

투과 스펙트럼을 이용한 토마토 수확 후 저장일자 예측모형 개발

김영태 서상룡

Development of Prediction Model to Estimate the Storage Days of Tomato Using Transmittance Spectrum

Y. T. Kim S. R. Suh

Abstract

The goal of this study was to develop prediction models to estimate the storage days of tomato. The transmittance spectral data measured on tomato were preprocessed through normalization, SNV, Savitzky-Golay, and Norris Gap and then were used to build the prediction models using partial least square (PLS) method.

For the experiments, the tomato samples of different varieties were collected at different harvest time. The samples were taken right after harvest from the field and then were stored in a low-temperature storage room in which room temperature was maintained at 10°C. The transmittance spectral data of the tomato samples were measured at three-day intervals for 16 days.

The performance of the prediction models was affected by the preprocessing techniques as well as the varieties and harvest time of the tomato. The best model was found when SNV was applied. The accuracy of the best model was 90.2%. It can be concluded that the transmittance spectra are useful information for predicting the period of storage of tomato.

Keywords : Period of storage, PLS, Tomato, Transmittance spectrum

1. 서론

과채류의 신선도는 소비자에게 있어 내부품질과 마찬가지로 주요한 선택 요인 중 하나이다. 이러한 과채류의 신선도는 저장하는 동안 빠르게 변화한다. 보통 신선도는 표면 윤기나 경도의 정도로서 결정한다. 최근 많은 과채류 비파괴 품질 평가 방법들이 개발되고 있다(Jah와 Matsuoka, 2001). Clark 등(2004)은 300~1140 nm 영역의 VIS/NIR 스펙트럼을 이용하여 키위의 저장 기간 동안에 발생하는 손상을 예측하고자 하였다. 모형 개발을 위해서 키위의 건물함량, 당도, 그리고 과육의 색을 파괴적인 방법으로 측정하였으며, 이를 실제 측정 한 품질 변수와 NIR 스펙트럼 데이터를 CDA(canonical

discriminant analysis) 분석하여 모형을 개발하였다. 그 결과 수확당시 NIR 분석을 이용하여 과실을 분류함으로써 저장 기간 동안에 발생하는 손상을 줄일 수 있을 것이라고 보고하였다. Hsieh와 Lee(2005)는 400~2498 nm의 영역의 VIS/NIR 반사 스펙트럼을 측정하여 Pearl guava의 당도와 경도(생체 항복점)를 측정 하였다. 측정된 스펙트럼을 평활화(smoothing) 전처리를 실시하여 광 산란과 베이스 라인 이동의 영향을 제거 한 후 모형을 개발하였다. 모형 개발은 MLR(Multiple Linear Regression), PLS, MPLSR(Modified Partial Least Square Regression)의 분석을 통하여 개발하였다. 그 결과 MPLSR로 개발한 모형을 이용한 당도의 예측은 상관계수가 0.947이고, SEP는 0.721 Brix라고 보고하였다. 경도의 예측은 PLSR 분

This study was supported financially by the Agricultural R&D Promotion Center (ARPC). The article was submitted for publication on 2008-07-23, reviewed on 2008-08-30, and approved for publication by editorial board of KSAM on 2008-09-22. The authors are Young Tae Kim, KSAM member, Graduate Student, and Sang Ryong Suh, Professor, KSAM member, Chonnam National University. Corresponding author: Y. T. Kim, KSAM member, Graduate Student, Dept. of Biosystems Engineering, Chonnam National University, Gwangju, 500-757, Korea; Fax: +82-62-530-2159; E-mail: <expertsyskr@yahoo.co.kr>.

석을 통하여 개발한 모형에서 가장 낮은 SEP를 나타내는 것으로 분석되었으며, 3.427 N이라고 보고한 바 있다. Kim과 Suh 등(2008)은 분광분석 방법으로 토마토 경도를 예측하는데 효과적인 분광자료 수집방법과 예측모형 개발방법을 찾기 위해 반사광과 투과광 스펙트럼을 측정하고 수종의 스펙트럼 전처리 방법을 적용한 다음 PLS와 MLR 방법에 의해 경도 예측 모형을 개발하고 그 성능을 비교하였다. 그 결과, 투과광 스펙트럼을 이용하여 경도를 예측하는 것이 효과적인 것으로 나타났고, 투과광 분광자료를 PLS와 MLR에 의해 개발한 경도 예측모형의 성능을 비교한 결과 두 개발방법의 성능은 비슷한 것으로 보고하였다.

