

بازشناسی الگوی اسپکتروسکوپی فرو سرخ نزدیک برای تفکیک غیر مخرب پرتقال‌ها بر اساس شاخص مزه

بهاره جمشیدی^{۱*}، سعید مینایی^۲، عزالدین مهاجرانی^۳ و حسن قاسمیان^۴

۱- دانش‌آموخته دکتری تخصصی، دانشگاه تربیت مدرس (استادیار، موسسه تحقیقات فنی و مهندسی کشاورزی)

b.jamshidi@aeri.ir

۲- استاد، دانشکده کشاورزی، دانشگاه تربیت مدرس

۳- استاد، پژوهشکده لیزر و پلاسما، دانشگاه شهید بهشتی

۴- استاد، دانشکده مهندسی برق و کامپیوتر، دانشگاه تربیت مدرس

چکیده

در سال‌های اخیر، کاربرد اسپکتروسکوپی فروسرخ نزدیک (NIR) به عنوان یک روش غیر مخرب در ترکیب با روش‌های شیمی‌سنجی به منظور ارزیابی کیفیت محصولات کشاورزی و غذایی به طور گسترده‌ای مورد توجه قرار گرفته است. در روش‌های شیمی‌سنجی، آنالیزهای کیفی موضوعات مهمی هستند که می‌توانند به مسئله بازشناسی الگو نسبت داده شوند. در این پژوهش، توانایی روش‌های بازشناسی الگو در ترکیب با اسپکتروسکوپی NIR بازتابی به منظور تفکیک غیر مخرب مزه پرتقال‌ها بررسی شد. برای این منظور، روش‌های بازشناسی الگوی نظارت‌نشده و نظارت‌شده، خوشه‌بندی سلسله مراتبی (HCA) و مدل‌سازی مستقل نرم شباهت‌های طبقه (SIMCA)، به ترتیب برای امکان‌سنجی تفکیک واریته‌های پرتقال و طبقه‌بندی (بر اساس مزه آنها) بر پایه اطلاعات طیفی محدوده ۱۶۵۰-۹۳۰ nm استفاده شدند. آنالیزهای کیفی نشان داد که طیف‌های NIR واریته‌های پرتقال به درستی با بازشناسی الگوی نظارت‌نشده HCA خوشه‌بندی شدند. همچنین، بازشناسی الگوی نظارت‌شده SIMCA برای طیف‌های NIR پرتقال‌ها نتایج عالی طبقه‌بندی واریته بر اساس شاخص Brima را در سطح احتمال ۵ درصد در بر داشت (دقت طبقه‌بندی ۹۸/۵۷ درصد). همچنین، طول موج‌های ۱۰۴۷/۵ nm، ۱۵۰۲ nm و ۱۴۷۵ nm سهم بیش‌تری نسبت به سایر طول موج‌ها در تفکیک دو طبقه بر عهده داشتند. نمونه‌های دارای مقادیرهای یکسان شاخص Brima نیز به درستی و با دقت طبقه‌بندی بالا (۹۵/۴۵ درصد) در سطح احتمال ۵ درصد طبقه‌بندی شدند. قدرت تفکیک طول موج‌های ۱۴۷۵ nm، ۱۵۸۳ nm و ۱۴۳۶/۷۵ nm برای دستیابی به این طبقه‌بندی بسیار بیش‌تر از سایر طول موج‌ها بود. بنابراین، اسپکتروسکوپی NIR بازتابی در ترکیب با روش‌های بازشناسی الگو می‌تواند برای تشخیص سایر پارامترهای مرتبط با مزه نیز به کار رود.

واژه‌های کلیدی

اسپکتروسکوپی فروسرخ نزدیک، بازشناسی الگو، طبقه‌بندی، غیر مخرب، مزه

مقدمه

از بیش از دو دهه گذشته، اسپکتروسکوپی^۱ فرورسرخ نزدیک^۲ (NIR) به عنوان یک روش تحلیلی و عملی غیر مخرب در گستره پهناوری از علم و صنعت بکار گرفته شده و پیشرفت‌های اخیر تکنولوژی در طراحی سخت‌افزار و روش‌های جمع‌آوری، پردازش و تحلیل داده، توانمندی آن را دو چندان نموده است. کاربرد این روش در آنالیزهای غذایی به ویژه به منظور ارزیابی کیفیت درونی میوه‌ها و سبزی‌ها نیز بسیار مورد توجه پژوهشگران و صنعتگران قرار گرفته است (Jamshidi *et al.*, 2011b, 2012a, 2013a).

طیف NIR یک ترکیب می‌تواند اطلاعات مهمی در مورد ساختار مولکولی و طبیعت شیمیایی ماده ارائه کند؛ با این حال همیشه استخراج داده‌ها به طور مستقیم از طیف ممکن نیست. هم‌چنین، طیف‌های NIR ممکن است به دلیل اثر پراکندگی، ناهمگن بودن بافت، نویز دستگاهی، اثرهای محیطی و غیره پیچیده‌تر شوند. بنابراین، روش‌های آماری چندمتغیره^۳ (شیمی‌سنجی^۴) برای استخراج اطلاعات مفید از یک طیف NIR نیاز هستند (Nicolai *et al.*, 2007).

در سال‌های اخیر، تلفیق روش غیر مخرب اسپکتروسکوپی NIR با روش‌های چندمتغیره به طور گسترده‌ای در آنالیزهای کمی و تعیین ویژگی‌های درونی میوه‌ها و سبزی‌ها به کار گرفته شده است (Flores *et al.*, 2009; Jamshidi *et al.*, 2012c; Moghimi *et al.*, 2010; Penchaiya *et al.*, 2009; Pissard *et al.*, 2013; Subedi and Walsh, 2011). آنالیزهای کیفی که می‌توانند به مسئله بازشناسی الگو نسبت داده شوند نیز به اندازه آنالیزهای کمی در اسپکتروسکوپی NIR میوه‌ها و سبزی‌ها دارای اهمیت هستند (Cen and He, 2007). تاکنون، کاربرد این آنالیزها بر پایه اسپکتروسکوپی NIR به منظور تشخیص کیفی، تفکیک و طبقه‌بندی میوه‌ها بر اساس یک یا چند ویژگی خاص از قبیل وارپته (Cao *et al.*, 2010; Luo *et al.*, 2013); Jamshidi *et al.*, 2012b; Luo *et al.*, 2011; Suphamitmongkol *et al.*, 2013); آسب‌دیدگی (Luo *et al.*, 2013); شدت رنگ (Camps and Christen, 2009); مرحله رسیدگی (Mireei *et al.*, 2010); و کیفیت پس از انبارداری (Manley *et al.*, 2007) مورد توجه قرار گرفته است.

