Nondestructive maturity determination in green tomatoes using a handheld visible and near infrared instrument

Gopal Tiwari\textsuperscript{a,}\textsuperscript{*}, David C. Slaughter\textsuperscript{a}, Marita Cantwell\textsuperscript{b}

\textsuperscript{a} Department of Biological and Agricultural Engineering, University of California, Davis, One Shields Avenue, Davis, CA 95616, USA

\textsuperscript{b} Department of Plant Sciences, University of California, Davis, One Shields Avenue, Davis, CA 95616, USA

A R T I C L E   I N F O

Article history:
Received 19 June 2012
Accepted 4 July 2013

Keywords:
Nondestructive sensing
Spectroscopy
Handheld spectrometer
Green tomatoes
Maturity

A B S T R A C T

A nondestructive method based on visible and near infrared spectroscopy, was investigated for determining the maturity of green tomatoes at harvest. The interactance spectrum of tomato fruit at the firm green stage were measured in less than 5 s with a handheld spectrometer (400–1000 nm) utilizing an optical design with a 0° angle between the incident illumination and the detection view. Results showed that the predominant change in the interactance spectra in green tomatoes useful for predicting changing maturity levels occurred in the 600–750 nm portion of the 400–1000 nm region, typically associated with chlorophyll. Variety-specific Bayesian classification models and a joint variety ‘global’ classification model were developed to predict tomato maturity after 7 d of storage in a 20°C ripening room using canonical discriminant analysis techniques applied to the interactance spectra from 600 nm to 750 nm. Variety specific models correctly identified 75–85% of immature tomatoes and 82–86% of mature green tomatoes in internal cross-validation, however external validation performance decreased when applied to predict maturities in a cultivar external to the training set. False positive rates of these models in identifying immature and mature green tomatoes varied from 3% to 40% and 0% to 31% respectively. A ‘global’ model, trained on two cultivars showed more stability and correctly predicted 71% of immature and 85% of mature green tomatoes, with false positive error rates of 13% and 22%, respectively, in internal cross-validation of both varieties. This handheld system showed good potential as a rapid, nondestructive technique to aid tomato production managers in the identification of immature green tomatoes at harvest and could be a valuable tool in delivering more flavorful fruit to consumers by reducing the amount of immature fruit harvested by workers.

© 2013 Published by Elsevier B.V.

1. Introduction

To minimize transit and other supply-chain related injury, tomatoes are commercially harvested when they are mature, but still firm and green in color (Kader et al., 1977; Wills and Ku, 2002; Alonso et al., 2010). After transport, green tomatoes are allowed to ripen at retail distribution centers so that they can develop uniform red peel color, and characteristic tomato flavor, and aroma before consumption. Ripe tomato quality attributes, such as color, flavor, and texture are greatly influenced by maturity at harvest (Kashmere and Kader, 1978; Kader et al., 1978a). Maturity in green tomatoes before ripening is characterized by the development of gel in their locules (Brecht et al., 1991; Atta-Aly et al., 2000). As the fruit matures, cell walls of locular tissues of green tomatoes undergo autolysis to form this gel. Although the USDA (United States Department of Agriculture) classifies tomatoes into six different groups based on the development of red color on their skin during ripening, no official method is currently used to distinguish the maturities of green tomatoes at the time of harvest. Kashmere and Kader (1978) developed a classification system for green tomatoes that is comprised of four maturity groups: immature, partially mature, fully mature, and advanced mature, however it requires cutting the fruit. They noted that immature tomatoes completely lack gel in their locules and their seeds can be cut by a sharp knife, while slicing. Partially mature tomatoes can have small amounts of gel in some of their locules, but their seeds are well developed. Fully mature and advanced mature tomatoes have gel in all their locules, and their seeds are not typically cut when slicing the fruit. Advanced mature green tomatoes also develop internal red color.

Immature green tomatoes will ripen and turn red if exposed to ethylene gas, but they will not develop good flavor (Kader et al., 1977, 1978a,b; Maul et al., 1998; Qin et al., 2012). The immature fruit are also, susceptible to water loss, decay, and surface and chilling injuries (Kashmere and Kader, 1978; Kader, 1986). External factors associated with tomato maturity such as green color, size or shape, are not reliable indicators for identifying and

* Corresponding author. Tel.: +1 530 7549776; fax: +1 530 7522640.
E-mail addresses: gopal.wsu@gmail.com, goptiwari@ucdavis.edu (G. Tiwari).

