Postharvest Biology and Technology 52 (2009) 103–109

Contents lists available at ScienceDirect

Postharvest Biology and Technology

journal homepage: www.elsevier.com/locate/postharvbio

# Rapid estimation of lycopene concentration in watermelon and tomato puree by fiber optic visible reflectance spectroscopy

R. Choudhary[a,*], T.J. Bowser[a], P. Weckler[a], N.O. Maness[b], W. McGlynn[b]

[a] Department of Biosystems and Agricultural Engineering, Oklahoma State University, Stillwater, OK 74078, USA
[b] Department of Horticulture and Landscape Architecture, Oklahoma State University, Stillwater, OK 74078, USA

ARTICLE INFO

*Article history:*
Received 27 March 2008
Accepted 11 October 2008

*Keywords:*
Lycopene
Carotenoid
Watermelon
*Citrullus lanatus* tomato
*Lycopersicon esculentum*
Visible reflectance spectroscopy
Carotenoid measurement
Fiber optic sensing
Chemometry

ABSTRACT

Chemometric models were developed for prediction of lycopene concentration in watermelon and tomato puree from their visible reflectance spectra acquired by a fiber optic reflectance probe. A fiber optic spectrometer was used to acquire reflectance spectra from puree samples in the wavelength range of 500–750 nm. Least squares (LS) and partial least squares (PLS) regression were used to correlate spectral data with lycopene concentration measured by hexane extraction and spectrophotometry. An apparent absorbance index (AAI) obtained by subtracting apparent absorbance at 700 nm from that at 565 nm showed linear correlation with lycopene concentration ($R^2$ = 0.90 for watermelon puree and 0.62 for tomato puree). A normalized apparent absorbance index (NAAI) obtained by dividing the AAI by the sum of apparent absorbances at 565 and 700 nm, also had linear correlation with lycopene concentration ($R^2$ = 0.90 and 0.61 for watermelon and tomato, respectively). The LS linear regression model for watermelon puree could predict lycopene concentration with $R^2$ of 0.93, and standard error of prediction (SEP) of 5.1 mg kg$^{-1}$. The LS linear regression model for tomato puree could predict lycopene concentration with $R^2$ of 0.54 and an SEP of 5.2 mg kg$^{-1}$. The PLS model for watermelon puree could predict lycopene concentration with an $R^2$ of 0.97 and an SEP of 3.4 mg kg$^{-1}$. The PLS model for tomato puree could predict lycopene concentration with an $R^2$ of 0.88 and an SEP of 2.5 mg kg$^{-1}$. The high linear correlations between spectral parameters and lycopene concentration of samples (with lycopene concentration between 10 and 80 mg kg$^{-1}$) suggest that this method can be reliably used for fast and safe quantification of lycopene concentration in watermelon and tomato puree.



## 1. Introduction

Lycopene is a hydrocarbon carotenoid, $C_{40}H_{56}$, with molecular weight of 537. In nature, it is most abundantly found as the red pigment of watermelon and tomato. It has recently received attention for its potential role in preventing prostate cancer and cardiovascular disease in humans (Arab and Steck, 2000; Hadley et al., 2002; Heber and Lu, 2002). It is a natural antioxidant due to its ability to act as a free radical scavenger. It has the highest singlet oxygen-quenching rate of all carotenoids in biological systems (Di Mascio et al., 1989; Tinker et al., 1994).

Due to the increasing popularity of lycopene for use in food and nutritional supplements, scientists are interested in developing lycopene rich products and ingredients by extracting lycopene from watermelon and tomato. The process of recovery and use of lycopene requires measurement of lycopene in raw materials, intermediate products and final products. Present methods of lycopene assay require extraction of lycopene from samples in hazardous organic solvents, which are tedious, hazardous and destructive. A rapid method for lycopene quantification in fruits and vegetables and their products is urgently required, which should be portable and easily usable in the field or at a postharvest handling facility.

Although lycopene may be present in small quantities in apricot, pink fleshed guava, and pink grapefruit its availability is another issue (Mangels et al., 1993). In western diets, major natural sources are red-fleshed watermelon and tomato (USDA, 2007). Lycopene concentration in red-fleshed watermelon of different cultivars can be as low as 12 mg kg$^{-1}$ in some seeded cultivars to 100 mg kg$^{-1}$ in some individual seedless watermelons (Tomes et al., 1963; Perkins Veazie et al., 2001). Lycopene concentration in fresh tomatoes may vary from 5 to 50 mg kg$^{-1}$, depending on the cultivar, ripening stage,

*Abbreviations:* AA, apparent absorbance; $AA_\lambda$, apparent absorbance at wavelength $\lambda$; AAI, apparent absorbance index; LS, least squares; NAA, normalized apparent absorbance; NAAI, normalized apparent absorbance index; nm, nanometer; PLS, partial least squares; S.D., standard deviation; SEC, standard error of calibration; SEP, standard error of prediction; mg kg$^{-1}$, milligram per kilogram.

* Corresponding author. Present address: Agriculture 176, Southern Illinois University, 1205 Lincoln Drive, Carbondale, Illinois 62901-4415, USA.
Tel.: +1 618 453 2496.
*E-mail address:* ruplal.choudhary@yahoo.com (R. Choudhary).

0925-5214/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.postharvbio.2008.10.002

**Nomenclature**

| | |
|---|---|
| $R_\lambda$ | reflectance at wavelength λ |
| $R^2$ | coefficient of determination |

and temperature during crop growth (Scott and Hart, 1995; Cadoni et al., 2000; Hadley et al., 2002).

Most studies of lycopene assay methods have been conducted using tomatoes (Beerh and Siddappa, 1959; Adsule and Dan, 1979; Sadler et al., 1990; D'Souza et al., 1992; Arias et al., 2000), and watermelons (Fish et al., 2002; Perkins Veazie et al., 2001; Davis et al., 2003a,b). The majority of lycopene assay studies focused on methods involving extraction in organic solvents followed by use of either a spectrophotometer or a high-performance liquid chromatograph (HPLC).