یک پژوهش مروری به بررسی تحقیقات انجام‌شده تاکنون در زمینه کاربرد اسپکتروسکوپی NIR در آنالیزهای کمی و کیفی مرکبات پرداخته است (Magwaza *et al.*, 2012). با این حال و با توجه به این که کاربرد روش مذکور در آنالیزهای کمی و کیفی میوه‌های با پوست ناصاف و ضخیم مانند مرکبات به مراتب مشکل‌تر و پیچیده‌تر از کاربرد آن در میوه‌های با پوست صاف و نازک است (Jamshidi *et al.*, 2011a, 2013b)، پژوهش‌های پیش‌تر در این زمینه به منظور دستیابی به نتایج قابل اعتماد با به‌کارگیری روش‌های مختلف شیمی‌سنجی نیاز است. از سوی دیگر، مزه یکی از ویژگی‌های مهم میوه‌ها و سبزی‌هاست که تاثیر

1 - Spectroscopy
2 - Near-infrared
3 - Multivariate
4 - Chemometrics



بسیاری در تعیین کیفیت، بازارپسندی و جلب رضایت مشتری چه برای تازه‌خوری و چه در صنعت فراوری دارد. از این رو در این پژوهش، از ترکیب روش‌های بازشناسی الگوی نظارت‌شده و نظارت‌نشده با اسپکتروسکوپی NIR در محدوده طیفی ۹۳۰-۱۶۵۰ nm برای اولین بار به منظور بررسی امکان تفکیک و طبقه‌بندی غیر مخرب واریته‌های پرتقال (والنسیا و تامسون) بر اساس شاخص مزه مبتنی بر شیرینی و ترشی میوه (BrimA)، همچنین سایر پارامترهای مربوط به مزه استفاده شد.

مواد و روش‌ها

اسپکتروسکوپی NIR بازتابی از ۳۰۰ نمونه پرتقال شامل دو واریته والنسیا و تامسون (۱۵۰ عدد از هر واریته) با استفاده از چیدمان مناسب طراحی‌شده (Jamshidi *et al.*, 2012c) و با یک اسپکترومتر مدل EPP2000-NIR دارای قدرت تفکیک ۲/۵ nm و مجهز به آشکارساز آرایه فتودیود^۱ (PDA) انجام و طیف‌ها در محدوده طیفی ۹۳۰-۱۶۵۰ nm ذخیره شدند. به منظور حذف اطلاعات ناخواسته پس‌زمینه و نویزهای تصادفی موجود در طیف‌ها، پس از تبدیل آنها به طیف‌های جذبی (Log 1/R) از یک فیلتر میانگین‌گیر متحرک^۲ (MA) با عرض پنجره ۵ استفاده شد. نرمال‌سازی طیف‌ها نیز برای حذف اثرات جمعی و افزایش طیفی ناشی از پخش غیر یکنواخت نور، اندازه نمونه، و تغییر فاصله نوری به روش تصحیح پخش افزایشده^۳ (MSC) یا خطی‌سازی طیف هر نمونه به طیف میانگین نمونه‌ها انجام شد (Fu *et al.*, 2007).

پس از انجام اندازه‌گیری‌های طیفی، نمونه‌های پرتقال آب‌گیری و میزان مواد جامد حل‌شده^۴ (SSC) آنها (معرف مزه شیرین پرتقال) با یک دستگاه رفرکتومتر مدل DR-A1 و مقدار اسیدیته قابل تیتر^۵ (TA) آنها بر حسب اسید سیتریک (معرف مزه ترش پرتقال) به روش تیتراسیون با سود ۰/۱ نرمال اندازه‌گیری شدند. سپس، شاخص برتر مزه ترکیبی پرتقال (BrimA) مبتنی بر پارامترهای SSC و TA (که معرف مزه واقعی میوه است) از رابطه (۱) به‌دست آمد (Jordan *et al.*, 2001):

$$\text{BrimA} = \text{SSC} - k \times (\text{TA}) \quad (1)$$

در این رابطه، k ثابت حساسیت زبان^۶ است. مقدار این ثابت برابر ۴ در نظر گرفته شد که بیش‌ترین هم‌بستگی بین BrimA و امتیاز هیدونیک داورهای چشایی را برای پرتقال سبب می‌شود (Obenland *et al.*, 2009).

آنالیزهای کیفی به منظور تفکیک و طبقه‌بندی واریته‌های پرتقال ابتدا بر اساس مزه (شاخص BrimA) و سپس بر اساس سایر پارامترهای مربوط به مزه، بر پایه ترکیب روش‌های بازشناسی الگو با داده‌های به‌دست آمده از آزمون‌های اسپکتروسکوپی انجام شد. در این راستا، نخست از یک روش بازشناسی الگوی نظارت‌نشده، آنالیز خوشه سلسله مراتبی^۷ (HCA)، برای بررسی امکان خوشه‌بندی واریته‌ها بر پایه ویژگی‌های طیفی آنها استفاده شد. به این ترتیب که، پس از حذف داده‌های پرت (برای هر واریته) و

- 1 - Photodiode Array
- 2 - Moving Average
- 3 - Multiplicative Scatter Correction
- 4 - Soluble Solids Content
- 5 - Titratable Acidity
- 6 - Tongue's Sensitivity Constant
- 7 - Hierarchical Cluster Analysis