0925-5214/$ – see front matter © 2013 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.postharvbio.2013.07.009
sorting immature green tomatoes at harvest or in the packing shed. As a result, significant quantities of immature tomatoes are often delivered to the fresh market (Kader et al., 1977). A handheld management tool based on a rapid, non-destructive internal quality measurement method that could aid in the identification of immature green tomatoes would be useful as a worker training aid for pickers to reduce the harvest of immature fruit, thereby eliminating the shipping and handling costs of marketing poor quality fruit, and in improving quality and consumer satisfaction.

Near infrared (NIR) spectroscopy, as a nondestructive method, has been extensively used to measure the quality of ripe tomatoes. Slaughter et al. (1996) used NIR spectroscopy to determine the soluble solid content of fresh tomatoes of more than thirty varieties. Subsequent studies by Peiris et al. (1998a,b), Shao et al. (2007), Clement et al. (2008), and Xie et al. (2009a) used NIR spectroscopy to measure tomato color, ethylene content, firmness, pH, soluble solids content, and acidity. The effect of storage temperature on the firmness, moisture content and the concentration of pectin degrading enzymes in tomatoes were investigated by measuring their optical properties (Van Dijk et al., 2006a,b). NIR spectroscopy has also been explored to identify the difference between transgenic and non-transgenic varieties of tomatoes (Xie et al., 2007). Qin and Lu (2008) used a hyperspectral imaging system to measure the optical properties of several fruit including tomatoes within the wavelength band 500–1000 nm. They reported that the spectral data could characterize the changes in chlorophyll and lycopene contents in tomatoes with ripening. Recently, mid infrared spectroscopy (2500–25,000 nm) has also been explored to measure tomato quality in terms of dry matter, soluble solids, acidity, citric acid and individual sugar contents (Schisz et al., 2011).

Despite all these research efforts, very few studies have been reported for predicting maturity in green tomatoes and none with handheld instruments. Worthington (1974) used light transmittance spectroscopy to determine maturity in green tomatoes. They reported that light transmittance through whole green tomatoes was difficult, as they were optically very dense. Natuvetty and Chen (1980) attempted to predict green tomato maturity by measuring light transmittance passing through a small region of the fruit. A high density, non-portable, bench-type spectrometer was used to measure light transmission between wavelength 450 and 800 nm. Fruit were classified separately into four or two maturity groups. In the four group classification system, the maturity classes were defined as tomatoes, which exhibited a 75% red/green color ratio on their external surface within 0–7, 8–14, 15–21, and 22–28 d, respectively. In the two group system, the maturity classes were defined as tomatoes, which exhibited the 75% red/green color ratio in 0–14 or 15–28 d, respectively. Simple linear regression models showed that performance to sort green tomatoes in two maturity groups were better (class accuracies ranging from 61% to 98%) than when classified into four maturity groups (class accuracies ranging from 55% to 79%). However, these spectral experiments were run in a laboratory setting and their linear models only used absorption values at two wavelengths (615 and 730 nm). Brecht et al. (1991) used X-ray computed tomography images to identify green immature tomatoes. In that study, absorption values of X-ray images of freshly harvested green tomatoes were determined. Fruit were then immediately cut and classified into four maturity groups, based on the visual appearance of gel in their locules. A multivariate discriminant analysis was used to correlate calculated absorption values of X-ray images of the fruit with their maturity classes. However, only forty-eight tomatoes were used in their study. Saltveit (1991) used nuclear magnetic resonance (NMR) imaging techniques on green tomatoes to determine their maturities. Images of green tomatoes were captured and analyzed for their inner structural details. He identified different maturity stages in green tomatoes based on the available water content in them, but reported that NMR incurs high operational cost and long processing times. Recently, Qin et al. (2012) applied spatially offset Raman spectroscopy to investigate the internal maturity in intact tomatoes. The spectra of tomatoes at different ripeness stages (from green to red) were collected in the lab from wavelengths 4000 to 50,000 nm using a Raman imaging spectrometer mounted on a camera. It was reported that carotenoids were absent in green immature tomatoes. All these prior efforts to determine maturity in green tomatoes were performed using bench top instruments. A portable handheld instrument should be useful to the produce industry as a management tool for identifying and removing immature green tomatoes at the time of harvest or in packing shed. To address this need, a portable handheld instrument capable of rapid, non-destructive visible and NIR interac tance spectroscopic measurements on intact tomato fruit was investigated. The specific objectives of the research were:

1. to develop, calibrate, and validate visible and NIR-based statistical models for maturity classification in green tomatoes using whole fruit interac tance measurements, and
2. to investigate the feasibility of using a handheld spectrometer for maturity determination in green tomatoes.