이러한 광학적 측정 방법 외에 Jah와 Matsuoka(2002)는 가지의 신선도 지표를 예측할 수 있는 공식을 개발하기 위해 표면 윤기, 경도 및 가지의 밀도의 물리적 매개변수들을 이용하였고 수분 손실과 밀도의 비율과 저장기간과 관련된 다양한 지수를 찾아내고 이들로부터 중량과 밀도 비율이 신선도 지표를 결정하는 최적의 관계를 나타내는 것으로 보고하였다. Cho와 Hwang 등(1998)에 따르면 사과와 토마토의 신선도 평가에 적합한 물리 화학적 인자를 탐색하고자 수확 후 경과한 시간에 따른 여러 가지 물리 화학적 인자들의 변화를 관찰하고, 인자들 상호간의 상관관계를 분석하였다. 연구결과 사과와 토마토는 신선도와 관련하여 상이한 물리 화학적 특성을 보인 것으로 보고하였다. Ketelaere 등(2003)은 13품종의 토마토를 2주 동안 저장하면서 상업용 음파 경도 측정기를 사용하여 경도 변화를 측정하고 선형 혼합모형을 개발하였다. 그 결과, 모형 개발에 있어서 가장 큰 변수는 품종이었고, 경도 측정 시 측정 오차가 모형 개발에 영향을 주었으며, 음파를 이용한 경도 측정기는 경도가 낮은 토마토에 있어서 그 재현성은 낮은 것으로 보고한 바 있다.

이상의 연구 결과를 기초로 하여 본 연구는 분광분석을 이용하는 비파괴적인 방법으로 토마토 수확 후 경과일자를 측정할 수 있는 예측모형을 개발하기 위한 연구로서, 토마토 수확 후 경과일자 별 투과광 스펙트럼을 측정한 다음 이들 스펙트럼을 수종의 전처리(preprocessing)를 적용한 후, PLS 분석하여 예측모형을 개발하고 그 성능을 검증하여 토마토 수확 후

경과일자를 예측할 수 있는 모형을 개발하고자 수행하였다.

2. 재료 및 방법

가. 공시 재료

본 연구에서 사용한 공시재료는 전라남도 장성군 황룡면에서 재배하고 2007년 4월, 5월 그리고 7월에 수확한 것으로 그 품종이 4월과 5월에 수집한 토마토는 ‘수퍼선로드’ 품종의 것이었고, 7월에 수집한 토마토는 ‘호용’ 품종이었으며 그 수는 각각 311개, 367개 그리고 360개였다.

나. 실험장치 및 방법

1) 투과 스펙트럼 측정 장치

토마토 수확 후 경과일자별 투과광 스펙트럼은 Kim과 Suh 등(2008)이 사용한 것과 동일한 그림 1과 같은 측정 장치를 사용하여 측정하였다.

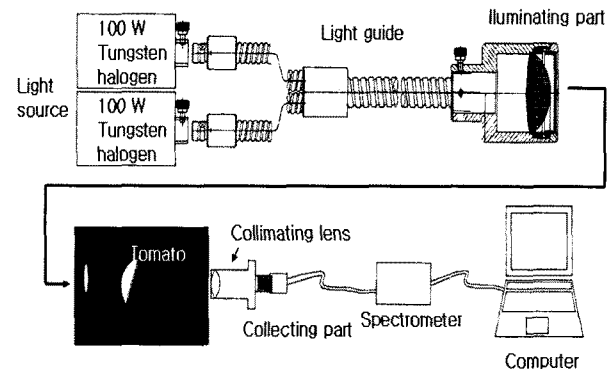


Fig. 1 Schematic diagram of experimental setup to measure transmittance spectrum from a tomato.

2) 실험 방법

수확 후 경과일자별 투과 스펙트럼은 본 연구에서 수집한 실험별 토마토를 표 1과 같이 각 실험별로 6개 그룹으로 구분한 다음, 10℃ 냉장고에 보관하면서 실험일로부터 16일(3일간격)까지 측정하였다.

Table 1 Number of samples used to develop a model to predict the storage days after harvest of tomato

Exp.	Cultivar	Number of samples						
		date 1	date 4	date 7	date 10	date 13	date 16	Total
1	Super-sunload	53	54	51	52	49	52	311
2	Super-sunload	73	72	64	53	61	44	367
3	Ho-yong	60	60	60	60	60	60	360
	Total	186	186	175	165	170	156	1038

본 연구에서는 투과 스펙트럼 측정 전 적외선 온도계(IR-thermometer, Fluke Co., U.S.A.)를 이용하여 토마토 표면 온도를 측정하였으며, 스펙트럼 측정 시 토마토 표면 온도가 상승할 경우, 이들 토마토를 일정수준의 온도로 유지하기 위하여 냉장고에 다시 보관 한 다음 온도를 재측정하여 온도가 10°C로 유지한 경우에만 측정하였다. 이는 스펙트럼 측정 시 토마토 온도의 영향을 배제하기 위함이었다. 또한 본 연구에서 사용한 분광분석기(USB4000, OceanOptics Inc., U.S.A.)에 의해 측정된 전체 파장(650~1100 nm)의 스펙트럼 3648개 중 2270개(659~1100 nm)를 사용하고 이들을 10개씩 평균하여 총 227개 스펙트럼으로 압축하여 사용하였다. 전체 스펙트럼 중 659 nm 파장 이하의 스펙트럼은 측정장치의 측정범위에 속하지만 정확한 스펙트럼의 측정이 이루어지지 않는 것으로 판단되어 사용하지 않았다.