اِعمال پیش‌پردازش ترکیبی MA + MSC بر روی طیف‌های NIR، خوشه‌بندی بر اساس نزدیک‌ترین همسایگی یا کوتاه‌ترین فاصله اقلیدسی^۱ انجام گرفت. برای این منظور، ابتدا فرض شد هر نمونه به تنهایی یک خوشه است؛ سپس نمونه‌ها به طور متوالی به هم پیوستند تا مطابق با فاصله بین آنها، خوشه‌ها تشکیل شوند. به عبارت دیگر، ابتدا نزدیک‌ترین دو نمونه برای تشکیل یک خوشه به هم پیوستند و در مراحل سلسله مراتبی بعدی، بقیه نمونه‌ها خوشه‌بندی شدند تا یک تک‌گروه تشکیل شده از همه نمونه‌های آنالیزشده، شکل بگیرد. در ادامه آنالیزهای کیفی، از یک روش بازشناسی الگوی نظارت‌شده، مدل‌سازی مستقل نرم شباهت‌های طبقه^۲ (SIMCA)، برای طبقه‌بندی واریته‌های پرتقال ابتدا بر اساس مزه (شاخص BrimA) استفاده شد. طبق این روش، یک نمونه زمانی به یک طبقه اختصاص داده شد که به اندازه کافی با دیگر اعضای آن طبقه شباهت داشت. برای این منظور پس از حذف داده‌های پرت، نمونه‌های هر طبقه (واریته) به دسته‌های آموزش (۷۵ درصد نمونه‌ها) و آزمون (۲۵ درصد باقی‌مانده) تقسیم شدند. مدل‌سازی دسته آموزش یا واسنجی هر طبقه (واریته) بعد از اِعمال پیش‌پردازش ترکیبی MA + MSC با روش حداقل مربعات جزئی^۳ (PLS) برای متغیر وابسته BrimA انجام شد. سپس، مدل‌های PLS تدوین شده برای هر طبقه با روش اعتبارسنجی متقاطع حذف تک‌نمونه^۴ ارزیابی شدند. پس از اعتبارسنجی مدل‌های واسنجی هر طبقه، تعداد بهینه متغیرهای نهان^۵ (LVs) برای هر طبقه بر پایه داشتن کم‌ترین ریشه میانگین مربعات خطای اعتبارسنجی متقاطع^۶ (RMSECV) مشخص و در نهایت، توانایی مدل طبقه‌بندی SIMCA برای تفکیک نمونه‌های دسته آزمون، در قالب دقت طبقه‌بندی درست (صحت) و اعتبار طبقه‌بندی بررسی شد. به این ترتیب که، فاصله هر نمونه ناشناخته از هر مدل تعیین و نمونه به طبقه‌ای تعلق یافت که فاصله آن از مدل طبقه، زیر مقدار بحرانی آن باشد (Varmuza and Filzmoser, 2009). هم‌چنین، توانایی طول موج‌های مختلف محدوده طیفی ۹۳۰-۱۶۵۰ nm در تفکیک پرتقال‌ها بر اساس مزه (شاخص BrimA) با بررسی انحراف استاندارد طول موج‌ها هنگام برازش نمونه‌ها از مدل یک طبقه روی مدل طبقه دیگر، ارزیابی شد. در مرحله بعد، طبقه‌بندی SIMCA تنها برای نمونه‌های والنسیا و تامسون دارای مزه یکسان بر پایه شاخص BrimA ($4 < \text{BrimA} < 6$) انجام شد. ایده اصلی در این مرحله این بود که اگر امکان طبقه‌بندی واریته‌های مختلف پرتقال (والنسیا و تامسون) با مزه یکسان بر اساس شاخص BrimA و بر پایه طیف‌های NIR بازتابی آنها وجود داشته باشد، روش اسپکتروسکوپی NIR قادر به تشخیص غیر مخرب دیگر پارامترهای مربوط به مزه پرتقال (مانند اسیدهای آمینه) نیز است. بنابراین پس از حذف داده‌های پرت، تقسیم‌بندی نمونه‌ها به دسته‌های آموزش و آزمون برای هر طبقه (در مجموعه جدید نمونه‌ها)، و پیش‌پردازش ترکیبی MA + MSC طیف‌ها انجام شد.

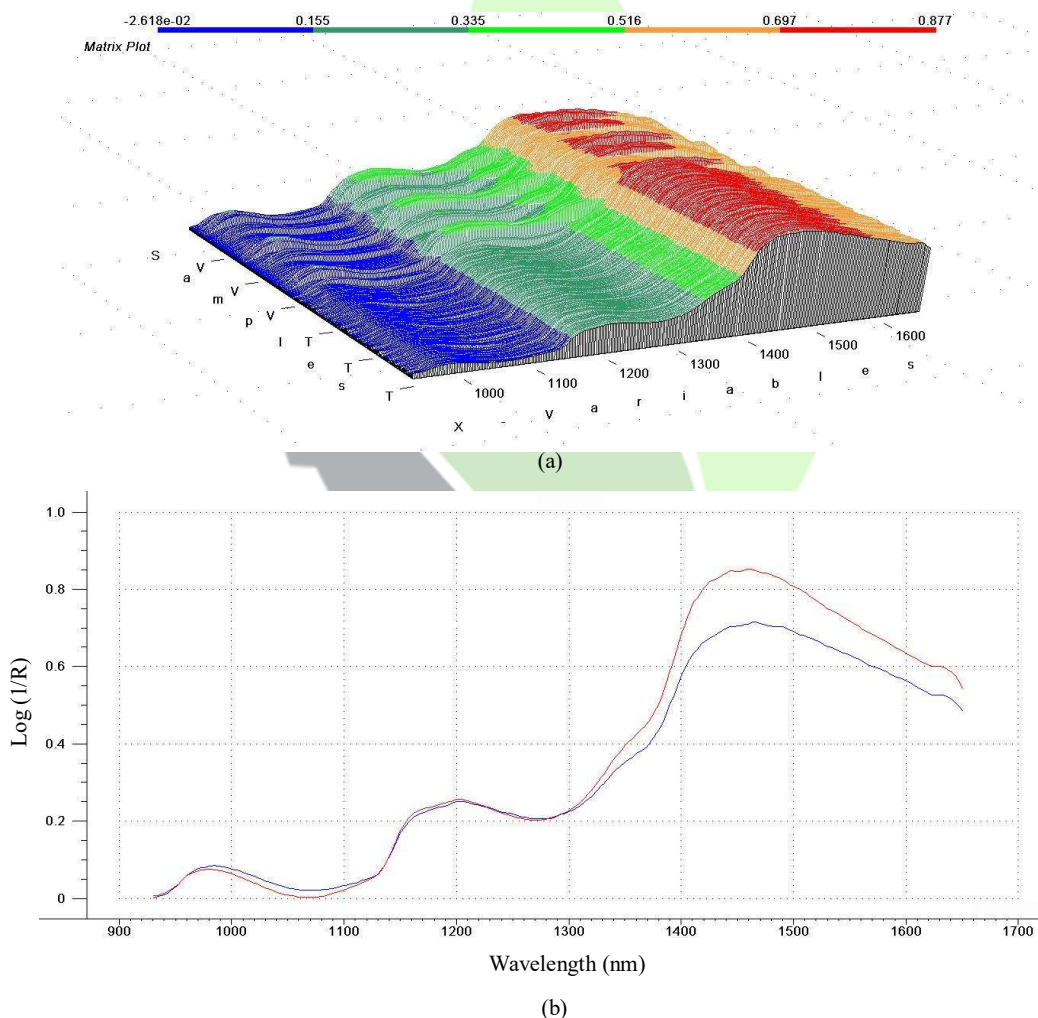
-
- 1 - Euclidean Distance
 - 2 - Soft Independent Modeling of Class Analogies
 - 3 - Partial Least Squares
 - 4 - Leave-one-out Cross Validation
 - 5 - Latent Variables
 - 6 - Root Mean Square Error of Cross Validation