2. Materials and methods

2.1. Sample collection and data acquisition

Three hundred fifty green tomatoes of each of two cultivars ‘901’ and ‘Bobcat’ were collected from two different commercial packaging houses in California’s Central Valley during the month of June (cultivar ‘901’) and again in October (cultivar ‘Bobcat’). Though only cartons of green tomatoes were selected, it was noticed that some of the tomatoes, especially in the ‘901’ variety, were at the breaker stages (between 0% and 10% red color in their skins). After exclusion of breakers, only green tomatoes, 180 fruit of cultivar ‘901’ and 342 fruit of cultivar ‘Bobcat’, were used in the study.

A handheld spectrometer (Model Nirvana-Analytical Spectrometer, Integrated Spectronics, Sydney, Australia) was used to measure the interac tance spectrum of the intact tomatoes between the wavelengths of 400 nm and 1000 nm at a data resolution of 3 nm and an optical bandwidth ranging from 8 nm to 13 nm (Fig. 1). The instrument used a 0° illumination angle, 0° detection view angle, where a tungsten light source illuminated the fruit through a sampling window of diameter 30 mm. The projected area of the
illumination on the fruit surface by the optical beam was about 7 cm². A small size detector acceptance probe was placed at the center of the optical beam, which obscured the optical beam and caused a shadow on the fruit surface (Greensill and Walsh, 2000). Under this configuration, the detector only received the interactance spectra that emerged from the shadowed area. The spectrophotometer was controlled by a built-in pocket-pc style computer running the Windows CE operating system. Once a model is developed by the user to determine a desired quality index (e.g., soluble solids content of a fruit) the user can upload the model coefficients into the instrument. The instrument is then able to predict the desired quality index of future fruit samples based on the uploaded model for real-time use by managers in the produce industry. Unlike many bench top NIR instruments, this instrument is small and low mass (~1 kg.), and can be carried easily for the spectral measurements of fruit in different locations along the postharvest supply chain.

To measure the interactance spectrum, each fruit was gently held against the outer rim of the sampling window of the spectrometer. Typically, the samples completely covered the sampling window in order to minimize the amount of incoming ambient light and adversely affecting the spectral measurements. The measurement sequence automatically inserted an internal gold reference plate in front of the fruit for use as the optical standard just prior to each spectral measurement and after the fruit was in the measurement position. This feature helped to automatically correct the spectra for variation in ambient light effects. Onsite spectral measurements were carried out on each tomato fruit at two randomly selected equatorial positions (180° apart). The complete time for each spectral measurement was about 5–6 s, including the time to take the reference spectrum. Processing of the interactance spectra into absorbance and second derivative of absorbance was automatically conducted within the instrument during this time period as well.

After onsite optical measurements were collected, the green tomatoes were transported to the University of California, Davis and placed in a ripening room (20 °C, 85% RH). Fruit were then allowed to ripen for a week. For this study, tomatoes that were still green after one week in the ripening room were considered to be immatures.

During the week of ripening, visual observations of external appearance were conducted and tomatoes were classified into six ripeness stages. These stages followed the color classification requirements of the USDA standard for grading fresh tomatoes (CFR, 1991; Choi et al., 1995). The stages are called green (fruit skins are completely green), breaker (less than 10% red color in fruit skins), turning (over 10% but less than 30% red color in fruit skins), pink (over 30% but not more than 60% red color in fruit skins), light red (over 60% but not more than 90% red color in fruit skins), and red (more than 90% red color in fruit skins). Scores on a scale of 1–6 were assigned to the tomatoes based on the development of red color in their skin. Scores 1 and 6 were assigned to the tomatoes, which were either completely green or completely red, respectively.

Tomatoes were then classified into three different maturity groups based on their visual color scores after ripening. Fruit with a score of less than 2 were classified as immature, fruit having scores from 2 to 4 were placed in the ‘intermediate’ maturity group (called Intermediate), and fruit having scores greater than 4 were placed in the ‘advanced’ maturity group (called Advanced). Data sets were formed by merging second derivative of absorbance (D2A) values of fruit with their respective maturity classes. Nattuvetty and Chen (1980) adopted a similar approach where green tomatoes were classified into four maturity classes, based on the development of red color on their surfaces during ripening. Though the present study only aimed to identify immature tomatoes, it was nonetheless desirable to classify the tomatoes into three maturity classes as it should help growers and packing shed managers to manage harvest operations and determine market suitability of the tomatoes based on their maturity.