Some studies have looked at developing non-extraction methods for lycopene quantification by using chromaticity of tomatoes (D'Souza et al., 1992) and tristimulus color reading of cut watermelons (Perkins Veazie et al., 2001), but these methods are not precise enough to replace chemical extraction methods.

Davis et al. (2003a,b) developed a color absorbance method for lycopene assay of watermelon and tomato puree. They used a xenon flash colorimeter and spectrophotometer in transmittance mode to obtain color absorbance spectra of watermelon and tomato puree. The xenon flash spectrophotometer had an integrating sphere sensor, which acquired transmittance spectra of liquid samples in cuvette. They correlated scatter adjusted absorbance at 560 nm (absorbance at 560 nm–absorbance at 700 nm) with lycopene concentration in tomato and watermelon puree. A linear relationship ($R^2 = 0.98$) was reported between absorbance difference ($A_{560\,nm} - A_{700\,nm}$) and lycopene content of watermelon puree, with mean standard error of correlation of 2.85% (Davis et al., 2003a). For tomatoes, they obtained an $R^2$ value of 0.96 and a mean standard error of correlation of 14.4% (Davis et al., 2003b). This model was validated using over 200 tomato samples (Perkins Veazie et al., 2006). The primary limitation of this method is that the machine can be used only in a lab setting, due to its bulky, integrating sphere type sensing unit.

Recently mid-IR and Raman spectroscopy techniques were reported for direct determination of lycopene in tomato and tomato products (Schulz et al., 2006; Halim et al., 2006). These authors used lab instruments which are not portable for field application. Most recently, application of a portable NIR spectrometer was reported by Kusumiyati et al. (2008). A partial least squares (PLS) regression model was developed to calibrate the NIR spectra of tomato with its lycopene content. Although portable NIR spectrometers based on InGaAs sensors are available in market, their prices are prohibitive for field application. A visible spectrometer using silicon sensor or a dual band photodiode-based lycopene sensor will prove economical for field assays.

An economical, fast and miniature sensing device is desired to implement a rapid, regretless, lycopene sensing technique for field assays. Due to the miniature size and easy portability of fiber optic sensors, we have investigated the use of a fiber optic diffuse reflectance sensor for quantification of lycopene in watermelon and tomato puree. Results of this study will help in the design of a fiber optic sensor for rapid measurement of lycopene in tomato and watermelon purees to assess the quality and price of these produce based on their lycopene concentration. The specific objective of this study was to develop chemometric models for rapid quantification of lycopene concentration in watermelon and tomato puree from their visible reflectance spectra acquired by a fiber optic reflectance probe.

## 2. Materials and methods

### 2.1. Samples

Watermelons used for this study were grown at Oklahoma Vegetable Research Station, Bixby, Oklahoma in 2003 or were purchased from retail stores. Fruits grown at the research station at Bixby were the seedless cultivar 'Sugar Shack' and were undermature (harvested 7–10 d prior to maturity) or mature. The origin and growing condition of purchased fruits were unknown. Tomatoes used for this study were purchased from local retail stores. The cultivars and postharvest history were not available for tomatoes.

Watermelon flesh and whole tomato were pureed using an Omni mixer homogenizer (Omni Macro, Omni International Inc., Marietta, GA) to produce homogeneous slurry with particle size smaller than 3 $mm^3$, as per the sample preparation method suggested by Davis et al. (2003b).

### 2.2. Sample statistics

A total of 110 samples were used for development and validation of regression models for each fruit type. One fruit sample was used for each assay. Table 1 shows the descriptive statistics for the samples used for calibration and validation of regression models. Lycopene concentrations in 70 samples of the calibration set for watermelon puree varied from 19.9 to 80.7 mg $kg^{-1}$ and that in 40 samples of the validation set varied from 19.3 to 79.0 mg $kg^{-1}$. For tomato puree, 67 samples were used for calibration with a lycopene concentration ranging from 10.0 to 50.5 mg $kg^{-1}$ and 43 samples were used for validation with a lycopene concentration of 15.4–42.9 mg $kg^{-1}$.

### 2.3. Instrumentation

A miniature fiber optic spectrometer (S2000 VIS-NIR, Ocean Optics Inc., Dunedin, FL) was used for spectral measurements. The spectrometer acquired spectra in wavelengths ranging from 350 to 1000 nm at 0.38 nm intervals. It had an optical resolution of 1.5 nm full width half maximum (FWHM). The spectrometer used a diffraction grating to separate incoming light into separate wavelength components. The separated light then fell onto a high-sensitivity linear Charge-Coupled Device detector (CCD, 2048-pixel Sony ILX511 linear sensor) through a 25-μm entrance slit. The CCD collected photons at each pixel during an integration time that

**Table 1**
Descriptive statistics of sample sets used for calibration and validation of lycopene concentration in watermelon and tomato puree.

| Sample set | No. of samples | Lycopene concentration of samples (mg $kg^{-1}$) | | |
|---|---|---|---|---|
| | | Mean | S.D. | Range |
| Watermelon: calibration | 70 | 49.9 | 15.4 | 19.9–80.7 |
| Watermelon: validation | 40 | 42.9 | 18.7 | 19.3–79.0 |
| Tomato: calibration | 67 | 25.6 | 10.9 | 10.0–50.5 |
| Tomato: validation | 43 | 27.8 | 6.7 | 15.4–42.9 |

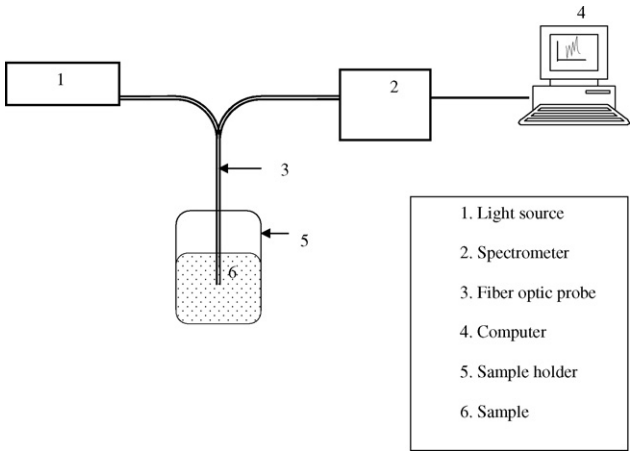


**Fig. 1.** Experimental setup for acquiring visible reflectance spectra of watermelon and tomato puree samples with various lycopene concentrations.