مدل‌سازی نمونه‌های آموزش هر طبقه به صورت مجزا و بر پایه آنالیز مولفه‌های اصلی^۱ (PCA) انجام و تعداد بهینه مولفه‌های اصلی (PCs) برای هر طبقه نیز بر اساس سازش متوازن بین نیاز به تشریح واریانس در داده‌های اصلی و نیاز به عدم بیش‌برازش مدل مشخص شد. در نهایت، توانایی مدل طبقه‌بندی SIMCA برای تفکیک نمونه‌های دسته آزمون (در قالب دقت طبقه‌بندی درست و اعتبار طبقه‌بندی)، همچنین توانایی طول موج‌های مختلف در تفکیک پرتقال‌های دارای مزه یکسان بر اساس شاخص BrimA بررسی شد. برای اجرای آنالیزها از نرم‌افزار Unscrambler X 10.1 (CAMO Software AS, Norway) استفاده شد.

نتایج و بحث

شکل ۱ طرح ماتریسی طیف‌های جذبی NIR (در محدوده طیفی ۹۳۰-۱۶۵۰ nm) همه نمونه‌های پرتقال شامل والنسیا (V) و تامسون (T) و میانگین طیف‌های جذبی دو واریته را نسبت به یکدیگر نشان می‌دهد.

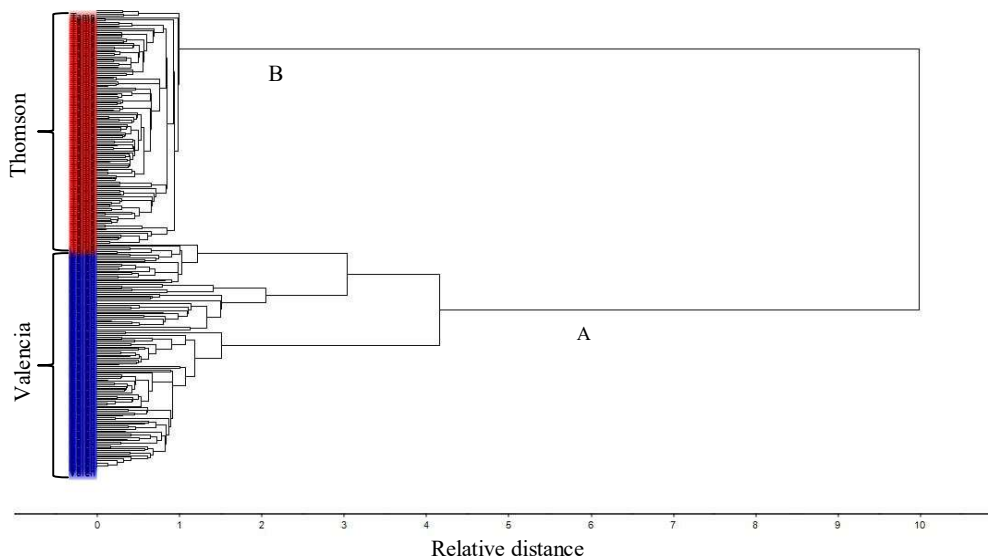


شکل ۱- (a) طرح ماتریسی، (b) میانگین طیف‌های جذبی NIR واریته‌های پرتقال
Fig.1. (a) Matrix plot, (b) the mean of absorbance NIR spectra of orange varieties



مطابق شکل، طیف‌های هر دو وارسته شامل چند پیک جذبی در فرکانس‌های خاص به دلیل ارتعاش‌های کششی اورتون‌های گروه‌های عاملی C-H, O-H، و (یا) N-H متناسب با غلظت بعضی ترکیبات شیمیایی درونی میوه شامل این گروه‌های عاملی مانند قند و اسید بودند (Jamshidi *et al.*, 2013a). این نتیجه با یافته‌های سایر پژوهشگران هم‌خوانی داشت (Cayuela and Weiland, 2010; Magwaza *et al.*, 2012; Suphamitmongkol *et al.*, 2013)