2.2. Development of the multivariate models

2.2.1. Selection of spectral waveband

Selection of the proper wavelength region is important in analyzing the relationship between the model and chemical compositions of the sample (Xie et al., 2009b) and model predictability and robustness can be increased by excluding the irrelevant and noisy regions of spectra (Han et al., 2008) as optical measurements often contain noise and irrelevant information (such as scattered ambient light) related to the response variable (Vigneau et al., 1997). To select an appropriate waveband for the model, waveband sections along with the full waveband 400–1000 nm were evaluated using canonical discriminant analysis. The analysis of these preliminary models indicated that including the interactance spectra between 400 nm and 600 nm and between 750 nm and 1000 nm to models using the 600–750 nm waveband region did not appear to improve models accuracies for maturity classification of these two cultivars. Therefore, the 600–750 nm waveband region was selected for canonical discriminant classification model development and validation in this study.

2.2.2. Canonical discriminant analysis and Bayesian classification

Canonical discriminant analysis (CDA) was performed on the spectral variables of the datasets to reduce the multicollinearity among variables. High levels of multicollinearity among spectral variables reduces model robustness and its performance, unless methods like CDA are utilized (Rencher, 1992; Graham, 2003), which frequently reduces the dimensions of data set by providing fewer, more relevant canonical variables as the linear combinations of original spectral variables. Since the number of maturity classes in the present study was three, CDA provided only two canonical variables by projecting high dimensional data (45 wavebands between 600 nm and 750 nm) on to a two dimensional space in such a way that separation among three maturity classes was maximum relative to their within class variances. PROC CANDISC in the SAS statistical software (version 9.2, SAS Institute Inc., Cary, NC, USA) was used to perform CDA.

After CDA, Bayesian discriminant functions were used to classify fruit into three maturity classes. The two canonical variables from CDA were used as a two-dimensional feature vector in order to develop a Bayesian discriminant function for each class. If the features are normally distributed and x is a d dimensional feature vector, the Bayesian discriminant function for the ith class can be expressed as (Duda et al., 2001):

$$g_i(x) = -\frac{1}{2} (x - \mu_i)\Sigma_i^{-1} (x - \mu_i) - \frac{1}{2} \Sigma_i - \ln P(w_i)$$

where \(\mu_i\) is the mean of featured vector x, \(\sum_i\) and \(\Sigma_i\) are the inverse and determinant of the covariance matrix \(\Sigma_i\) of d x d dimensions, and \(P(w_i)\) is the prior probability for the class i.

The statistical procedure PROC DISCRIM in SAS was used to generate Bayesian discriminant functions. The procedure also computes the posterior probabilities for each spectral measurement belonging to each of the three maturity classes. In cases when a fruit had positive posterior probabilities in two or three different maturity classes based on the spectral measurements on its two different cheeks, these posterior probabilities were compared and the fruit was placed in the maturity class, which had the highest posterior probability.

To validate the performance of the developed models, a custom built SAS macro was used to combine procedure PROC CANDISC and PROC DISCRIM together. The custom built macro, which would sequentially combine these procedures, was required for the
3. Results and discussion

3.1. Mean spectra of green tomatoes of different maturity

The mean D2A spectra of green tomatoes for the three maturity classes of varieties ‘901’ and ‘Bobcat’ for the full 400–1000 nm waveband are shown in Fig. 2. It was clear from these figures that the mean D2A spectra of these three maturity classes in both varieties differed mainly in the 400–750 nm region, and that the absorbance patterns were different for all three maturity classes in both varieties. It was also noticeable that the mean D2A spectral values for all three maturity classes in variety ‘Bobcat’ were lower than those of variety ‘901’. The difference in mean D2A spectra between these two varieties could be variety specific, however in this study, variety and growing conditions were confounded and this hypothesis could not be tested. Fig. 3 shows a more detailed view of the mean D2A spectra of all three maturity classes of these two varieties in a narrow waveband range (660–690 nm) typically associated with chlorophyll content, which absorbs light at these wavelengths (Choi et al., 1995; Clement et al., 2008). In both varieties, the mean D2A value of the ‘Immature’ class was the lowest in the 675–678 nm region, suggesting that the amount of chlorophyll content in green tomatoes was the highest in the ‘Immature’ maturity class and consistently changed during the tomato maturation process in both varieties. Fig. 3 also showed that the mean D2A values at the chlorophyll peak for the maturity classes of variety ‘Bobcat’ were more equally spaced than those of variety ‘901’. A possible reason for the dissimilar rate of change in absorbance at the chlorophyll peak during the tomato maturation process could be a characteristic of the variety or because they were harvested in different months from farms in which the growing conditions (particularly temperature) were quite different, or an interaction of these two factors.