was set by the operating software (OOI Base32, Version 2.0, Ocean Optics Inc., Dunedin, FL). After the integration time (as described in next section) was elapsed, the number of photons collected at each pixel location of the CCD (represented by a voltage) was transmitted from the CCD to an A/D converter board in the computer, where the voltage was converted to a 12-bit digital value. The data were displayed and stored using the operating software OOI Base32 from Ocean Optics. A tungsten halogen light source (LS-1, Ocean Optics Inc., Dunedin, FL) was used for this study. It had a 6.5 W (5 V input) bulb with color temperature of 3100 K. The light from LS-1 was conveyed to the sample through a 200 μm diameter, 2 m long optical fiber assembly (BIF200-VIS-NIR, Ocean Optics Inc., Dunedin, FL).

### 2.4. Acquisition of reflectance spectra

The experimental setup for obtaining reflectance spectra of puree samples has been shown in Fig. 1. A 2-m bifurcated optical fiber cable (BIF200-VIS-NIR, Ocean Optics Inc., Dunedin, FL) was used as a reflectance probe. It was a Y shaped assembly with two optical fibers of 200 μm diameter side by side in the common end. One of the fibers was connected to the light source and the other one to the spectrometer. The common end of the optical fiber cable, thus having a light emitting fiber and a light sensing fiber, was used as a reflectance probe. The optical probe was dipped manually in puree samples contained in 100 mL sample holder (amber glass bottle covered by silver foil) for obtaining diffuse reflectance of puree. It was dipped in the sample such that a sample (puree) thickness of 30 mm was below the probe tip. This was done to ensure that there was enough thickness of sample in the field of view of the probe so as to obtain only the diffuse reflectance of sample and prevent any reflection from the bottom surface of the sample holder. All spectra were acquired under low ambient light level (approximately 300 lx). The response of the spectrometer sensor was observed at varying ambient illumination. There was no affect in the reflectance spectra of samples measured by the spectrometer up to the ambient illumination of 500 lx. This was because the ambient light was attenuated partly by the wall of the sample holder and remaining by the sample thickness, hence the effect of ambient light was not detected by the spectrometer. A white diffuse reflectance standard (WS-1, Ocean Optics Inc.) made of polytetrafluoroethylene (PTFE), a diffuse white plastic that provided a highly Lambertian surface, was used for taking reference spectra at an integration time of 5 ms. The probe tip was placed at a height of 30 mm, in perpendicular to the surface of the reflectance standard. Higher integration time saturated the spectrometer detector. The reflectance values of a puree spectrum were converted to percentage of white reference reflectance. The bifurcated fiber optics probe was dipped in the pureed sample and three spectra for each sample were obtained. Higher integration time (300 ms) was used for obtaining reflectance ($R_\lambda$) of samples. Different integration times were taken for white standard and samples because white standard had very high reflectance and therefore required less integration time (5 ms) to get optimum level of light signal by the CCD sensor of the spectrometer. The puree samples had much lower reflectance compared to the white standard. Therefore to get the optimum level of signal at the sensor, the integration time was increased to 300 ms. Since the diffuse reflectance values have nonlinear relationship with the concentration of analytes in samples due to reflections from several layers inside a thick sample (Hruschka, 2001), reflectance values ($R_\lambda$) were linearized by taking the log of $1/R_\lambda$ and referred as apparent absorbance. Further data analyses were performed on these $\log(1/R_\lambda)$ values, named as apparent absorbance values.

### 2.5. Lycopene assay by hexane extraction method

The lycopene concentration in puree samples was determined by extraction of lycopene in hexane, followed by absorbance measurement of lycopene solution. The lycopene extraction method of Sadler et al. (1990) was used by taking 1 g samples in triplicate. The absorbance of lycopene solution in hexane was obtained by using a spectrophotometer (Shimadzu UV 160U, Shimadzu Scientific Instruments Inc., Columbia, MD) at 503 nm. This wavelength was chosen because lycopene solution in hexane has major absorbance peak at 503 nm, as reported by Sadler et al. (1990) and Davis et al. (2003a,b). Lycopene concentration in each sample was obtained by averaging the lycopene concentrations of triplicate subsamples.

### 2.6. Spectral data analysis

The spectra of watermelon and tomato samples were plotted as 'log $(1/R_\lambda)$' versus 'wavelength $(\lambda)$' in nm. The major peaks of tomato and watermelon spectra were obtained at 565 nm. Therefore the peak height at this wavelength band was chosen for possible correlation with lycopene concentration. One more band at 700 nm was chosen to make a scatter correction for turbidity in samples (Davis et al., 2003a,b). Reflectance values were converted to apparent absorbance units by taking log of $1/R_\lambda$, i.e., $AA_\lambda = \log(1/R_\lambda)$. Using the apparent absorbances at 565 and 700 nm, two indices were developed and evaluated for their possible correlation with lycopene concentration. An apparent absorbance index (AAI) of a sample was obtained as follows:

$$AAI = AA_{565} - AA_{700} \quad (1)$$

In addition to the AAI, which is based on difference between the selected wavelengths, a normalized apparent absorbance index (NAAI) was developed. This index is based on the ratio of difference and sum of the absorbances at the selected wavelengths. This index not only includes the difference between absorbance at selected wavelengths but also includes the effect of average spectra height of a sample. The normalized difference vegetative index (NDVI) popularly used in remote sensing is based on this concept and is a very effective spectral index in crop health monitoring. It is therefore hypothesized that a normalized difference index could be better correlated with lycopene concentration. The normalized apparent absorbance index was calculated using the following equation:

$$NAAI = \frac{AA_{565} - AA_{700}}{AA_{565} + AA_{700}} \quad (2)$$

### 2.7. Least squares regression

Least squares (LS) linear regression analysis of the AAI and NAAI to lycopene concentration was conducted. The linear regression models developed using calibration samples were used to predict lycopene concentration in validation samples. Coefficient of determination, standard error of prediction (SEP) and bias were used to evaluate the prediction ability of the models. SEP was obtained by calculating standard deviation of residuals (difference between predicted and actual values of lycopene concentration), as follows:

$$\mathrm{SEP} = \sqrt{\frac{\sum_{i=1}^{n}(Y_{\mathrm{actual}} - Y_{\mathrm{predicted}})^2}{n-1}} \quad (3)$$

where $n$ is the number of samples, $Y_{\mathrm{actual}}$ is the actual lycopene concentration measured by reference method (hexane extraction), and $Y_{\mathrm{predicted}}$ is the predicted lycopene concentration by the regression model.

Mean values of the residuals (difference between actual and predicted lycopene concentration) were reported as bias. The LS regression models for tomato and watermelon were compared using PROC GLM procedure of SAS V8.2 (SAS Institute Inc., Cary, NC) taking fruit type (tomato, watermelon) as a class variable.

### 2.8. Partial least squares regression

Partial least squares regression has been extensively used for non-destructive quality evaluation of food and agricultural products by NIR spectroscopy (Williams and Norris, 2001). PLS regression is a multivariate statistical analysis tool for chemometric applications. It uses apparent absorbance values at each wavelength of the spectrum of a sample as independent variables and correlates them with its lycopene concentration. PLS regression is expected to give better correlation than the LS regression because it incorporates the effects of all the independent variables as compared to the LS regression, which uses a only a few selected wavelengths (in AAI and NAAI). PLS algorithm uses a data compression by projection of each independent variable (apparent absorbance values at each wavelength) onto imaginary axes called 'principal components', or factors, which account for most of the variations in the data. These new variables (factors), which are orthogonal to each other, thus remove the collinearity in the original variables. The PLS method creates new variables (principal components) by decomposing the original spectral data which are correlated and weighed proportionally with the analyte (lycopene in our case) concentration. Thus PLS uses single step decomposition and regression. The first principal component (or a factor) captures maximum variation in the spectral data which are correlated to the concentration. The subsequent factors capture successively less variation in spectral data relevant to the concentration. Details of the PLS algorithm has been explained in Williams and Norris (2001). PLS regression was performed by a commercial software package (Unscrambler® v7.8, Camo Process AS, Norway). PLS calibration models were developed for the wavelength region of 500–750 nm. Though the spectral data were acquired in the range of 350–1000 nm, preliminary data analysis using correlation plots of the apparent absorbance spectra showed lower correlation coefficients for the wavelength range of 350–500 nm and 750–1000 nm. This could be due to the lycopene is sensitive only in the visible range of spectrum. Using PLS regression in the whole spectrum increased the errors of the PLS model. Hence, the wavelength range of 500–750 nm was used for PLS regression analysis because this spectral range gave PLS models with lower errors. The optimal numbers of factors (principal components) were determined by minimizing the predicted residual sum of squares (PRESS). The

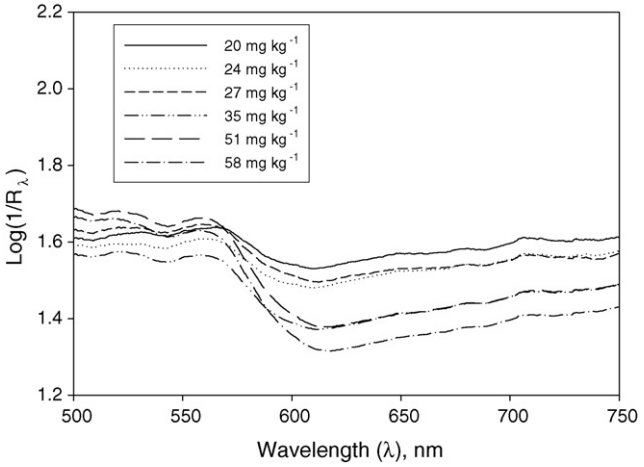


**Fig. 2.** Log response of visible reflectance spectra of watermelon puree samples with various lycopene concentrations.

PLS models were evaluated by coefficient of determination, $R^2$, and standard errors of calibration (SEC) for the calibration sample set. The SEC for PLS regression model was calculated as follows:

$$\mathrm{SEC} = \sqrt{\frac{\sum_{i=1}^{n}(Y_{\mathrm{actual}} - Y_{\mathrm{predicted}})^2}{n-f-1}} \quad (4)$$

where $n$ is the number of samples, $f$ is the number of principal components, $Y_{\mathrm{actual}}$ is the actual values of lycopene concentration, and $Y_{\mathrm{predicted}}$ is the predicted lycopene concentration by the model.

Prediction ability of PLS models were evaluated by $R^2$, SEP (Eq. (3)), and bias from the validation data set. The bias values were obtained as the mean of the residuals.