3) 수확 후 경과일자 예측 모형 개발

일반적으로 VIS/NIR 스펙트럼은 외부환경 및 측정 장치에 의해 영향을 많이 받는 것으로 알려져 있다. 특히 시료의 크기 차이에 따른 경로 길이(path length)의 변화나, 산란(scattering) 차이 등에 의해 더 큰 변이를 일으킨다. 더군다나 스펙트럼 측정 시 주변 환경의 변화(조명, 온도 등)와 대상체 표면에서의 광 산란, 물체의 크기, 물체와 측정 센서와의 거리 차이, 측정 센서의 노이즈 등이 측정된 광 스펙트럼에 주요한 영향을 주게 된다. 이러한 광학 측정상의 노이즈를 제거하여 임의 물질의 광학 스펙트럼을 이용하는데 있어 오차를 줄이는 매우 기초적이고 중요한 기술로서 보다 안정적인 광학 스펙트럼의 특성을 구하는데 전처리가 사용된다. 이에 본 연구에서는 normalization(mean, maximum, range), SNV(standard normal variate), Savitzky-Golay & Norris Gap 1차 미분과 2차 미분 총 8종의 스펙트럼 전처리를 수행하였다. 이러한 전처리를 수행한 스펙트럼 자료로 저장일자를 예측하는 모형을 개발하기 위하여 분광분석에 많이 사용하는 통계분석 방법인 PLS(partial least square) 분석을 수행하였다. PLS분석은 The Unscrambler(Ver. 9.7, Camo Co., Norway)를 사용하였고, 수확 후 경과일자 예측 모형의 개발

은 다음과 같은 방법을 수행하였다. 첫째로 예측모형 개발 가능성을 판단하기 위해 실험별 8종 전처리와 전처리하지 않은 원 자료(raw data)를 PLS 분석하여 총 9종 예측 모형을 개발하고 그 개발 가능성을 파악한 다음, 개발한 예측모형을 다른 실험 토마토에 적용하여 예측정확도로써 성능을 검증하여 그 적용 가능성을 판단하였다. 둘째로 실험별 투과광 스펙트럼을 혼합한 다음 혼합한 자료를 모형 개발용 실험 자료와 검증용 실험 자료로 구분한 후 모형 개발용 실험 자료로서 예측 모형을 개발하고 개발 모형 검증용 실험 자료로서 그 성능을 검증하는 방법으로 예측모형을 개발하고 그 성능을 검증하였다. 표 2는 혼합 자료를 모형 개발 및 개발 예측모형의 검증을 위해 구분한 후 사용한 스펙트럼 자료의 수이다.

3. 결과 및 고찰

가. 수확 후 경과일자 예측모형 개발 가능성 실험 결과

각 실험별로 본 연구에서 적용한 8종 전처리를 수행한 자료와 전처리하지 않은 자료를 PLS 분석하여 총 9종 예측모형을 개발하고 개발모형의 결정계수와 검증용 자료의 예측정확도로 그 개발 가능성을 파악한 결과는 표 3과 같다. 표와 같이 Exp. 1의 경우, 개발한 9종 모형의 결정계수(R_c^2)는 0.62~0.83으로 나타났으며, 개발 모형의 예측정확도는 최소 87.7% (114/130)부터 최대 96.2%(125/130)로 나타났다. Exp. 2의 경우, 개발한 9종 모형의 결정계수는 0.65~0.76으로 나타났으며, 예측정확도는 최소 94.0%(140/149)부터 최대 97.3% (145/149)로 매우 높은 수준인 것으로 나타났다. Exp. 3의 경우, 개발한 9종 모형의 결정계수가 0.69~0.84로 나타났으며, 예측정확도는 최소 93.9%(169/180)부터 최대 97.2%(175/180)로 매우 우수한 것으로 나타났다.

이상의 결과로부터 투과 스펙트럼을 이용한 토마토 수확 후 경과일자 예측모형은 매우 높은 성능의 수준으로 개발이 가능할 것으로 판단되었다.

Table 2 Number of commingled spectrums used to develop a model to predict the storage days after harvest regardless of harvest time and cultivar of tomato

	Number of spectrums					
	date 1	date 4	date 7	date 10	date 13	date 16
Calibration	110	102	99	86	97	84
Validation	76	84	76	79	73	72

Table 3 Results of calibration and validation data obtained by the developed experiment

Preprocessing		Calibration		
		Exp. 1	Exp. 2	Exp. 3
		R_c^2	R_c^2	R_c^2
Raw data		0.83	0.75	0.70
Normalization	Mean	0.82	0.76	0.84
	Maximum	0.83	0.76	0.80
	Range	0.83	0.76	0.80
SNV		0.83	0.75	0.79
Savitzky-Golay	1st deri.	0.63	0.66	0.76
	2nd deri.	0.64	0.65	0.69
Norris Gap	1st deri.	0.62	0.65	0.72
	2nd deri.	0.66	0.65	0.76

