

가시광선/근적외선을 이용한 쌀의 정백수율 예측

Prediction of Rice Milling Ratio by Visible/ Near-Infrared Spectroscopy

김재민 [*]	민봉기 ^{**}	최창현 [*]
정희원	정희원	정희원
J.M.Kim	B.K.Min	C.H.Choi

1. 서론

전자산업과 컴퓨터 산업의 발전으로 농업 분야에서도 첨단 측정 장비의 도입과 새로운 기계 및 제어장치의 개발로 농업 생산 시스템의 자동화에 관한 연구는 물론 농산물의 광학적 특성을 측정하는 비파괴 검사에 대한 연구가 활발히 진행되고 있다. 측정 재료에 광선을 照射하면 일부는 재료의 표면에서 반사되고, 나머지는 흡수되거나 투과하게 되며, 이러한 반사, 흡수, 투과되는 에너지의 양은 재료의 표면상태와 내부상태에 따라 변화하게 된다. 이러한 재료의 고유 특성에 따른 파장대역별 스펙트럼은 통계적 처리에 의하여 정성 및 정량화할 수 있으며 농산물의 품질 또는 이화학적 특성의 측정 및 선별에 사용된다.

곡물의 가공, 저장 및 유통과정에서의 손실을 방지하기 위한 연구는 곡물의 생산 못지 않게 중요하게 취급되고 있다. 특히 주식인 쌀의 경우, 도정과정에서 곡물의 양적, 질적 손실을 절감할 수 있는 도정과정의 자동화가 절실하게 요구되고 있다. 이를 위해서는 정백수율에 대한 신속하고 객관적인 측정 방법의 개발이 필요하다. 국내에서의 정백수율 판정은 대부분 경험에 의한 주관적 판단에 의존하고 있으나, 일본의 경우, 일본정미공업회(재)에서는 반드시 백도계(whiteness tester)와 수분계등을 사용하여 품질 측정을 하도록 규정하고 있으며 우리나라에서도 농협이나 중소형 도정공장에서 품질 측정을 위하여 백도계의 사용이 활발하다. 백도계는 가시광선 영역에서의 특정 파장에 의한 도정된 쌀 표면의 반사율을 측정하여 백도로 나타내며 이를 정백수율로 환산할 수 있다. 그러나 단순히 쌀 표면의 색상을 측정하여 나타내므로 정백수율을 표현하는 참고 자료로는 활용할 수 있으나, 동일 정백수율에서도 산지 및 품종에 따른 백도의 차이가 다소 있으므로 단순히 백도를 이용한 정백수율의 측정은 적합하지 않은 것으로 판단된다. 미국은 일부 대형 도정공장에서는 근적외 분광 분석기를 이용하여 지방함량을 측정함으로써 정백수율을 측정하고 있다.

* 성균관대학교 생물기전공학과

** 한국식품개발연구원 쌀연구사업단

본 연구는 가시광선/근적외선 영역(400 - 2500nm)을 이용하여 쌀의 정백수율을 판정함에 있어서 사람의 시각과 백도계가 사용하고 있는 가시광선은 물론 쌀 표면의 이화학적 특성 측정이 가능한 근적외선 영역을 사용하여 반사 및 투과특성을 분석하므로써 비파괴적 방법으로 현장에서 신속하게 정백수율을 예측할 수 있는 방법을 개발하기 위하여 수행하였다.

2. 재료 및 방법

가. 재료

공시재료는 농촌진흥청 산하의 도진흥원에서 재배한 1995년산 원원종의 벼로서 1996년 2월에 구입하였다. 각 지역별로 선택된 품종은 동진(충남, 경북, 경남, 전북), 화성(강원, 전남, 경북, 경기), 일품(경북, 충남, 충북, 경기)등으로서 모두 12개의 재료를 이용하였다.

시료의 장기 보관을 위하여 진공포장용 비닐팩으로 밀봉한 후 4℃의 저온저장고에 보관하였으며, 벼의 제현시에는 비닐팩에 넣은 상태로 상온에서 24시간 동안 노출시켜 곡물의 온도를 상온에 도달하도록 하였다.

나. 도정 수율별 시료 제조

벼는 제현기(Satake, THU 35A, Japan)를 사용하여 왕겨를 제거한 후 1.7mm 줄체와 썰미 정선기(Satake, TRCO5A, Japan)를 사용하여 미숙립과 썰미를 제거하므로써 정립의 현미를 선별하였다.