## 3. Results and discussion

### 3.1. Reflectance spectra of watermelon and tomato puree

The reflectance spectra of watermelon and tomato puree are shown in Figs. 2 and 3, respectively. The major peaks at 565 nm were due to the high apparent absorbance by lycopene in the aqueous environment of fruit tissue (Choudhary et al., 2004; Davis et al., 2003a,b). The lycopene has an absorbance peak at 503 nm in hex-

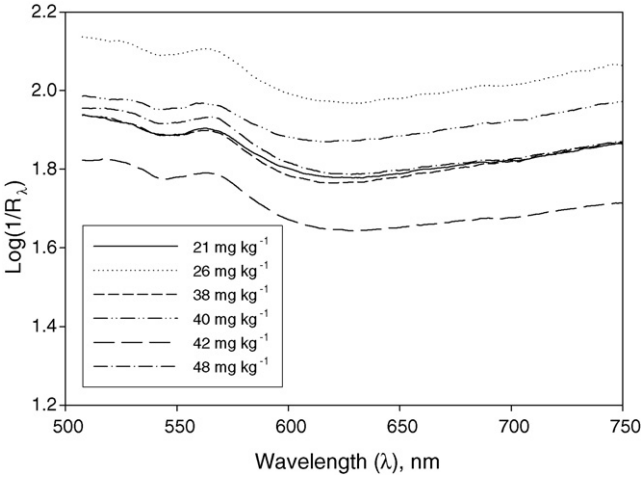


**Fig. 3.** Log response of visible reflectance spectra of tomato puree samples with various lycopene concentrations.

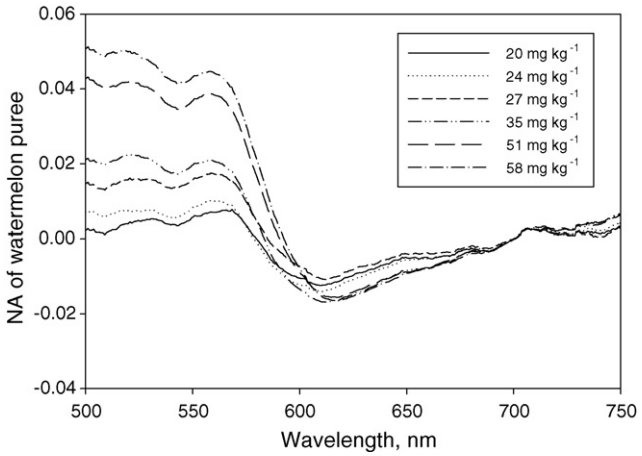


**Fig. 4.** Normalized reflectance spectra of watermelon puree samples with various lycopene concentrations (normalized apparent absorbance versus wavelength).

ane, and a peak of 565 nm when surrounded by the polar molecules of water. The spectra in Figs. 2 and 3 were normalized and baseline corrected by calculating normalized apparent absorbance values, NAA, at each wavelength of spectra, using the following equation:

$$\mathrm{NAA} = \frac{\mathrm{AA}_\lambda - \mathrm{AA}_{700}}{\mathrm{AA}_\lambda + \mathrm{AA}_{700}} \quad (5)$$

where $\mathrm{AA}_\lambda$ is the apparent absorbance at wavelength $\lambda$ (nm), $\mathrm{AA}_{700}$ is the apparent absorbance at 700 nm. The waveband at 700 nm was chosen as a reference band for base line correction and normalization as it is in the red band and could represent redness of lycopene in watermelon and tomato samples. The normalized spectra of watermelon and tomato puree samples are shown in Figs. 4 and 5, respectively. By normalizing the spectra, baseline shifts in the spectra of individual samples were removed and peak heights at 565 nm in the normalized spectra gave the values of NAAI.

### 3.2. LS regression of watermelon and tomato puree reflectance

Apparent absorbance index and normalized apparent absorbance index from reflectance spectra of watermelon and tomato puree were correlated with lycopene concentration. As shown in Figs. 6 and 7, both the AAI and NAAI had a high linear correlation with lycopene concentration of watermelon puree ($R^2$ = 0.90 for both AAI and NAAI) and tomato puree ($R^2$ = 0.62 for

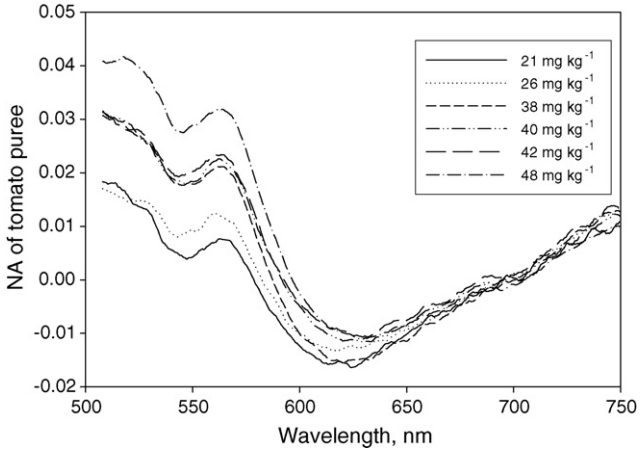


**Fig. 5.** Normalized reflectance spectra of tomato puree samples with various lycopene concentrations (normalized apparent absorbance versus wavelength).

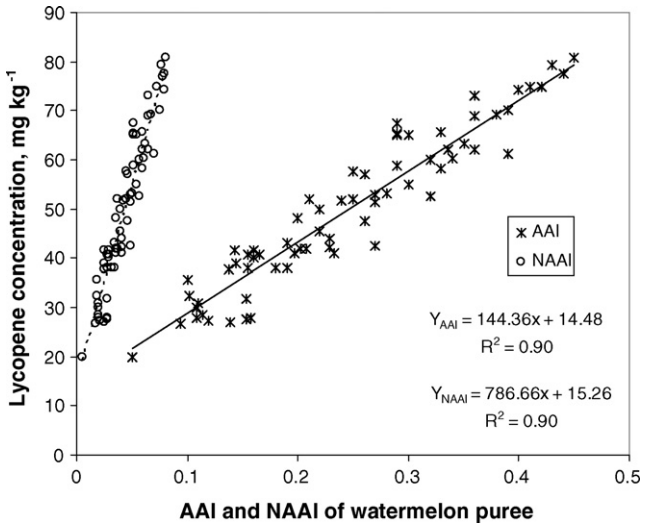


**Fig. 6.** Correlation of apparent absorbance index (AAI) and normalized apparent absorbance index (NAAI) with lycopene concentration of watermelon puree.