Preprocessing		Validation		
		Exp. 1	Exp. 2	Exp. 3
Raw data		89.2% (116/130)	96.6% (144/149)	95.0% (171/180)
Normalization	Mean	94.6% (123/130)	96.6% (143/149)	97.8% (176/180)
	Maximum	93.8% (122/130)	96.6% (144/149)	93.9% (169/180)
	Range	96.2% (125/130)	96.6% (144/149)	95.0% (171/180)
SNV		96.2% (125/130)	97.3% (145/149)	91.7% (165/180)
Savitzky-Golay	1st deri.	87.7% (114/130)	96.6% (144/149)	94.4% (170/180)
	2nd deri.	87.7% (114/130)	96.6% (144/149)	97.2% (175/180)
Norris Gap	1st deri.	88.5% (115/130)	94.0% (140/149)	93.9% (169/180)
	2nd deri.	88.5% (115/130)	96.0% (143/149)	97.2% (175/180)

이상의 결과와 같이 토마토 수확 후 경과일자 예측모형의 개발이 가능할 것으로 판단되어 실험별 예측모형을 다른 실험자료로서 예측정확도를 파악하여 그 적용 가능성을 파악한 결과는 표 4와 같다. 표와 같이 Exp. 1의 실험자료로부터 개발한 예측모형을 수확시기는 다르나 품종이 같은 Exp. 2와

수확시기와 품종이 다른 Exp. 3의 실험자료의 예측 정확도를 파악한 결과, Exp. 2의 경우 예측 정확도가 49.0%(180/367) ~ 94.3%(346/367)로 나타났으며, 이 중 normalization-range 전처리하고 개발한 모형의 예측정확도가 전체 367개 중 346개를 정확하게 예측하여 94.3%(346/367)의 수준으로 매우 우

Table 4 Results of validation for experiment 2 and 3 by the prediction model (obtained using the data of experiment 1)

a) Experiment 2

Preprocessing		date 1	date 4	date 7	date 10	date 13	date 16	total
Raw data		47.9% (35/73)	55.6% (40/72)	51.6% (33/64)	52.8% (28/53)	34.4% (21/61)	52.3% (23/44)	49.0% (180/367)
Normalization	Mean	82.2% (60/73)	77.8% (56/72)	89.1% (57/64)	92.5% (49/53)	95.1% (58/61)	93.2% (41/44)	87.5% (321/367)
	Maximum	91.8% (67/73)	83.3% (60/72)	92.2% (59/64)	92.5% (49/53)	98.4% (60/61)	93.2% (41/44)	91.6% (336/367)
	Range	81.8% (67/73)	88.9% (64/72)	95.3% (61/64)	96.2% (51/53)	98.4% (60/61)	97.7% (43/44)	94.3% (346/367)
SNV		86.3% (63/73)	88.9% (64/72)	93.8% (60/64)	90.6% (48/53)	98.4% (60/61)	93.2% (41/44)	91.6% (336/367)
Savitzky-Golay	1st deri.	52.1% (38/73)	59.7% (43/72)	54.7% (35/64)	62.3% (33/53)	36.1% (22/61)	45.5% (20/44)	52.0% (191/367)
	2nd deri.	64.4% (47/73)	80.6% (58/72)	90.6% (58/64)	90.6% (48/53)	90.2% (55/61)	84.1% (37/44)	82.6% (303/367)
Norris Gap	1st deri.	53.4% (39/73)	58.3% (42/72)	53.1% (34/64)	64.2% (34/53)	44.3% (27/61)	61.4% (27/44)	55.3% (203/367)
	2nd deri.	63.0% (46/73)	65.3% (47/72)	62.5% (40/64)	79.2% (42/53)	63.9% (39/61)	77.3% (34/44)	67.6% (248/367)

Table 4 Results of validation for experiment 2 and 3 by the prediction model (obtained using the data of experiment 1) (Continued)