Fig. 2. The average second derivative of absorbance spectra of green tomatoes of varieties (a) ‘901’ and (b) ‘Bobcat’ under three maturity classes from wavelength band 400 to 1000 nm. The average and standard deviation of the spectra of the green tomatoes of variety ‘901’ and variety ‘Bobcat’ are based on 62 Immature, 88 Intermediate, 30 Advanced mature and 39 Immature, 145 Intermediate, and 158 Advanced mature fruit, respectively.

Fig. 3. The average second derivative of absorbance spectra of green tomatoes of varieties (a) ‘901’ and (b) ‘Bobcat’ under three maturity classes from wavelength band 660 to 690 nm.
Table 1
Internal calibration with one-out cross validation performance of ‘901’ specific model.

<table>
<thead>
<tr>
<th>Maturity class</th>
<th>Actual number of fruit in each maturity class</th>
<th>Predicted maturity class in percent (with number of fruit below)</th>
<th>False negative error rate (%)</th>
<th>False positive error rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Immature</td>
<td>Intermediate</td>
<td>Advanced</td>
</tr>
<tr>
<td>Immature</td>
<td>62</td>
<td>75.81</td>
<td>22.58</td>
<td>1.61</td>
</tr>
<tr>
<td>Intermediate</td>
<td>88</td>
<td>22.72</td>
<td>54.55</td>
<td>22.73</td>
</tr>
<tr>
<td>Advanced</td>
<td>30</td>
<td>3.33</td>
<td>40.00</td>
<td>56.67</td>
</tr>
<tr>
<td>Mature (Intermediate + Advanced)</td>
<td>118</td>
<td>17.80</td>
<td>82.20</td>
<td></td>
</tr>
</tbody>
</table>

3.2. Prediction of maturity of green tomatoes using variety-specific models

3.2.1. Prediction performance of the ‘901’ model

Tables 1 and 2 show the prediction accuracy of the ‘901’ model for identifying maturity in green tomatoes for both varieties ‘901’ and ‘Bobcat’. Internal calibration with one-out cross-validation results from Table 1 show that the ‘901’ specific model correctly predicted about 76% of ‘Immature’ green tomatoes when it was used on the fruit of variety ‘901’. In comparison, Nattuvetty and Chen (1980) were able to detect 60–64% immature green tomatoes using light transmission in four selected varieties, when fruit were classified into four maturity groups by color after ripening. Using X-ray images, Brecht et al. (1991) identified 91% (11 out of 12 fruit) immature green tomatoes, in a study of 48 fruit. The maturity of the green tomatoes in the Brecht et al. (1991) study were based on the gel present in their locules before ripening, which may impact a direct comparison with the results here.

Table 1 also shows that nearly all (14 of 15 fruit) misclassified immature fruit were wrongly placed in the adjacent Intermediate maturity class. The classifying error between these two neighboring classes could be partially due to sampling error given the fact that the two spectral data per fruit were collected from two random spots on its different cheeks. The total measured area of these two spots was only between 16% and 22% of the total surface area of the fruit. Kashmire and Kader (1978) reported that Immature green tomatoes completely lack gel in their locules, whereas partially mature tomatoes may have gel in some of their locules. Since the interactance spectral measurement used in this study was only a spot measurement and only provides spectral information about the tissue of the fruit in the immediate proximity of the measurement, it is possible that in some green fruit, the measured spectra were not taken over locular areas and could not accurately reflect their actual maturities in this case. A similar, but reversed, misclassification pattern was observed for the Advanced maturity fruit where twelve out of thirteen green tomatoes were falsely classified as Intermediate, also a neighboring maturity class error. The total false positive identification of immature fruit was predominantly fruit from the Intermediate class (20 of 21 false positives) showing that two-step (‘Immature’ vs. Advanced) errors were uncommon.