AAI and 0.61 for NAAI). In Fig. 7, the negative apparent absorbance index of tomato puree samples could be due to higher scattering of light at 700 nm caused by higher particulate concentration, which included tomato seeds; and lower absorbance at 565 nm due to the comparatively lower lycopene concentration of tomato puree samples. This lower value of lycopene/particulate ratio in tomato samples compared to watermelon samples might have caused lower correlation coefficient for tomato samples. For both watermelon and tomato puree, the slopes of least squares models for NAAI was higher than that for the AAI due to normalization of AAI. For AAI models, the regression coefficient (slope of the linear model) for tomato puree was higher than that of watermelon puree but there was no significant difference between the slopes of NAAI models of both the fruits ($P$ = 0.0941). Hence the NAAI regression model can be used for prediction of lycopene in either fruit types. The validation of least squares regression models for prediction of lycopene from AAI of watermelon and tomato are shown in Figs. 8 and 9. The LS linear regression model for watermelon puree could predict lycopene concentration in validation samples with $R^2$ = 0.93, SEP = 5.1 mg kg$^{-1}$ and bias = 2.1 mg kg$^{-1}$ (Fig. 8). The least

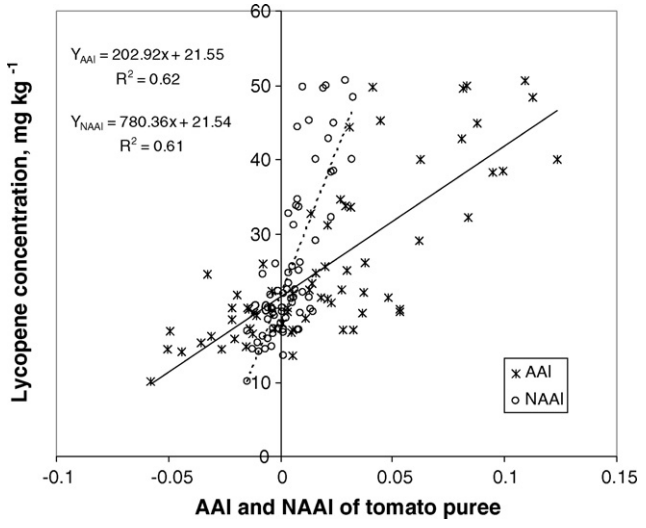


**Fig. 7.** Correlation of apparent absorbance index (AAI) and normalized apparent absorbance index (NAAI) with lycopene concentration of tomato puree.

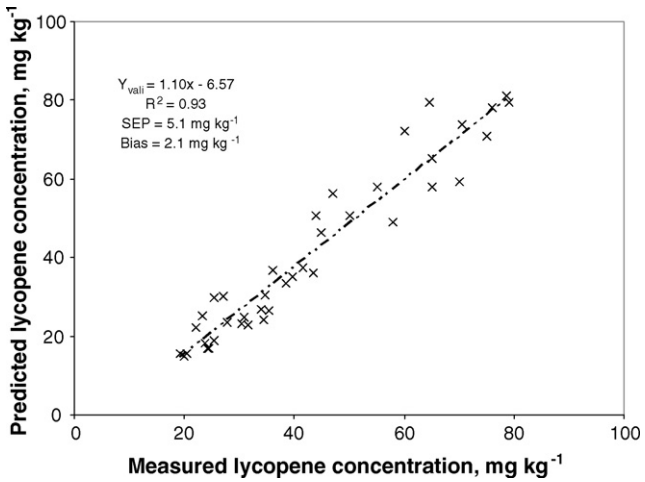


**Fig. 8.** Prediction of lycopene concentration from least squares models of AAI for watermelon puree samples.

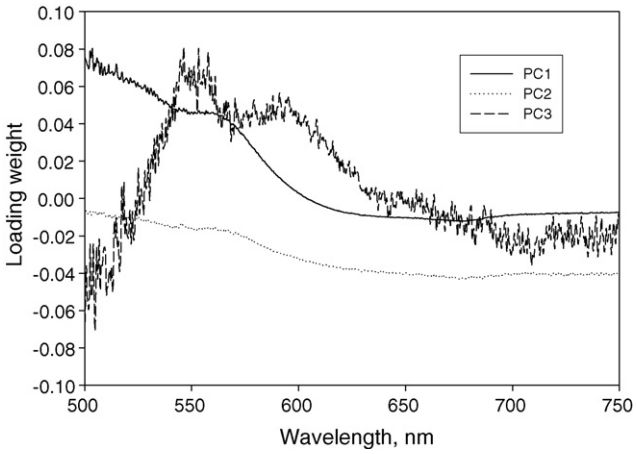


**Fig. 10.** Loading plot of three latent variables (principal components) of PLS regression model used for tomato puree samples.

squares model for tomato puree predicted lycopene concentration in tomato puree validation samples with an $R^2$ value of 0.54, SEP of 5.2 mg kg$^{-1}$ and bias of −3.3 mg kg$^{-1}$ (Fig. 9).

### 3.3. PLS regression for watermelon and tomato puree reflectance

PLS regression models were developed for watermelon and tomato puree samples using all the wavelengths in the spectrum range of 500–750 nm. By minimizing predicted residual error sum of squares, three latent variables (known as factors or principal components) were found to be optimal, and hence used for PLS prediction model. Fig. 10 shows a loading plot of first three principal components of tomato puree samples. The first principal component (PC1) which explained 74% variation in data had high loading values for wavelengths lower than 600 nm, with a peak at 565 nm. This shows the significance of 565 nm for prediction of lycopene concentration. The PC1 loading values for 650 nm and higher wavelengths were very small and almost flat. Absorbance values at 700 nm were thus useful for scatter correction in absorbance index calculation. The PC2, which explained 13% variation in data, had lower loading values than the PC1, without any noticeable peak. The third latent variable, PC3, which explained 10% variance, had loading plot with peaks at about 550 and 600 nm, and a valley at 565 nm. The high loading values of PC3 in the wavelength range of 525–625 nm shows the importance of these bands in lycopene sensing. Since PLS regression model included all wavelengths including these bands, the PLS model provided higher precision of lycopene prediction than the least squares regression models. The PLS model for watermelon puree with three principal components predicted lycopene concentrations in validation samples with an $R^2$ value of 0.97, and a SEP of 3.4 mg kg$^{-1}$, with zero bias (Fig. 11). The PLS model for tomato puree with three principal components predicted lycopene concentration in validation samples with an $R^2$ value of 0.88, an SEP of 2.5 mg kg$^{-1}$ and a bias of 0.5 mg kg$^{-1}$ (Fig. 12). Compared to the least squares models, the PLS models had better prediction ability with less error. This was possible because the PLS algorithm has the capability to capture the relevant significant variance in the independent variables in the first few important principal components (factors) and send the noise values to the later factors (Williams and Norris, 2001).