b) Experiment 3

Preprocessing		date 1	date 4	date 7	date 10	date 13	date 16	total
Raw data		18.3% (11/60)	61.7% (37/60)	60.0% (36/60)	73.3% (44/60)	33.3% (20/60)	11.7% (7/60)	43.1% (155/360)
Normalization	Mean	56.7% (34/60)	61.7% (37/60)	86.7% (52/60)	83.3% (50/60)	90.0% (54/60)	86.7% (52/60)	77.5% (279/360)
	Maximum	66.7% (40/60)	83.3% (50/60)	93.3% (56/60)	91.7% (55/60)	96.7% (58/60)	90.0% (54/60)	86.9% (313/360)
	Range	70.0% (42/60)	88.3% (53/60)	93.3% (56/60)	98.3% (59/60)	98.3% (59/60)	90.0% (54/60)	89.7% (323/360)
SNV		68.3% (41/60)	90.0% (54/60)	96.7% (58/60)	98.3% (59/60)	98.3% (59/60)	90.3% (54/60)	90.3% (325/360)
Savitzky-Golay	1st deri.	41.7% (25/60)	70.0% (42/60)	83.3% (50/60)	76.7% (46/60)	70.0% (42/60)	78.3% (47/60)	70.0% (252/360)
	2nd deri.	50.0% (30/60)	85.0% (51/60)	95.0% (57/60)	98.3% (59/60)	80.0% (48/60)	83.3% (50/60)	81.9% (295/360)
Norris Gap	1st deri.	38.3% (23/60)	66.7% (40/60)	80.0% (48/60)	75.0% (45/60)	63.3% (38/60)	63.3% (38/60)	64.4% (232/360)
	2nd deri.	41.7% (25/60)	71.7% (43/60)	86.7% (52/60)	86.7% (52/60)	70.0% (42/60)	76.7% (46/60)	72.2% (260/360)

수한 것으로 나타났다. 그리고 Exp. 3의 경우, 예측 정확도는 43.1%(155/360) ~ 90.3%(325/360)로 나타났으며 이 중 SNV 전처리하고 개발한 모형의 예측정확도가 90.3%(325/360)로 나타나 Exp. 1 실험자료로부터 개발한 예측 모형의 수확 후 경과일차 예측은 높은 수준으로 가능할 것으로 판단되었으나 수확시기와 품종에 관계없는 예측모형의 개발이 필요할 것으로 판단되었다.

Exp. 2의 실험자료로 부터 개발한 예측모형의 예측 정확도

를 품종은 같으나 수확시기가 다른 Exp. 1과 품종과 수확시기가 다른 Exp. 3의 실험자료로 그 정확도를 분석한 결과는 표 5와 같다. Exp. 1의 경우, 표와 같이 예측 정확도가 10.3%(32/311) ~ 95.8%(298/311)로 나타났으며, 이 중 Savitzky-Golay 2차 미분 전처리하고 개발한 모형의 전체 예측정확도가 가장 높은 95.8%(298/311)로 나타났다. 또한 Exp. 3의 경우, 예측 정확도는 71.1%(258/360) ~ 96.9%(349/360)로 나타났으며, 이 중 가장 높은 예측 정확도를 보인 예측 모형은 Exp. 1의

Table 5 Results of validation for experiment 1 and 3 by the prediction model (obtained using the data of experiment 2)

a) Experiment 1

Preprocessing		date 1	date 4	date 7	date 10	date 13	date 16	total
Raw data		98.1% (52/53)	50.0% (27/54)	39.2% (20/51)	36.5% (19/52)	44.9% (22/49)	32.7% (17/52)	50.5% (157/311)
Normalization	Mean	62.3% (33/53)	0.0% (0/54)	0.0% (0/51)	0.0% (0/52)	0.0% (0/49)	0.0% (0/52)	10.6% (33/311)
	Maximum	66.0% (35/53)	0.0% (0/54)	0.0% (0/51)	0.0% (0/52)	0.0% (0/49)	0.0% (0/52)	11.3% (35/311)
	Range	62.3% (33/53)	0.0% (0/54)	0.0% (0/51)	0.0% (0/52)	0.0% (0/49)	0.0% (0/52)	10.6% (33/311)
SNV		60.4% (32/53)	0.0% (0/54)	0.0% (0/51)	0.0% (0/52)	0.0% (0/49)	0.0% (0/52)	10.3% (32/311)
Savitzky-Golay	1st deri.	96.2% (51/53)	63.0% (34/54)	56.9% (29/51)	44.2% (23/52)	57.1% (28/49)	40.4% (21/52)	59.8% (186/311)
	2nd deri.	98.1% (52/53)	94.4% (51/54)	98.0% (50/51)	96.2% (50/52)	89.8% (44/49)	98.1% (51/52)	95.8% (298/311)
Norris Gap	1st deri.	96.2% (51/53)	63.0% (34/54)	56.9% (29/51)	48.1% (25/52)	59.2% (29/49)	42.3% (22/52)	61.1% (190/311)
	2nd deri.	96.2% (51/53)	70.43% (38/54)	86.3% (44/51)	63.5% (33/52)	67.3% (33/49)	59.62% (31/52)	74.0% (230/311)