The ‘901’ model correctly identified only 55% Intermediate green tomatoes. The comparatively high prediction error for the Intermediate maturity green tomatoes could be due to maturity assessment subjective decision based on the red color development in their peels after a week of ripening. Nattuvetty and Chen (1980) also reported lower model prediction accuracies (ranging between 55% and 78%) for identifying intermediate maturities in green tomatoes for four tomatoes varieties. Brecht et al. (1991) reported prediction accuracies of 58% and 75% of their model to detect green tomatoes placed in the two intermediate classes. In case of detecting advanced maturity in green tomatoes, the model only identified 57% fruit of this class. Previous studies (Nattuvetty and Chen, 1980; Brecht et al., 1991) reported that the prediction performances of their models in identifying advanced maturity tomatoes varied from 66% to 84%. The mean spectral D2A graphs in Fig. 3(a) indicated that in variety ‘901’, the tomato maturation process was not uniform and mean spectral graphs for the Intermediate and the Advanced maturity classes were close. This could be a possible reason for the reduced prediction performance of the ‘901’ model to detect maturities for these two classes. The model showed high misclassification rates for these two classes as it wrongly placed 23% of Intermediate mature green tomatoes in the Advanced class, and 40% of Advanced maturity fruit were wrongly placed in the Intermediate class. Misclassified Intermediate fruit were equally split between the ‘Immature’ and Advanced classes. Part of the error in model prediction could also be due to the incorrect subjective visual assessment of maturities in fruit after a week of ripening. It is important to mention that for commercial purposes, growers and packaging shed managers may require only two maturity categories (immature and mature green tomatoes) at the time of harvest. Therefore, the model performance was also investigated by relaxing the classification and combining Intermediate

Table 2
Prediction performance of ‘901’ specific model, when it was tested externally on the variety ‘Bobcat’.

<table>
<thead>
<tr>
<th>Maturity class</th>
<th>Actual number of fruit in each maturity class</th>
<th>Predicted maturity class in percentage (with number of fruit below)</th>
<th>False negative error rate (%)</th>
<th>False positive error rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Immature</td>
<td>Intermediate</td>
<td>Advanced</td>
</tr>
<tr>
<td>Immature</td>
<td>39</td>
<td>100.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Intermediate</td>
<td>145</td>
<td>68.96</td>
<td>26.90</td>
<td>4.14</td>
</tr>
<tr>
<td>Advanced</td>
<td>158</td>
<td>13.29</td>
<td>23.42</td>
<td>63.29</td>
</tr>
<tr>
<td>Mature (Intermediate + Advanced)</td>
<td>303</td>
<td>45.21</td>
<td>54.79</td>
<td></td>
</tr>
</tbody>
</table>
Table 3

<table>
<thead>
<tr>
<th>Maturity class</th>
<th>Actual number of fruit in each maturity class</th>
<th>Predicted maturity class in percentage (with number of fruit below)</th>
<th>False negative error rate (%)</th>
<th>False positive error rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Immature</td>
<td>Intermediate</td>
<td>Advanced</td>
</tr>
<tr>
<td>Immature</td>
<td>39</td>
<td>84.62 (33)</td>
<td>15.39 (6)</td>
<td>0.00 (0)</td>
</tr>
<tr>
<td>Intermediate</td>
<td>145</td>
<td>24.83 (36)</td>
<td>58.62 (85)</td>
<td>10.55 (24)</td>
</tr>
<tr>
<td>Advanced</td>
<td>158</td>
<td>3.16 (5)</td>
<td>13.30 (21)</td>
<td>83.54 (132)</td>
</tr>
<tr>
<td>Mature (Intermediate + Advanced)</td>
<td>303</td>
<td>86.47 (262)</td>
<td>3.53 (41)</td>
<td>15.33</td>
</tr>
</tbody>
</table>

and Advanced maturity classes together. In that scenario, the ‘901’ specific model was able to correctly distinguish 82% of green tomatoes of these classes (Table 1) with a false negative and positive error rate of ∼24%.

When the ‘901’ specific maturity model classifier was externally applied on the fruit of variety ‘Bobcat’, the model correctly detected 100% of immature green tomatoes of that variety (Table 2). But the model also incorrectly placed 70% of Intermediate and 13% of Advanced mature green tomatoes in the ‘Immature’ class, as reflected by its very high false positive error rate (∼40%) for this class. The model only detected 27% Intermediate maturity and 63% Advanced maturity in green tomatoes. It is clear that variety specific model ‘901’ performed poorly for the maturity prediction in ‘Bobcat’ variety. However, when the model classifiers were relaxed to detect Intermediate and Advanced maturity classes together, its predictive performance to identify these groups together was about 55%.

3.2.2. Prediction performance of ‘Bobcat’ model

Table 3 shows the performance of the ‘Bobcat’ specific model, when it was internally calibrated and cross-validated on the variety ‘Bobcat’. The model performed fairly well in detecting immature green tomatoes as it correctly detected about 85% of ‘Immature’ green tomatoes. The misclassified immature fruit were placed entirely into the neighboring Intermediate maturity class. The false positive error rate (∼14%) for the ‘Immature’ class was primarily caused by misclassification of Intermediate tomatoes as was observed in Table 1. The ‘Bobcat’ model was also able to detect 59% Intermediate and 84% of Advanced maturity fruit. When these classes were combined together as a single mature class, model detected 86% of mature green tomatoes. Overall, the ‘Bobcat’ model had better performance in predicting maturities in green tomatoes of the same cultivar than the ‘901’ model. These results are consistent with the differences between the cultivar spectra observed in Fig. 3.