The prediction error of the PLS model for watermelon puree was nearly equal to the precision of the hexane extraction method (±3.0 mg kg$^{-1}$). The fiber optic sensing method tested in this study can be used for rapid sensing of lycopene concentration in tomato and watermelon puree for their rapid quality assessment before their use as raw materials for manufacturing of lycopene rich food and nutraceuticals products.

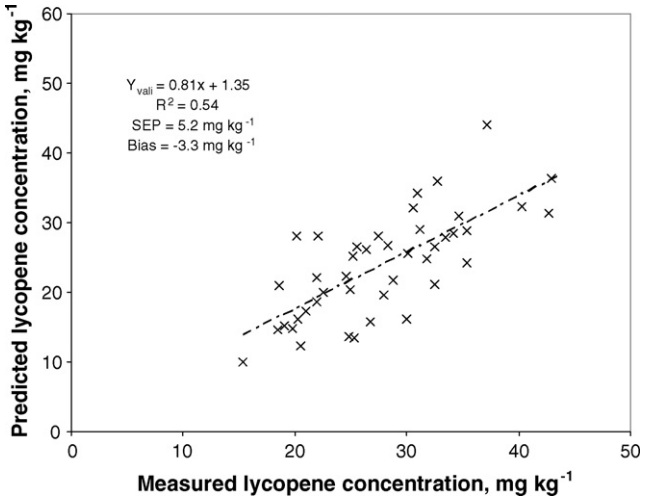


**Fig. 9.** Prediction of lycopene concentration from least squares models of AAI for tomato puree samples.

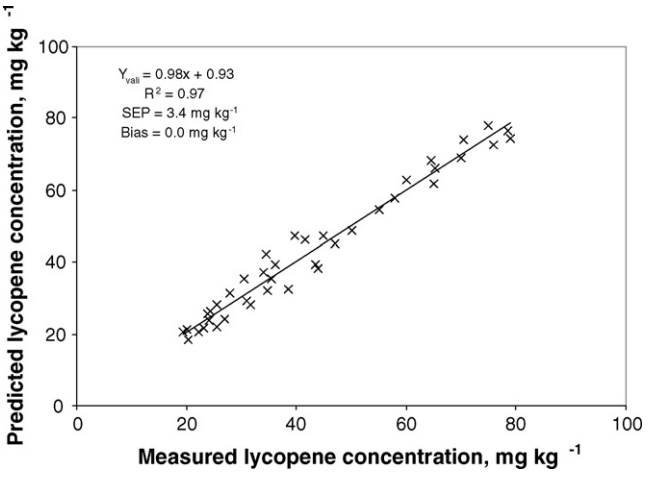


**Fig. 11.** Performance of PLS regression model of watermelon puree samples.

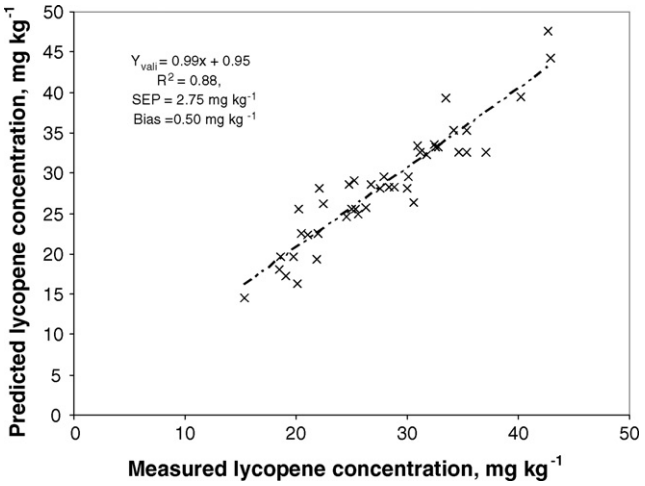


**Fig. 12.** Performance of PLS regression model of tomato puree samples.

## 4. Conclusions

The apparent absorbance index and normalized apparent absorbance index obtained from apparent absorbance values at 565 and 700 nm from reflectance spectra of watermelon and tomato purees were linearly correlated with the lycopene concentration of watermelon and tomato puree samples ($R^2$ = 0.90 for watermelon and 0.62 for tomato). This is informative for the design of a fiber optic reflectance sensor for estimating lycopene concentration of watermelon and tomato puree, as only two monochromatic light sources of 565 and 700 nm can be used for lycopene sensing. The PLS models developed by using all apparent absorbance values from reflectance spectra in the wavelength range of 500–750 nm performed well during validation with the validation sample sets. The PLS model for watermelon puree predicted lycopene concentrations in validation samples with an $R^2$ value of 0.97, an SEP of 3.4 mg kg$^{-1}$. The PLS model for tomato puree predicted lycopene concentration in validation samples with an $R^2$ value of 0.88, an SEP of 2.5 mg kg$^{-1}$ and a bias of 0.5 mg kg$^{-1}$. Based on these results, it can be concluded that fiber optic visible reflectance spectroscopy can be used reliably for rapid and safe quantification of lycopene in watermelon and tomato puree samples.