Table 5 Results of validation for experiment 1 and 3 by the prediction model (obtained using the data of experiment 2) (Continued)

b) Experiment 3

Preprocessing		date 1	date 4	date 7	date 10	date 13	date 16	total
Raw data		51.7% (311/60)	71.7% (437/60)	88.3% (53/60)	98.3% (59/60)	78.3% (47/60)	38.3% (23/60)	71.1% (258/360)
Normalization	Mean	78.3% (47/60)	80.07% (48/60)	95.07% (57/60)	85.0% (51/60)	78.3% (47/60)	70.0% (42/60)	81.1% (292/360)
	Maximum	78.3% (47/60)	86.7% (52/60)	90.03% (54/60)	85.0% (51/60)	78.37% (47/60)	65.0% (39/60)	80.6% (290/360)
	Range	78.3% (47/60)	83.3% (50/60)	90.0% (54/60)	85.0% (51/60)	78.3% (47/60)	61.7% (37/60)	79.4% (286/360)
SNV		73.3% (44/60)	73.30% (44/60)	91.7% (55/60)	86.7% (52/60)	78.3% (47/60)	65.0% (39/60)	78.1% (281/360)
Savitzky-Golay	1st deri.	90.0% (54/60)	95.0% (57/60)	96.7% (58/60)	100.0% (60/60)	91.7% (55/60)	91.7% (55/60)	94.2% (339/360)
	2nd deri.	91.7% (55/60)	98.3% (59/60)	96.70% (58/60)	100.0% (60/60)	98.3% (59/60)	96.7% (58/60)	96.9% (349/360)
Norris Gap	1st deri.	90.0% (54/60)	93.3% (56/60)	96.70% (58/60)	100.0% (60/60)	86.7% (52/60)	90.0% (54/60)	92.8% (334/360)
	2nd deri.	75.0% (45/60)	80.0% (48/60)	88.37% (53/60)	91.7% (55/60)	71.7% (43/60)	80.0% (48/60)	81.1% (292/360)

결과와 같은 Savitzky-Golay 2차 미분 전처리하고 개발한 모형으로 그 예측 정확도는 96.9%(349/360)로 나타나 Exp. 1의 결과와는 달리 Exp. 2의 실험자료로부터 개발한 예측모형을 이용하여 수확시기와 품종에 관계없이 높은 수준의 정확도로 수확 후 경과일자의 예측이 가능한 것으로 판단되었다.

위의 Exp. 1과 Exp. 2에서 개발한 예측모형의 정확도 분석 실험과 동일하게 Exp. 3 실험자료로 부터 개발한 예측모형을 수확시기와 품종이 다른 Exp. 1과 Exp. 2에 적용하여 그 예측 정확도를 분석한 결과는 표 6과 같다. 표와 같이 Exp. 3으

로부터 개발한 예측모형을 Exp. 1과 Exp. 2의 실험자료에 대하여 예측한 결과, 그 정확도는 각각 3.2%(10/311) ~ 81.4%(253/311), 43.3%(159/367) ~ 66.5%(244/367)이었으며 이 중 Savitzky-Golay 2차 미분 전처리하고 개발한 모형의 전체 예측정확도가 Exp. 1의 경우 81.4%(253/311)이었으며, Exp. 2의 경우는 66.5%(244/367)로 나타났다. 이상의 결과로 볼 때, Exp. 3 실험자료로부터 개발한 예측모형의 경우 Exp. 1과 Exp. 2의 실험자료의 예측 정확도로서 그 성능을 파악하면, 수확시기와 품종에 관계없는 예측모형의 개발은 가능할 것으

Table 6 Results of validation for experiment 1 and 2 by the prediction model (obtained using the data of experiment 3)

a) Experiment 1

Preprocessing		date 1	date 4	date 7	date 10	date 13	date 16	total
Raw data		86.8% (46/53)	83.3% (45/54)	76.5% (39/51)	61.5% (32/52)	55.1% (27/49)	44.2% (23/52)	68.23% (212/311)
Normalization	Mean	18.9% (10/53)	0.0% (0/54)	0.0% (0/51)	0.0% (0/52)	0.0% (0/49)	0.0% (0/52)	3.2% (10/311)
	Maximum	54.7% (29/53)	0.0% (0/54)	0.0% (0/51)	0.0% (0/52)	0.0% (0/49)	0.0% (0/52)	9.3% (29/311)
	Range	58.5% (31/53)	0.0% (0/54)	0.0% (0/51)	0.0% (0/52)	0.0% (0/49)	0.0% (0/52)	10.01% (31/311)
SNV		71.73% (38/53)	37.07% (20/54)	43.1% (22/51)	38.5% (20/52)	14.3% (7/49)	3.8% (2/52)	35.0% (109/311)
Savitzky-Golay	1st deri.	81.1% (43/53)	51.9% (28/54)	37.3% (19/51)	34.6% (18/52)	36.7% (18/49)	28.8% (15/52)	45.3% (141/311)
	2nd deri.	90.6% (48/53)	88.9% (48/54)	88.2% (45/51)	73.1% (38/52)	73.5% (36/49)	73.1% (38/52)	81.4% (253/311)
Norris Gap	1st deri.	83.0% (44/53)	66.7% (36/54)	64.7% (33/51)	48.1% (25/52)	55.1% (27/49)	30.8% (16/52)	58.2% (181/311)
	2nd deri.	73.6% (39/53)	46.3% (25/54)	37.3% (19/51)	30.8% (16/52)	38.8% (19/49)	28.2% (15/52)	42.8% (133/311)