선별된 현미는 시험용 도정기(Toyo Tester Miller, MC-90A, Japan)를 이용하여 정백수율 94% - 86%까지 0.5% - 1%내외의 간격으로 도정을 실시한 후 무게를 측정하였다. 정백수율은 도정전 현미의 중량과 도정후 백미의 중량의 비로서 나타낼 수 있으므로 식 (1)과 같이 계산하였다.

$$\text{정백수율(Milling ratio)} = \frac{\text{도정 후 백미의 중량(g)}}{\text{현미의 중량(g)}} \times 100 \dots \dots \dots (1)$$

다. 근적외선 분광광도계

백미의 정백수율별 흡광 스펙트럼의 측정은 근적외선 분광광도계(NIRSystems, Model 6500, Silver Spring, U.S.A.)를 이용하였다. 스펙트럼은 각 품종에 대하여 10~20개의 반사 및 투과 스펙트럼을 측정하므로써 총 298개(반사 및 투과 스펙트럼 각 149개)를 측정하였다.

본 연구에 이용된 근적외선 분광광도계는 반사도 및 투과도를 측정할 수 있으며, 시료의 종류에 따라 각각 다른 모듈을 설치하여 측정할 수 있다. 본 연구에서는 백미의 정백수을 별 반사 및 투과 스펙트럼을 측정하기 위하여 시료이송장치(Sample Transport Module, NIRSystems, NR-6511, USA)를 근적외선 분광광도계에 부착하여 측정하였다. 시료이송장치는 넓은 면적의 스펙트럼을 측정할 수 있으므로 시료채취시의 오차를 줄이고 다량의 시료로 대표성이 향상되며 시료의 전처리에서 발생하는 오차를 줄일 수 있다.

라. 정백수을 예측모델 개발 방법

스펙트럼을 이용한 정백수을 예측 모델을 개발하기 위하여 149개의 스펙트럼은 교정부 92개와 검증부 57개로 분류하였으며, 교정부의 스펙트럼은 모델의 개발 단계에서 반복하여 사용되었고 검증부는 예측 모델의 최종평가에 사용되었다. 정백수을 예측 모델은 다중회귀 분석법(Multiple Linear Regression; MLR), 부분최소자승법(Partial Least Square method; PLS), 신경회로망기법(Artificial Neural Network; ANN)을 이용하여 개발하였다. MLR과 ANN 모델 개발은 NSAS 4.0(NIRSystems, Silver Spring, U.S.A.)을 사용하였으며, PLS 모델은 상용 프로그램인 GRAMS/32(Galactic Industries, Salem, USA)를 사용하였다. 각 모델은 교정부의 표준오차(Standard Error of Calibration, SEC), 검증부의 표준오차 SEP(Standard Error of Prediction), Bias를 이용하여 모델의 예측 성능을 분석하였다.

각 모델의 개발은 원시 스펙트럼과 gap 및 평활화 간격을 변화시킨 1차($d\log(1/R)$) 및 2차 미분($d^2\log(1/R)$) 스펙트럼을 사용하였다. 이들 스펙트럼은 빛의 산란효과를 제거하기 위하여 전처리로서 다분산보정(Multiplicative Scatter Correction; MSC)과 표준정규화(Standard Normal Variate and Detrending; SNV)를 적용시킨 스펙트럼과 전처리를 실시하지 않은 스펙트럼으로 분리하여 분석하였다.

마. 다중회귀분석을 이용한 정백수을 예측모델(다중회귀 모델)의 결정조건

다중회귀 모델은 step-up, stepwise 두가지 방법으로 모델을 개발하였으며, 최대 6개의 파장을 선택할 수 있도록 설정하였다. 이때, 최적의 모델은 실측치와 예측치 사이의 결정계수(r^2)가 높고, 교정부의 표준오차가 낮으며, 최소의 파장을 가진 모델로 결정하였다.

바. PLS를 이용한 정백수을 예측모델의 결정조건

PLS 모델의 적정 팩터 수는 PRESS (Prediction Residual Error Sum of Squares)의 수치를 이용하여 결정하였다. PRESS 최소값과 비교하여 F-test의 결과가 0.75미만의 PRESS값을 나타내는 팩터의 수를 최적팩터로 결정하였다. 개발된 모델은 검증부의 예측에 적용하여 예측치의 표준오차(SEP)를 이용하여 분석하였다.

