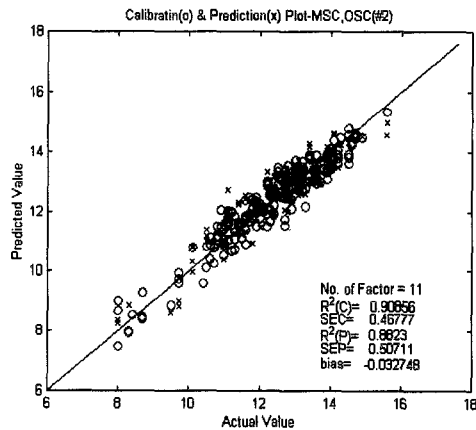


Fig. 9 Plot of calibration and validation results (MSC+OSC on spectral data)

4. 결론

본 연구는 초당 2개의 속도로 이송되는 사과를 대상으로 측정된 VIS/NIR영역의 투과 에너지 스펙트럼을 이용하여, 내부품질인자 중 당도 예측모델을 개발하기 위해 스펙트럼 전처리의 효과 및 최적 전처리 방법을 구명하고자 수행되었다. 결과를 요약하면 다음과 같다.

- 1) 스펙트럼의 산란보정, 노이즈감소 등을 위해 1차미분, MSC, SNV, OSC 및 이들의 조합으로 구성된 각종 전처리 알고리즘을 개발하고, 이를 투과 에너지 스펙트럼에 적용한 결과 특히 MSC와 SNV 전처리에 의해 스펙트럼 데이터의 변이가 크게 감소

하고, 나아가서 각 파장에서의 투과에너지와 당도와 의 상관관계가 전처리를 하지 않은 경우에 비해 현저히 증가하였다.

- 2) 각종 전처리를 수행한 후 PLSR을 이용하여 당도 예측 회귀 모델을 개발하고, 검증 실험을 수행한 결과, 전처리 방법에 따라 예측모델의 SEP가 최대 1.265 %brix에서 최소 0.507 %brix로 큰 차이를 나타내었다. 이는 SEP를 최소화 하기 위해 주어진 스펙트럼 데이터의 특성에 알맞는 전처리 방법이 개발 또는 선택되어야 함을 의미한다.
- 3) 본 연구에서 사용된 실시간 스펙트럼 데이터에 대한 1차미분 전처리는 오히려 모델 예측 성능을 저하시키는 것으로 나타났다. 이는 실시간으로 측정된 투과스펙트럼에 상대적으로 많이 포함되어있는 노이즈 성분이 미분에 의해 강조되기 때문인 것으로 판단되었다.

참고문헌

- [1] 김우기, "분광반사특성을 이용한 주요 과실의 비파괴 당산도측정", 서울대학교 대학원 석사학위논문, (1997)
- [2] 노상하, 김만수, 장세권, "수출용 배의 부가가치 향상을 위한 선별포장시스템 개발에 관한 연구", 농림부 농림기술개발사업 최종연구보고서, (2001)
- [3] 류동수, "VIS/NIR투과 분광분석법을 이용한 감귤의 비파괴 내부품질 판정시스템 개발", 서울대학교 박사학위논문, (2001)
- [4] 이강진, 노상하, W. R. Hruschka, J. A. Abbott, Yud-Ren Chen, "근적외선을 이용한 사과와 당도예측 모델 개발과 비교", 한국농업기계학회지, Vol. 22, No. 1, pp. 206-212, (1997)
- [5] 최창현, 이강진, 박보순, "가시광선/근적외선 분광분석법을 이용한 사과와 당도 및 경도측정", 한국농업기계학회지, Vol. 22, No. 1, pp. 200-205, (1997)
- [6] 황인근, "VIS/NIR 실시간 분광 스펙트럼에 의한 후지 사과의 당산도 선별 시스템 개발", 서울대학교 박사학위논문, (2000)
- [7] R. J. Barnes, M. S. Dhanoa and S. J. Lister, "Standard normal variate transformation and de-trending of near-infrared spectra," Analyst. Vol. 113, pp. 1849-1854, (1989)
- [8] P. Geladi, D. MacDougal and H. Martens, "Linearization and scattering-correction for non-linear reflectance spectra of meat", Applied Spectroscopy, Vol. 39, pp. 491-500, (1985)
- [9] T. Issakson, and B. Kowalsky, "Piece-wise multiplicative scatter correction applied to near-infrared diffuse transmittance data from meat products", Applied Spectroscopy, Vol. 46, pp.772-777, (1993)
- [10] PLSplus/IQ manual, Galactic Industries Corp., (1996)
- [11] A. Savitzky and M. J. E. Golay, "Smoothing and differentiation of data by simplified least square procedures", Analytical Chemistry. Vol. 36, pp. 1627- 1638, (1964)
- [12] J. Sjoblom, O. Svensson and M. Josefson, "An evaluation of orthogonal signal correction applied to calibration transfer of near infrared spectra", Chemometrics and Intelligence Laboratory System, Vol. 44, pp. 229-244, (1998)
- [13] S. Wold, H. Antti, F. Lindgren and J. Ohman, "Orthogonal signal correction of near-infrared spectra", Chemometrics and Intelligence Laboratory System, Vol. 44, pp. 175-185, (1998)