Table 6 Results of validation for experiment 1 and 2 by the prediction model (obtained using the data of experiment 3) (Continued)

b) Experiment 2

Preprocessing		date 1	date 4	date 7	date 10	date 13	date 16	total
Raw data		41.1% (30/73)	51.4% (37/72)	45.3% (29/64)	56.6% (30/53)	32.8% (20/61)	43.2% (19/44)	45.01% (165/367)
Normalization	Mean	37.3% (27/73)	44.4% (32/72)	35.90% (23/64)	69.8% (37/53)	42.6% (26/61)	47.7% (21/44)	45.2% (166/367)
	Maximum	30.1% (22/73)	52.8% (38/72)	57.8% (37/64)	66.0% (354/53)	41.0% (25/61)	43.2% (19/44)	48.0% (176/367)
	Range	41.1% (30/73)	69.4% (507/72)	70.3% (45/64)	71.7% (38/53)	45.9% (28/61)	34.1% (15/44)	56.1% (206/367)
SNV		30.1% (22/73)	54.2% (39/72)	75.0% (483/64)	67.93% (36/53)	54.1% (33/61)	52.3% (23/44)	54.8% (201/367)
Savitzky-Golay	1st deri.	37.0% (27/73)	54.2% (39/72)	43.8% (28/64)	52.8% (28/53)	31.1% (190/61)	40.9% (18/44)	43.3% (159/367)
	2nd deri.	76.7% (56/73)	77.8% (56/72)	64.1% (41/64)	73.6% (39/53)	42.6% (26/61)	59.1% (26/44)	66.5% (244/367)
Norris Gap	1st deri.	38.4% (28/73)	58.37% (427/72)	50.0% (32/64)	58.5% (31/53)	31.1% (19/61)	38.6% (17/44)	46.0% (169/367)
	2nd deri.	34.2% (25/73)	51.4% (37/72)	34.4% (22/64)	56.6% (304/53)	27.9% (17/61)	38.6% (17/44)	40.3% (148/367)

로 판단되나 그 성능은 만족할 만한 수준의 것은 아닌 것으로 판단되었다.

위와 같이 각 실험별로 개발한 예측모형을 다른 실험자료의 수확 후 경과일자를 예측한 결과, 개발한 예측모형을 수확 시기나 품종이 다른 토마토의 수확 후 경과일자 예측에 사용은 가능할 것으로 나타났다. 그러나 개발 예측모형의 성능은 만족할 만한 수준의 것은 아닌 것으로 나타났다. 이와같은 결과는 품종과 수확시기가 예측모형 개발에 큰 영향을 주는 것으로 판단되어 이러한 영향을 최소화 할 수 있는 예측모형의 개발이 필요할 것으로 판단되었다. 또한 예측모형 개발에 있어 단일의 전처리 방법을 제시할 수는 없는 것으로 판단되었다.

이에 본 연구에서는 각 실험별 투과광 스펙트럼을 혼합하고 혼합한 자료를 예측모형 개발용 자료와 개발 모형 검증용 자료로 구분한 후, 모형 개발용 자료를 이용하여 PLS분석하여 예측모형을 개발하고 그 성능을 검증용 자료로서 예측모형의 예측정확도를 파악하여 수확시기와 품종에 관계없이 수확 후 경과일자를 예측할 수 있는 예측모형을 개발하고자 하였으며 그 결과는 표 7과 같다. 표와 같이 혼합자료 중 예측모형 개발용 자료를 이용하여 개발한 전처리별 개발 모형 중 SNV 전처리하여 개발한 모형을 이용한 검증용 자료의 수확 후 경과일자 예측정확도가 90.2%(415/460)로 나타나 개발모형을 이용하여 수확시거나 품종에 관계없이 수확 후 경과일자를 예측하는데 그 적용이 가능할 것으로 판단되었다.

이상의 결과로부터 각 실험별로 개발한 토마토 수확 후 경과일자 예측모형은 수확시기와 품종이 큰 영향을 주는 것으로

로 판단되었으며, 단일의 전처리 방법을 제시할 수는 없는 것으로 나타났다. 그러나 이러한 영향을 최소화하기 위해서는 많은 수확시기와 품종의 실험자료를 수집하고 예측모형을 개발하여야 할 것으로 판단되었고, 최적의 예측모형의 개발을 위해 다양한 전처리를 적용하여야 개발하여야 할 것으로 판단되었다.

4. 결론

본 연구는 토마토 수확 후 경과일자별 투과광 스펙트럼을 측정하고 다양한 전처리를 적용한 다음, PLS분석하여 예측모형을 개발하고 그 성능을 예측정확도로 파악하여 수확 후 경과일자를 비파괴적인 방법으로 예측할 수 있는 모형을 개발하기 위해 각 실험별로 개발한 토마토 수확 후 경과일자 예측모형을 개발하고 수확시기와 품종이 다른 실험 자료를 이용하여 그 성능을 평가하였으며, 수확시기와 품종이 다른 실험 자료를 혼합하여 예측모형을 개발하여 그 성능을 평가하고자 하였으며 그 결과는 다음과 같다.