شکل ۲ ساختار درختی سلسله مراتبی به دست آمده از HCA نمونه‌های پرتقال را بر پایه طیف‌های NIR (در محدوده طیفی ۱۶۵۰-۹۳۰ nm)، که بر اساس نزدیک‌ترین همسایگی یا کوتاه‌ترین فاصله اقلیدسی شکل گرفته‌اند، نشان می‌دهد. مطابق شکل، دو خوشه کاملاً مجزای A و B به ترتیب مربوط به وارسته‌های والنسیا و تامسون، به خوبی قابل تشخیص و تفکیک هستند. همچنین طبق این ساختار، هر وارسته پرتقال به خوشه‌های متراکم‌تری که از نمونه‌های همان وارسته شکل گرفته‌اند، سوق می‌یابد. به این ترتیب، بر پایه اسپکتروسکوپی غیر مخرب NIR و روش بازشناسی الگو نظارت نشده HCA، توانایی تشخیص و تفکیک وارسته‌های پرتقال با ویژگی‌های متفاوت مزه (والنسیا و تامسون) به خوبی امکان‌پذیر است. از سوی دیگر، الگوی خوشه‌بندی آشکار قابل مشاهده در ساختار درختی سلسله مراتبی نمونه‌ها، امکان تدوین مدل‌های طبقه‌بندی بر پایه طیف‌های NIR برای پیش‌گویی وارسته نمونه‌های ناشناخته را تایید می‌کند.

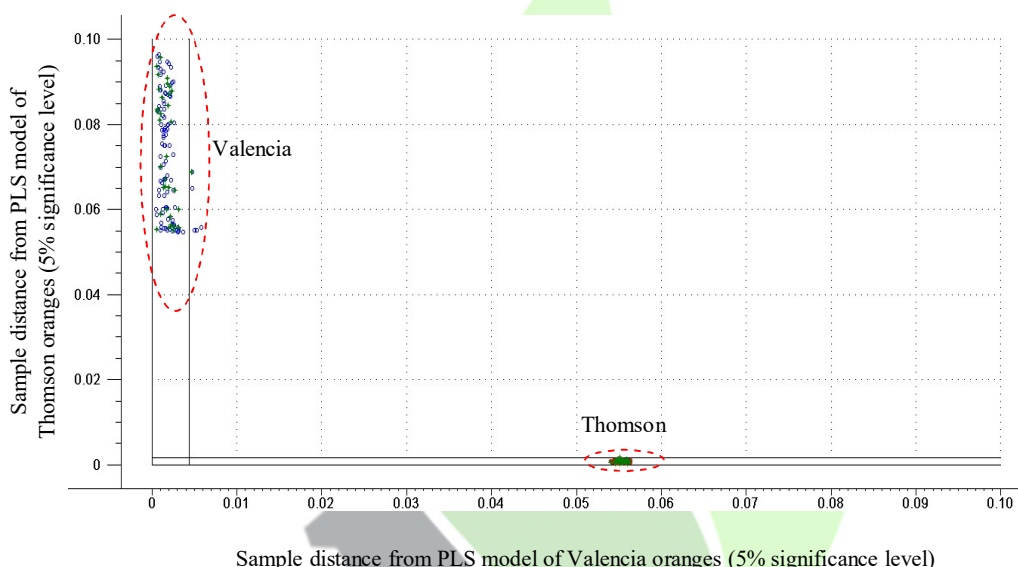


شکل ۲- ساختار درختی سلسله مراتبی به دست آمده از HCA نمونه‌های پرتقال بر پایه طیف‌های NIR
Fig.2. Hierarchical dendrogram obtained from HCA of orange samples based on NIR spectra

شکل ۳ فاصله نمونه‌های پرتقال را از مدل‌های PLS تدوین شده برای وارسته‌های والنسیا و تامسون در طبقه‌بندی SIMCA، بر پایه طیف‌های NIR نشان می‌دهد. همان‌گونه که در شکل دیده می‌شود، دو طبقه والنسیا و تامسون به طور کاملاً مجزا از



یک‌دیگر قرار گرفتند و نمونه‌های هر واریته به اندازه کافی از مدل PLS مربوط به واریته دیگر فاصله داشتند. همچنین، نمونه‌های ناشناخته طبقه‌بندی شده به روش SIMCA بر پایه مدل‌های PLS تدوین شده بر اساس شاخص مزه ترکیبی BrimA هر واریته، به خوبی به طبقه مربوط به خود تعلق گرفتند و یا (در صورت عدم تعلق به طبقه خود) به مدل PLS مربوط به طبقه خود نزدیک بودند. از سوی دیگر، هیچ نمونه‌ای به طبقه اشتباه یا به طور هم‌زمان به دو طبقه تعلق نگرفت که نشان می‌دهد طبقه‌بندی انجام شده از اعتبار ۱۰۰ درصد برخوردار است. به این ترتیب، طبقه‌بندی نظارت شده SIMCA بر پایه طیف‌های NIR نمونه‌ها، توانایی بسیار خوبی برای تفکیک واریته‌های پرتقال بر اساس شاخص BrimA نشان داد.



شکل ۳- فاصله نمونه‌ها از مدل‌های PLS مربوط به واریته‌های پرتقال در طبقه‌بندی SIMCA بر پایه طیف‌های NIR
Fig.3. Sample distances from PLS models of orange varieties in SIMCA classification based on NIR spectra

جدول ۱ نتایج طبقه‌بندی SIMCA را برای تفکیک واریته‌های پرتقال بر اساس مزه (شاخص BrimA) مبتنی بر مدل‌های PLS تدوین شده برای هر واریته نشان می‌دهد که شامل تعداد LVs بهینه، RMSECV، و ضریب همبستگی اعتبارسنجی متقاطع (r_{cv}) در مدل‌های تدوین شده؛ دقت طبقه‌بندی درست نمونه‌ها (صحت طبقه‌بندی)، تعداد نمونه‌هایی که به طبقه اشتباه تعلق گرفته‌اند (False negative)، و تعداد نمونه‌هایی که به طور هم‌زمان به دو طبقه تعلق داده شده‌اند (False positive)، می‌شود. از این رو، طبقه‌بندی SIMCA بر پایه طیف‌های NIR توانایی تفکیک پرتقال‌ها (بر اساس مزه) را با دقت‌های طبقه‌بندی ۹۷/۰۶ درصد برای واریته والنسیا و ۱۰۰ درصد برای واریته تامسون داشت بدون این‌که نمونه‌ای به طبقه اشتباه یا هم‌زمان به دو طبقه اختصاص داده شود (اعتبار طبقه‌بندی ۱۰۰ درصد). همچنین، دقت طبقه‌بندی کل نمونه‌های ناشناخته ۹۸/۵۷ درصد بود.