사. 신경회로망기법을 이용한 정백수율 예측모델(ANN모델)의 결정조건

가시광선/근적외선 스펙트럼을 이용한 정량 및 정성분석은 비선형 문제로 해석할 수 있으며, 이러한 비선형문제는 신경회로망기법을 이용하므로써 보다 우수한 모델을 개발할 수 있다. 따라서 본 연구에서는 하나의 은닉층(hidden layer)을 두었고, 입력층(input layer)은 흡광스펙트럼, 출력층(output layer)은 정백수율로 결정하였다. 스펙트럼의 학습시 입력층과 출력층의 전달함수는 시그모이드 함수를 사용하였으며, 시그모이드 함수의 전달율은 0.25, 학습율(learning rate)은 0.2, 모멘텀(momentum)은 0.3으로 하였다. 또한, ANN 모델의 개발은 PLS 모델을 기본으로 하여 파장영역, 수처리방법, 산란보정 방법등을 결정하였다.

3. 결과 및 고찰

가. 백미의 정백수율별 스펙트럼

본 연구에 이용된 시료는 현미의 정백수율을 100%로 하였을 때, 백미의 정백수율은 86.0~94.9%로서 약 0.5%간격으로 도정을 실시한 후 반사 및 투과 스펙트럼을 측정하였다. 반사 및 투과 스펙트럼은 가시광선/근적외선 영역이 포함된 400~2500nm 영역에서 2nm간격으로 측정하였다. 그러나 투과 스펙트럼의 경우 600~1400nm 이외의 파장에서는 시스템 감지부의 특성으로 나타나는 잡음이 많아 분석시 사용하지 않았다.

정백수율별 원시 스펙트럼을 분석한 결과 가시광선영역에서는 현미의 흡광도($\log(1/R)$)가 가장 높게 나타났으며, 정백수율이 감소할수록 흡광도는 감소하였다. 근적외선 영역에서는 현미의 흡광도가 가장 낮았으며, 백미의 경우에는 정백수율에 따라 유사한 흡광도를 나타내었다.

나. 다중회귀모델을 이용한 정백수율 예측

반사 스펙트럼의 경우 Stepwise방법으로 전파장 영역에서 전처리를 하지않고 평활화 간격 4nm, gap 8nm로 하여 1차 미분한 후 6개의 파장을 선택하였을 때 가장 적합한 모델을 구할 수 있었다. 사용된 파장은 404, 1184, 1216, 1732, 1734, 1184, 1846nm였으며, 이 결과를 이용하여 가시광선 영역을 제외한 1100~2500nm의 근적외선 영역에서 모델을 구하였지만 만족할 만한 결과를 얻지 못하였다.

본 연구에서 구한 다중회귀 모델에서 교정부의 r^2 는 0.984, SEC는 0.416%로 나타났으며, 57개의 미지시료에 대하여 검증한 결과 r^2 는 0.980, SEC는 0.286%, Bias는 0.069로 나타나 이 모델은 우수한 모델임을 알 수 있었다.

600~1400nm영역에서 스펙트럼을 측정한 투과 스펙트럼의 경우에도 반사 스펙트럼의 경우와 유사한 결과를 얻을 수 있었다. 산란보정을 실시하지 않고 평활화 간격 4nm, gap 8nm로 하여 1차 미분한 후 6개의 파장을 선택한 모델이 가장 우수하였으며, 교정부의 r^2 는 0.973, SEC는 0.549%로 나타났다. 이때, 사용된 파장은 610, 664, 684, 692,

968, 1228nm였다. 투과스펙트럼을 이용하여 57개의 미지시료에 대하여 검증한 결과 r^2 는 0.963, SEC는 0.524%, Bias는 -0.078로 나타났다.

위의 결과로부터 정백수율 예측을 위한 다중회귀모델은 반사 스펙트럼을 이용하여 전파장 영역에서 전처리를 하지 않고 1차 미분하여 6개의 파장을 선택한 모델이 정백수율을 가장 잘 예측할 수 있을 것으로 판단되었다.

다. PLS 모델을 이용한 정백수율 예측

400~2500nm의 전파장 영역에서 PLS를 실시한 결과, 산란 보정 및 전처리 방법에 따라 약간의 차이는 있었지만 정백수율과 반사 스펙트럼은 높은 상관관계를 보여 주었다.