When the ‘Bobcat’ specific model was externally tested to predict classification accuracies of green fruit in variety ‘901’, the model accurately detected 70% of Immature, 33% of Intermediate and 97% of Advanced classes of green tomatoes for variety ‘901’ (Table 4). It is clear from these results that this model did fairly well in detecting ‘immature’ and Advanced maturities of green tomatoes of the ‘901’ variety. Unlike the results in Table 2, the misclassified Intermediate fruit were mainly put into the Advanced class. When intermediate and advanced maturity green tomatoes were combined, the model identified 97% of the ‘901’ mature tomatoes, however the false positive rate for the mature class was still quite high (∼31%).

3.3. Posteriori probability contours and Bayesian decision boundaries

The difference in predictive performance of these two models can be visualized by examining the posterior probability contour maps of the maturity classes and models’ classification decision boundaries in canonical feature space the three classes result in two canonical feature vectors. Fig. 4(a) illustrates the scenario when the ‘901’ specific model was applied to fruit of the same variety. This contour map shows the overlap in the probability distribution contours of all three maturity classes. It is clear from this figure that the classification boundaries could not be placed to separate these classes completely. The ‘Immature’ class was the least overlapped and best separated by the classification boundaries. This explains the good predictive ability of the model for detecting immature fruit compared to other two classes (Table 1). On the other hand, the contours of the Intermediate maturity and Advanced group were highly overlapped with each-other. Here the classification boundaries could only separate about 55% and 57% of fruit belonged to these two maturity classes. The highly overlapped posterior probability distribution curves of these two classes also explained the model’s high false positive error rates shown in Table 1. When the two mature groups were combined, model’s predictive ability increased.

Table 4

<table>
<thead>
<tr>
<th>Maturity class</th>
<th>Actual number of fruit in each maturity class</th>
<th>Predicted maturity class in percentage (with number of fruit below)</th>
<th>False negative error rate (%)</th>
<th>False positive error rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Immature</td>
<td>Intermediate</td>
<td>Advanced</td>
</tr>
<tr>
<td>Immature</td>
<td>62</td>
<td>69.35 (43)</td>
<td>41.86 (18)</td>
<td>1.61 (1)</td>
</tr>
<tr>
<td>Intermediate</td>
<td>88</td>
<td>4.54 (4)</td>
<td>32.95 (29)</td>
<td>62.5 (55)</td>
</tr>
<tr>
<td>Advanced</td>
<td>30</td>
<td>0.00 (0)</td>
<td>3.33 (1)</td>
<td>96.67 (29)</td>
</tr>
<tr>
<td>Mature (Intermediate + Advanced)</td>
<td>118</td>
<td>96.61 (114)</td>
<td>3.39 (4)</td>
<td>3.39</td>
</tr>
</tbody>
</table>
When classification boundaries of the ‘901’ model were projected on the feature space with the posterior probability contour maps from the ‘Bobcat’ cultivar, the contours of the Intermediate and Advanced maturity classes were shifted into the neighboring class (Fig. 4(b)). External validation performance of a model depends upon the relative stability of both the shape of the probability contours and their respective positions with respect to decision boundaries. The non-optimal shape and shifting of probability contours of variety ‘Bobcat’ with respect to classifier ‘901’ was due to difference in spectral patterns between these two varieties for each maturity class (Figs. 2 and 3). There was more overlap between Immature and Intermediate maturity classes and they were shifted with respect to the class boundaries. The shift in the ‘Immature’ class away from the ‘901’ decision boundaries increased the model accuracy for identifying this class (Table 2) as probability contours of this class were well separated by the model classifier. But at the same time, the probability contours of the Intermediate class exhibited a large shift across the decision boundary with increasing the overlap with the contours of the ‘Immature’ class. These effects increased the model false positive error rate for the immature class (Table 2). The shift in the Intermediate class was also responsible for the improved efficacy for detection of fruit in the Advanced maturity class (Table 2).