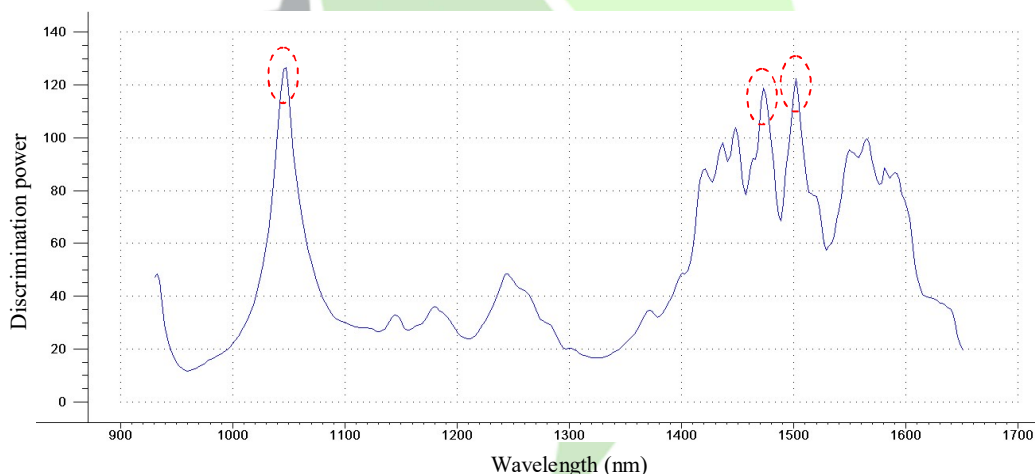


جدول ۱- نتایج طبقه‌بندی SIMCA برای تفکیک پرتقال‌ها بر اساس مزه مبتنی بر مدل‌های PLS تدوین شده برای هر طبقه

Table 1. SIMCA classification results for discrimination of oranges according to their taste based on developed PLS models of each class

| واریته Varieties | LVs | RMSECV | r_{cv} | دقت طبقه‌بندی Classification accuracy | False negative | False positive |
|---------------------|-----|--------|----------|--|----------------|----------------|
| والنسیا Valencia | 5 | 0.44 | 0.86 | 97.06 | 0 | 0 |
| تامسون Thomson | 8 | 0.71 | 0.75 | 100 | 0 | 0 |

شکل ۴ سهم طول موج‌های مختلف را در تفکیک دو طبقه پرتقال بر اساس مزه (شاخص BrimA) نشان می‌دهد. مطابق شکل، تمامی طول موج‌ها در تفکیک دو طبقه موثر ولی قدرت تفکیک طول موج‌های ۱۰۴۷/۵ nm، ۱۵۰۲ nm و ۱۴۷۵ nm به منظور دستیابی به طبقه‌بندی خوب به دست آمده بسیار بیش تر بود. به این ترتیب و بر اساس چگونگی توزیع اورتون‌های پیوندهای اصلی (Cen and He, 2007; Magwaza *et al.*, 2012)، تفکیک پرتقال‌ها بر پایه شاخص مزه BrimA را می‌توان بیش تر به اورتون‌های دوم C-H، CH₂ و (یا) CH₃ (در نزدیکی طول موج ۱۰۴۷/۵ nm)؛ و اورتون‌های اول O-H و (یا) NH₂ (در اطراف طول موج‌های ۱۴۷۵ nm و ۱۵۰۲ nm) نسبت داد.



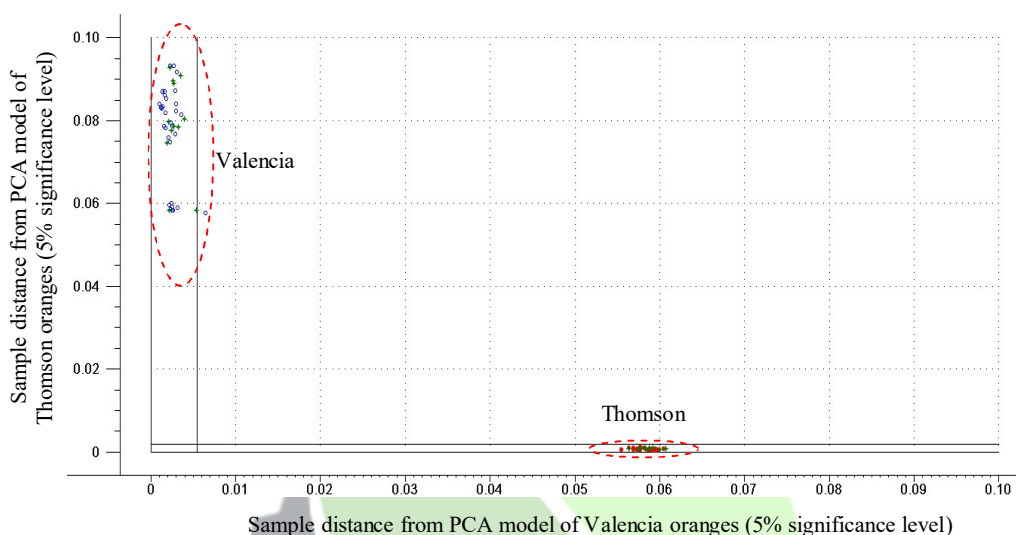
شکل ۴- قدرت تفکیک طول موج‌های مختلف در طبقه‌بندی پرتقال‌ها بر اساس شاخص مزه BrimA

Fig.4. Discrimination power of different wavelengths in classification of oranges according to taste index of BrimA