반사 스펙트럼의 경우, 원시 스펙트럼을 전파장 영역에서 다분산보정을 하여 적정 팩터를 4로 결정하였을 때 가장 양호한 모델을 개발할 수 있었다. 이 모델에서 교정부의 r^2 는 0.970, SEC는 0.564%로 나타났으며, 57개의 미지시료에 대하여 검증한 결과 r^2 는 0.976, SEC는 0.365%, Bias는 -0.145로 나타났다.

투과 스펙트럼에서는 원시 스펙트럼을 전파장 영역에서 전처리를 하지 않고 적정 팩터를 7로 결정하였을 때, 가장 양호한 모델을 작성할 수 있었다. 이 모델에서 교정부의 r^2 는 0.961, SEC는 0.651%로 나타났으며, 57개의 미지시료에 대하여 검증한 결과 r^2 는 0.964, SEC는 0.518%, Bias는 0.005로 나타났다.

위의 결과에서 알 수 있듯이 PLS 모델은 교정부보다는 검증부에서 오히려 오차가 줄어들었으며, 다중회귀 모델과 같이 반사 스펙트럼을 이용하는 것이 우수한 모델을 개발할 수 있을 것으로 판단되었다.

라. ANN 모델을 이용한 정백수율 예측

ANN 모델은 PLS 분석결과를 기초로하여 개발하였다. 원시 스펙트럼과 1차 및 2차 미분 스펙트럼을 이용하여 전처리를 하였을 때와 하지 않은 경우 각각에 대하여 모델을 개발하여 비교하였으며, 각 모델의 교정부의 결정계수와 SEC값을 검토하여 적정 모델을 선정하였다.

반사 스펙트럼의 경우, 원시 스펙트럼을 전파장 영역에서 전처리를 하지 않고 gap 4nm로 1차 미분한 스펙트럼에서 적정 팩터를 5로 결정한 PLS 모델을 기초로 하여 ANN 모델을 작성한 것이 가장 우수한 결과를 나타내었다. 이 모델에서 교정부의 r^2 는 0.997, SEC는 0.143%로 나타났으며, 57개의 미지시료에 대하여 검증한 결과 r^2 는 0.978, SEC는 0.320%, Bias는 0.031로 나타났다.

투과 스펙트럼에서는 전파장 영역에서 전처리를 하지 않은 원시스펙트럼에서 적정 팩터를 7로 결정한 PLS 모델을 기초로 하여 개발된 ANN 모델이 가장 양호하였다. 이 모델에서 교정부의 r^2 는 0.978, SEC는 0.439%로 나타났으며, 57개의 미지시료에 대하여 검증한 결과 r^2 는 0.968, SEC는 0.451%, Bias는 0.024로 나타났다.

따라서 ANN 모델은 다중회귀 모델 및 PLS 모델과 같이 반사 스펙트럼을 이용하는 것이 우수한 모델을 개발할 수 있었다.

4. 결론 및 요약

본 연구는 가시광선/근적외선 분광분석법을 이용하여 쌀의 정백수율을 예측할 수 있는 적정 모델을 개발하기 위하여 수행하였다. 공시재료는 농촌진흥청 산하의 도진흥원에서 재배하여 저장중인 1995년산 원원종의 벼로서 재배지역별로 구별하여 3개 품종을 각 4개 도별로 선택하여 1996년 2월에 구입하였다. 각 시료에 대하여 400~2500nm영역에서 반사 및 투과스펙트럼을 측정하였으며, 특히 투과 스펙트럼은 분석시 600~1400nm영역만을 이용하였다. 정백수율 예측 모델은 반사 및 투과 스펙트럼에 대하여 전처리를 하지 않은 원시 스펙트럼, 1차 및 2차 미분 스펙트럼과 각 스펙트럼을 전처리한 스펙트럼을 이용하였다. 예측 모델은 반사 및 투과 스펙트럼에 대하여 다중회귀분석법(Multiple Linear Regression; MLR), 부분최소자승법(Partial Least Square method; PLS), 신경회로망기법(Artificial Neural Network; ANN)을 이용하여 개발하였다.