## Acknowledgements

Research facilities were graciously provided by the Division of Agricultural Sciences and Natural Resources, Oklahoma State University, Stillwater through the Biosystems and Agricultural Engineering Department, Oklahoma Food and Agricultural Research and Technology Center, and the Department of Horticulture and Landscape Architecture. Instrumentation facilities were provided by Dr. Marvin Stone. Technical assistance of Mr. Joannis Oikonomakos and Ms. Donna Chrz is appreciated. Dr. Glenn Kranzler and Dr. Jeyamkondan Subbiah provided multivariate data analyses facilities.

## References

Adsule, P.G., Dan, A., 1979. Simplified extraction procedure in the rapid spectrophotometric method for lycopene estimation in tomato. J. Food Sci. Technol. 16, 216–218.
Arab, L., Steck, S., 2000. Lycopene and cardiovascular disease. Am. J. Clin. Nutr. 71, 1691S–1695S.
Arias, R., Lee, T.C., Logendra, L., Janes, H., 2000. Correlation of lycopene measured by HPLC with the L*, a*, b* color readings of a hydroponic tomato and the relationship of maturity with color and lycopene content. J. Agric. Food Chem. 48, 1697–1702.
Beerh, O.P., Siddappa, G.S., 1959. A rapid spectroscopic method for the detection and estimation of adulterants in tomato ketchup. Food Technol. 13, 414–418.
Cadoni, E., De Giorgi, M.R., Medda, E., Poma, G., 2000. Supercritical $CO_2$ extraction of lycopene and β-carotene from ripe tomatoes. Dyes Pigments 44, 27–32.
Choudhary, R., Bowser, T.J., Weckler, P., Maness, N.O., McGlynn, W., Stone, M.L., 2004. Rapid sensing of lycopene in watermelon and tomato by visible reflectance spectroscopy. In: Proceedings of the Research Symposium, Food and Agricultural Products Research and Technology Center. Oklahoma State University, Stillwater, OK, pp. 21–22.
D'Souza, M., Singh, S., Ingle, M., 1992. Lycopene concentration of tomato fruit can be estimated from chromaticity values. Hortscience 27 (5), 465–466.
Davis, A.R., Fish, W.W., Perkins-Veazie, P., 2003a. A rapid hexane-free method for analyzing lycopene content in watermelon. J. Food Sci. 68 (1), 328–332.
Davis, A.R., Fish, W.W., Perkins-Veazie, P., 2003b. A rapid spectrophotometric method for analyzing lycopene content in tomato and tomato products. Postharvest Biol. Technol. 28, 425–430.
Di Mascio, P., Kaiser, S.P., Sies, H., 1989. Lycopene as the most efficient biological singlet oxygen quencher. Arch. Biochem. Biophys. 274, 532–538.
Fish, W.W., Perkins-Veazie, P., Collins, J.K., 2002. A quantitative assay for lycopene that utilizes reduced volume of organic solvents. J. Food Comp. Anal. 15, 309–317.
Hadley, C.W., Miller, E.C., Schwartz, S.J., Clinton, S.K., 2002. Tomatoes, lycopene, and prostate cancer: progress and promise. Exp. Biol. Med. 227, 869–880.
Halim, Y., Schwartz, S.J., Francis, D., Baldauf, N.A., Rodriguez-Saona, L.E., 2006. Direct determination of lycopene content in tomatoes (*Lycopersicon esculentum*) by attenuated total reflectance infrared spectroscopy and multivariate analysis. J. AOAC Int. 89 (5), 1257–1262.
Heber, D., Lu, Q.Y., 2002. Overview of mechanisms of action of lycopene. Exp. Biol. Med. 227, 920–923.
Hruschka, W.R., 2001. Data analysis: wavelength selection methods. In: Williams, P., Norris, K. (Eds.), Near Infrared Technology in the Agricultural and Food Industries. American Association of Cereal Chemists Inc., St. Paul, MN, pp. 39–58.
Kusumiyati, Akinaga, T., Tanaka, M., Kawasaki, S., 2008. On-tree and after-harvesting evaluation of firmness, color and lycopene content of tomato fruit using portable NIR spectroscopy. J. Food Agric. Environ. 6 (2), 327–332.
Mangels, A.R., Holden, J.M., Beecher, G.R., Forman, M.R., Lanza, E., 1993. Carotenoid content of fruits and vegetables: an evaluation of analytic data. J. Am. Diet. Assoc. 93, 284–296.
Perkins Veazie, P.M., Collins, J.K., Pair, S.D., Roberts, W., 2001. Lycopene content differs among red-fleshed watermelon cultivars. J. Sci. Food Agric. 81, 983–987.
Perkins Veazie, P.M., Roberts, W., Collins, J.K., 2006. Lycopene content among organically produced tomatoes. J. Veg. Sci. 12 (4), 93–106.
Sadler, G., Davis, J., Dezman, D., 1990. Rapid extraction of lycopene and ß-carotene from reconstituted tomato paste and pink grapefruit homogenates. J. Food Sci. 55 (5), 1460–1461.
Scott, K.J., Hart, D.J., 1995. Development and evaluation of an HPLC method for the analysis of carotenoids in foods, and the measurement of the carotenoid content of vegetables and fruits commonly consumed in the UK. Food Chem. 54, 101–111.
Schulz, H., Schutze, W., Baranska, M., Baranska, M., 2006. Fast determination of carotenoids in tomatoes and tomato products by Raman spectroscopy. Acta Horticult. 712 (2), 901–905.
Tinker, J.H., Bohm, F., Schalch, W., Truscott, T.G., 1994. Dietary carotenoids protect human cells from damage. J. Photochem. Photobiol. B 26, 283–285.
Tomes, M.L., Johnson, K.W., Hess, M., 1963. The carotene pigment content of certain red fleshed watermelons. Proc. Am. Soc. Hortcult. Sci. 82, 460–464.
USDA, 2007. USDA National Nutrient Database for Standard Reference, Release 20. Available at: http://www.nal.usda.gov/fnic/foodcomp/Data/SR20/nutrlist/sr20w337.pdf (Accessed August 29, 2008).
Williams, P., Norris, K., 2001. Near Infrared Technology in the Agricultural and Food Industries. American Association of Cereal Chemists Inc., St. Paul, MN.