شکل ۵ فاصله نمونه‌های پرتقال را در طبقه‌بندی SIMCA از مدل‌های PCA تدوین شده برای دو واریته (که بر اساس شاخص BrimA دارای مزه یکسان هستند) بر پایه طیف‌های NIR نشان می‌دهد. دو طبقه والنسیا و تامسون با این که دارای مزه تقریباً یکسان (از نظر شاخص BrimA) بودند، به طور کاملاً مجزا از یکدیگر قرار گرفتند. به گونه‌ای که، نمونه‌های هر واریته به اندازه کافی از مدل PCA مربوط به واریته دیگر فاصله داشتند. از سوی دیگر، نمونه‌های ناشناخته طبقه‌بندی شده به روش



SIMCA بر پایه مدل‌های PCA تدوین شده برای هر واریته به خوبی به طبقه مربوط به خود تعلق گرفتند و یا (در صورت عدم تعلق به طبقه خود) به مدل PCA مربوط به طبقه خود نزدیک بودند. همچنین، هیچ نمونه‌ای به طبقه اشتباه یا به طور هم‌زمان به دو طبقه تعلق نگرفت (اعتبار طبقه‌بندی ۱۰۰ درصد). بنابراین، طبقه‌بندی نظارت شده SIMCA بر پایه طیف‌های NIR نمونه‌ها، توانایی بسیار خوبی برای تفکیک واریته‌های پرتقال که از نظر شاخص Brima دارای مزه تقریباً یکسان بودند، نشان داد.

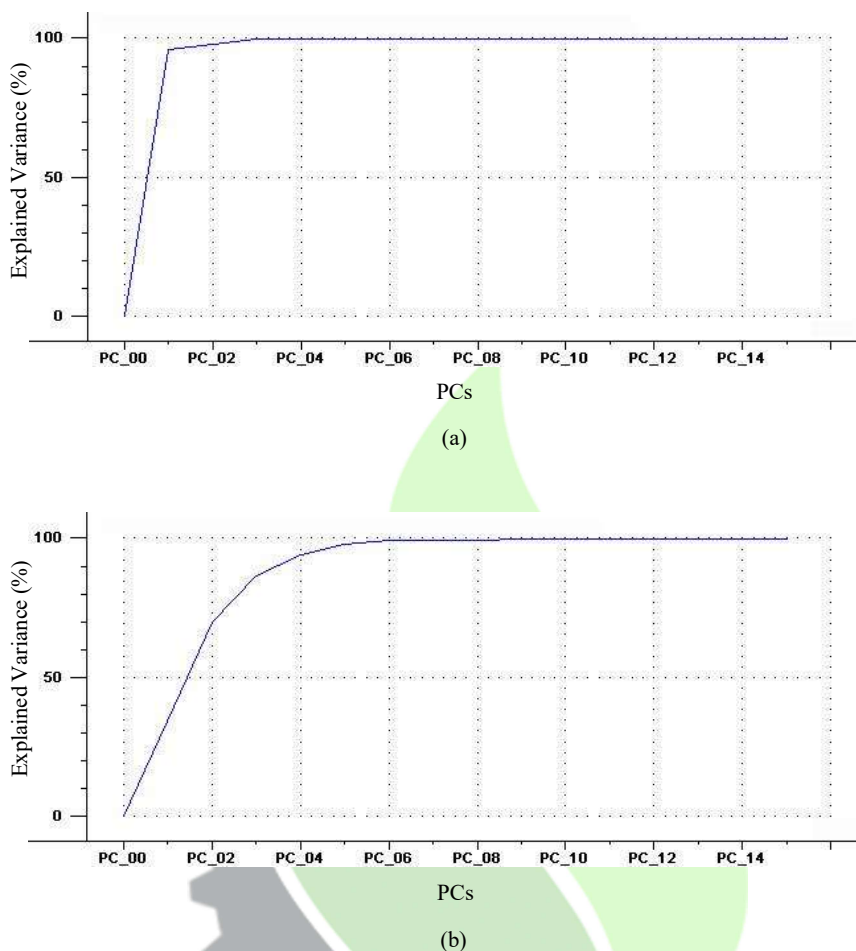


شکل ۵- فاصله نمونه‌ها از مدل‌های PCA مربوط به واریته‌های پرتقال دارای مزه یکسان در طبقه‌بندی SIMCA
Fig.5. Sample distances from PCA models of orange varieties having the same taste in SIMCA classification

نتایج طبقه‌بندی SIMCA برای تفکیک واریته‌های پرتقال دارای مزه یکسان بر پایه طیف‌های NIR و مدل‌های PCA تدوین شده برای هر واریته، در جدول ۲ (شامل تعداد PCs بهینه در مدل‌های تدوین شده برای هر واریته (شکل ۶)، دقت طبقه‌بندی درست نمونه‌ها، نمونه‌های False negative، و False positive) نشان داده شده است.

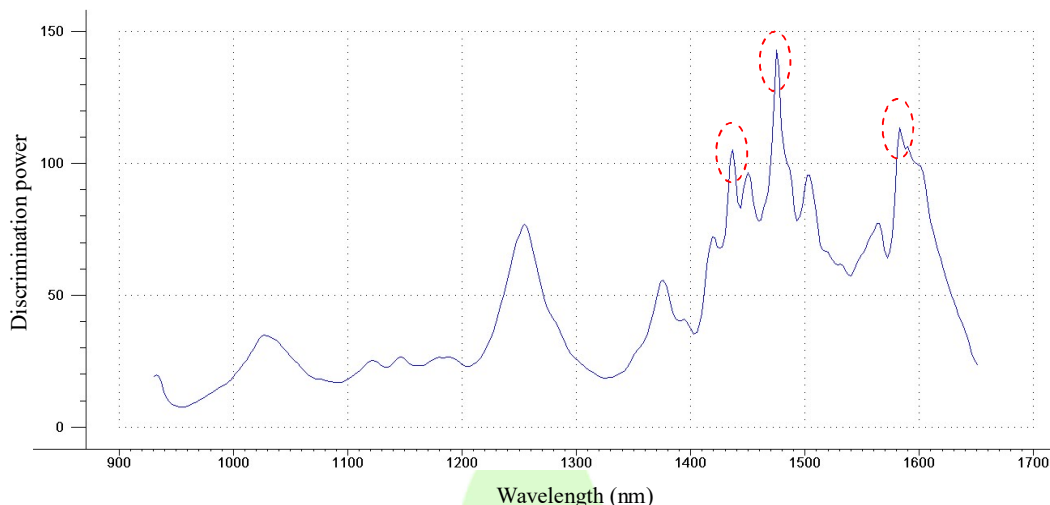
جدول ۲- نتایج طبقه‌بندی SIMCA برای تفکیک واریته‌های پرتقال دارای مزه یکسان بر پایه طیف‌های NIR
Table 2. SIMCA classification results for discrimination of oranges having the same taste based on NIR spectra