정백수율 예측을 위한 다중회귀 모델은 투과 스펙트럼을 이용하기보다는 반사 스펙트럼을 이용하여 전파장 영역에서 전처리를 않은 1차 미분 스펙트럼에 대하여 6개의 파장을 선택하여 모델을 개발하는 것이 정백수율을 가장 잘 예측할 수 있을 것으로 판단되었다. PLS 모델은 교정부보다는 검증부에서 오히려 오차가 줄어들었으며, 다중회귀 모델과 같이 반사 스펙트럼을 이용하는 것이 우수한 모델을 개발할 수 있을 것으로 판단되었다. ANN 모델은 PLS 분석결과를 기초로하여 개발하였다. ANN 모델은 다중회귀 모델 및 PLS 모델과 같이 반사 스펙트럼을 이용하는 것이 우수한 모델을 개발할 수 있었다.

위의 결과로부터, 본 연구에서 개발된 세가지 모델 중에서 다중회귀 모델이 정백수율 예측을 위한 가장 우수한 모델로 나타났지만, PLS 모델 및 ANN 모델 또한 정백수율 예측을 위한 시스템 개발에 적용시킬 수 있을 것으로 판단되었다. 다중회귀 모델은 6개의 특정 파장을 사용하므로 광다이오드와 대역필터를 이용한 시스템에 적용시킬 수 있으며, PLS 모델과 ANN 모델은 400~2500nm의 파장 영역에서 스펙트럼을 측정할 수 있는 회절격자형(grating type) 시스템에 적용시킬 수 있을 것으로 판단되었다.

5. 참고문헌

1. 김영배, 하덕모, 김창식. 1990. 우리나라 쌀의 도정 품위특성. 식품과학. 22(2):199-205.
2. 노상하, 也田善郎, 山下律也. 1990. 벼의 분광반사특성과 이를 이용한 정백미의 백도 측정기 개발. 한국농업기계학회지. 15(1):33-43.
3. Kao, C. 1986. Determination of Rice Milling Degree with Near Infrared Reflectance Method, Proceedings, 21st Rice Technical Working Group Feb. 24-26. 101-116.

4. Marshall, W.E and J.I. Wadsworth. 1993. Rice Science and Technology. Marcel Dekker, Inc. New York.
5. Siebenmorgen, T.J and H. Sun. 1994. Relationship Between Rice Surface Fat Concentration and Degree of Milling as Measured with a Commercial Milling Meter. American Association of Cereal Chem. 71(4) : 327-329.

Table 1 The result of calibration and prediction milling ratio using each model

Method	Math. Treatment	Preprocessing	Calibration (n=92)		Validation (n=57)		
			r ²	SEC	r ²	SEP	Bias
MLR	dlog(1/R)	None	0.984	0.416	0.980	0.286	0.069
PLS	log(1/R)	MSC	0.970	0.564	0.976	0.365	-0.145
ANN	dlog(1/R)	None	0.997	0.143	0.978	0.320	0.031

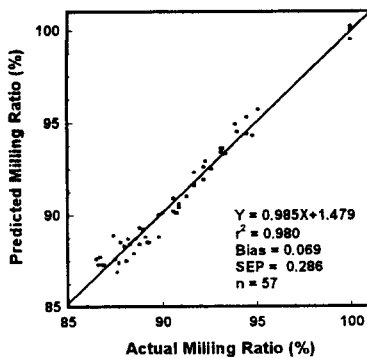


Fig. 1 Actual Vs Predicted Milling Ratio by Reflectance. (Multi-Linear Regression ; Stepwise)

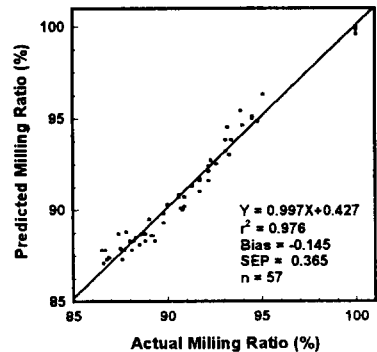


Fig. 2 Actual Vs Predicted Milling Ratio by Reflectance. (Partial Least Square Regression)

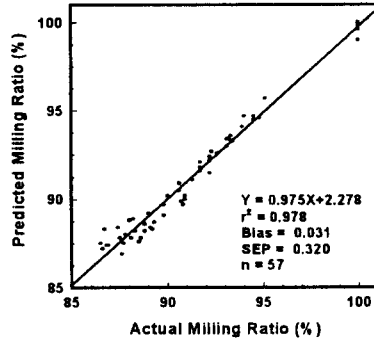


Fig. 3 Actual Vs Predicted Milling Ratio by Reflectance (Artificial Neural Network)