각 실험별로 개발한 토마토 수확 후 경과일자 예측모형은 수확시기와 품종이 큰 영향을 주는 것으로 판단되었고 단일의 전처리 방법을 제시할 수는 없는 것으로 나타났다. 그러나 이러한 영향을 최소화하기 위해서는 많은 수확시기와 품종의 실험자료를 수집한 다음 예측모형을 개발하면 그 예측 성능은 향상될 것으로 판단되었다. 또한 최적의 예측모형의 개발을 위해 다양한 전처리를 적용하여야 개발하여야 할 것으로

Table 7 Validation results for the models to predict freshness of tomato regardless of harvest time and cultivar by the developed PLS model

Preprocessing		date 1	date 4	date 7	date 10	date 13	date 16	total
Raw data		82.9% (63/76)	83.3% (70/84)	82.9% (63/76)	67.1% (53/79)	67.1% (49/73)	62.5% (45/72)	74.6% (343/460)
Normalization	Mean	93.4% (71/76)	90.5% (76/84)	94.71% (72/76)	88.6% (70/79)	87.7% (64/73)	81.9% (59/72)	89.6% (412/460)
	Maximum	93.4% (71/76)	88.1% (74/84)	89.5% (68/76)	88.6% (70/79)	83.6% (61/73)	79.2% (57/72)	87.2% (401/460)
	Range	93.4% (71/76)	88.1% (74/84)	94.7% (72/76)	87.3% (69/79)	87.7% (64/73)	81.9% (59/72)	88.9% (409/460)
SNV		93.4% (71/76)	90.5% (76/84)	94.7% (72/76)	89.9% (71/79)	89.0% (65/73)	83.3% (60/72)	90.2% (415/460)
Savitzky-Golay	1st deri.	82.9% (63/76)	81.0% (68/84)	84.2% (64/76)	74.7% (59/79)	69.9% (51/73)	69.4% (50/72)	77.21% (355/460)
	2nd deri	85.5% (65/76)	90.5% (76/84)	93.4% (71/76)	79.7% (63/79)	76.7% (56/73)	83.3% (60/72)	85.0% (391/460)
Norris Gap	1st deri	82.9% (63/76)	82.1% (69/84)	84.2% (64/76)	72.2% (57/79)	68.5% (50/73)	69.4% (50/72)	76.7% (353/460)
	2nd deri	84.2% (64/76)	76.2% (64/84)	85.5% (65/76)	75.9% (60/79)	69.9% (51/73)	69.4% (50/72)	77.0% (354/460)

판단되었다.

이상의 결과로 볼 때, 토마토의 투과광 스펙트럼을 이용한 비파괴적인 방법으로 수확 후 경과일자를 예측할 수 있는 모형의 개발은 가능할 것으로 판단된다. 개발 예측모형 중 SNV 전처리를 수행한 다음 개발한 예측모형의 전체 예측정확도가 90.2%인 점을 감안하면 이를 이용할 경우 비교적 높은 수준으로 토마토 수확 후 경과일자를 예측할 수 있을 것으로 판단되었다. 그러나 본 연구는 실험실 실험에 의한 결과이므로 실제 시장에서의 적용을 위한 보완이 필요할 것으로 판단되었다.

참고 문헌

1. Clark, C. J., V. A. McGlone, H. N. De Silva, M. A. Manning, J. Burdon and A. D. Mowat. 2004. Prediction of storage disorders of kiwifruit (*Actinidia chinensis*) based on visible-NIR spectral characteristics at harvest. *Postharvest Biology and Technology* 32:147-158.
2. Cho, Y. J. and J. K. Hwang. 1998. Physiochemical factors

for evaluating freshness of apple and tomato. *Journal of KSAM* 23(5):473-480. (In Korean)

3. Hsieh, C. and Y. Lee. 2005. Applied visible/near-infrared spectroscopy on detecting the sugar content and hardness of pearl guava. *Applied Engineering in Agriculture* 21(6):1039-1046.
4. Jha, S. N. and T. Matsuoka. 2001. Determination of freshness of eggplant using some physical parameters. *ASAE Paper No.* 016096.
5. Jha, S. N. and T. Matsuoka. 2002. Development of Freshness Index of Eggplant. *Applied Engineering in Agriculture* 18(5): 555-558.
6. Ketelaere Bart De, J. Lammertyn, G. Molenberghs, B. Nicolai and J. De Baerdemaeker. 2003. Statistical modeling of tomato firmness evolution during postharvest storage. *ASABE Paper No.* 036136.
7. Kim, Y. T. and S. R. Suh. 2008. Comparison of performance of models to predict hardness of tomato using spectroscopic data of reflectance and transmittance. *Journal of Biosystems Engineering* 33(1):63-68. (In Korean)