Fig. 5(a) illustrates the scenario when the ‘Bobcat’ model was used on the same variety. It is clear that the posterior probability contours of the ‘Immature’ and Advanced classes were better separated by the model decision boundaries than those of the Intermediate maturity class. This explained the increased performance in detecting maturities in these classes (Table 3). It is also clear from this figure that the probability contours of the Intermediate class were highly overlapped with the ‘Immature’ and Advanced maturity classes, which explained the model’s reduced performance in identifying maturity for this class. Similarly Fig. 5(b) shows the classification boundaries of the ‘Bobcat’ model projected on the feature space with the posterior probability contours from ‘901’ cultivar. It is clear from this figure that the model was least efficient in predicting Intermediate maturity in green tomatoes, but that the amount of shifting was reduced when compared to Fig. 4(b).

3.4. Global model accuracy to predict tomato maturity

It is clear from these results that variety-specific models performed better when they were validated on fruit from the same variety, though the ‘Bobcat’ model performed fairly well when it was used on the variety ‘901’. But constructing separate models for each and every tomato variety is not desirable; therefore a global model was created by training on the combined data sets of both varieties in an attempt to improve the model robustness and ease of use across both tomato varieties. Table 5 compares the predicted maturity of green tomatoes using the global model with
their maturity determined by visual assessment. Results show that the overall internal calibration and cross-validation performance of the global model in predicting the maturity of green tomatoes was more stable than the external validation performance of the two variety-based models. The global model was able to correctly detect about 71% of immature green tomatoes. The global model also correctly identified the maturity in 55% and 78% of maturity fruit in the other two classes. When the intermediate and advanced classes were combined, the global model correctly identified 85% of mature green tomatoes. It was also noticeable that maturity prediction performances for Intermediate and Advanced classes of variety ‘Bobcat’ under the global model were much better (57% and 82%) than the performances of these classes (27% and 63%) under the ‘901’ model (Table 2).

This pattern of cross-cultivar calibration performance in tomato is consistent with calibration transfer performance across cultivars or harvest dates reported in previous studies in other fruit. For example, Peiris et al. (1998a,b) developed NIR transmission models to determine the soluble solids content of four peach cultivars. When they applied cultivar-specific models across cultivars they observed up to a 5-fold increase in model bias. A similar declining performance pattern in the multiple-season application of a model developed from single season data was also observed. When a global model was developed using spectral data from all four cultivars, the validation performance of the global model was superior to the cross-cultivar performance of single cultivar models. A similar pattern of improved multi-cultivar performance of a global model as compared to cross-cultivar performance of single cultivar-based models was also observed by Peiris et al. (2003) in apple and by Golic and Walsh (2006) in peach. Peiris et al. (2003) hypothesized that cultivar differences in moisture content and other chemical constituents or light scattering properties may affect the model optimization process that then adversely affects their cross-cultivar behavior.

In an attempt to examine NIR model stability over fruit populations with different harvest dates, growing region and cultivar, Subedi et al. (2007) explored a sequential model updating process where the model was iteratively updated with a new population of data, then used in prediction of the next population before updating the model with that set. They observed that the model updating process produced unstable results until the inclusion of the fourth data population in the model. They also observed that cross-cultivars model prediction was more problematic for predicting Hunter b values than dry matter content.

Our results indicated that initial concentration of chlorophyll content and rate of ripening in tomatoes are variety dependent. Assuming that the cross-cultivar stability of NIR models in tomato will follow the pattern observed in other fruit, including more varieties in different calendar years should provide more stability to the model. Though it is difficult to state the exact number of tomato varieties to be included in the model in order to rightly predict the maturity levels in green tomatoes, the selection of varieties to incorporate a broad range of chemical characteristics and ripening patterns should increase the robustness of the model. As a handheld management tool for evaluating fruit maturity for determining harvest timing and as a training aid for workers picking the fruit, the observed level of accuracy should help farmers reduce the quantity of immature green tomatoes harvested.

4. Conclusions

Cross-validation results show that the variety-specific models were correctly able to identify 76% and 84% of green immature tomatoes when applied to their respective varieties. A global model was able to identify 71% immature green tomatoes with no prior knowledge of the variety and it had more stable performance than when the variety-specific models were applied to a variety outside the training set. The nondestructive measurement of this handheld device showed good potential for aiding production managers in identifying immature green tomatoes at harvest or in the packing shed. Further investigation using spectra from additional varieties of green tomatoes should help to develop a more robust global model.

Acknowledgements

This research is supported by the Specialty Crops Research Initiative Competitive Grants Program grant no 2009-51181-05783 from the USDA National Institute of Food and Agriculture. The authors thank Jeff Dolan (DiMare Fresh, Incorporation, Newman, CA, USA) for providing tomato fruit for this analysis.

References