| واریته Varieties | PCs | دقت طبقه‌بندی Classification accuracy | False negative | False positive |
|---------------------|-----|--|----------------|----------------|
| والنسیا Valencia | 3 | 91.67 | 0 | 0 |
| تامسون Thomson | 6 | 100 | 0 | 0 |



شکل ۶- واریانس توصیف‌شده داده‌ها در PCs مختلف برای (a) پرتقال‌های والنسیا و (b) تامسون
Fig.6. Explained variance of data in different PCs for (a) Valencia and (b) Thomson oranges

به این ترتیب، طبقه‌بندی SIMCA بر پایه طیف‌های NIR توانایی طبقه‌بندی پرتقال‌های دارای مزه یکسان را با دقت‌های طبقه‌بندی ۹۱/۶۷ درصد برای واریته والنسیا و ۱۰۰ درصد برای واریته تامسون داشت بدون این‌که نمونه‌ای به طبقه اشتباه یا هم‌زمان به دو طبقه اختصاص داده شود. در این حالت نیز، دقت طبقه‌بندی کل نمونه‌های ناشناخته ۹۵/۴۵ درصد بود. سهم طول موج‌های مختلف در تفکیک دو طبقه پرتقال که دارای مزه یکسان بودند، در شکل ۷ نشان داده شده است. طبق شکل، تمامی طول موج‌ها در تفکیک دو طبقه نقش موثر داشتند ولی قدرت تفکیک طول موج‌های ۱۴۷۵ nm، ۱۵۸۳ nm و ۱۴۳۶/۷۵ nm برای دستیابی به طبقه‌بندی خوب بسیار بیش‌تر بود. به این ترتیب و بر اساس چگونگی توزیع اورتون‌های پیوندهای اصلی (Cen and He, 2007; Magwaza *et al.*, 2012)؛ تفکیک پرتقال‌های دارای مزه یکسان (از نظر شاخص BrimA) می‌تواند بیش‌تر ناشی از اورتون‌های اول O-H و (یا) NH₂ (در اطراف طول موج‌های ۱۴۷۵nm و ۱۴۳۶/۷۵nm) و اورتون‌های اول C-H، CH₂ و (یا) CH₃ (در نزدیکی طول موج ۱۵۸۳nm) باشد.



شکل ۷- قدرت تفکیک طول موج‌های مختلف در طبقه‌بندی پرتقال‌های دارای مزه یکسان

Fig.7. Discrimination power of different wavelengths in classification of oranges having the same taste

نتیجه‌گیری

در این پژوهش، از ترکیب روش‌های بازشناسی الگوی نظارت‌شده (HCA) و نظارت‌نشده (SIMCA) با اسپکتروسکوپی NIR در محدوده طیفی ۹۳۰-۱۶۵۰ nm برای بررسی امکان تفکیک و طبقه‌بندی غیر مخرب پرتقال‌های والنسیا و تامسون بر پایه شاخص مزه ترکیبی (BrimA)، همچنین سایر پارامترهای مربوط به مزه (در نمونه‌های با مزه یکسان) استفاده شد. آنالیزهای کیفی برای تفکیک غیر مخرب پرتقال‌ها نشان داد که از اطلاعات طیفی نمونه‌ها در محدوده طیفی ۹۳۰-۱۶۵۰ nm (NIR) و روش بازشناسی الگو نظارت‌نشده HCA، می‌توان به خوبی برای تفکیک واریته‌های پرتقال با ویژگی‌های متفاوت مزه استفاده کرد. همچنین، طبقه‌بندی نظارت‌شده SIMCA بر پایه طیف‌های NIR توانایی تفکیک و جداسازی واریته‌های پرتقال را بر اساس مزه (شاخص BrimA) با دقت طبقه‌بندی عالی ۹۸/۵۷ درصد داشت. به گونه‌ای که، طول موج‌های ۱۰۴۷/۵ nm، ۱۵۰۲ nm و ۱۴۷۵nm بیش‌ترین سهم را در تفکیک دو طبقه بر عهده داشتند. از سوی دیگر، امکان تفکیک نمونه‌های پرتقال از واریته‌های مختلف که دارای مزه یکسان بودند نیز با طبقه‌بندی SIMCA و بر پایه طیف‌های NIR با دقت عالی ۹۵/۴۵ درصد وجود داشت و قدرت تفکیک طول موج‌های ۱۴۷۵ nm، ۱۵۸۳nm و ۱۴۳۶/۷۵ nm برای دستیابی به این طبقه‌بندی بسیار بیش‌تر بود. بنابراین، می‌توان از ترکیب روش‌های بازشناسی الگو با اسپکتروسکوپی غیر مخرب NIR برای تشخیص دیگر پارامترهای مربوط به مزه پرتقال مانند اسیدهای آمینه نیز استفاده کرد.

منابع

- Camps, C., and D. Christen. 2009. Non-destructive assessment of apricot fruit quality by portable visible-near infrared spectroscopy. *LWT-Food Science and Technology* 42: 1125-1131.
- Cao, F., D. Wu, and Y. He. 2010. Soluble solids content and pH prediction and varieties discrimination of grapes based on visible-near infrared spectroscopy. *Computers and Electronics in Agriculture* 71s: s15-s18.
- Cayuela, J. A., and C. Weiland, 2010. Intact orange quality prediction with two portable NIR spectrometers. *Postharvest Biology and Technology*. 58(2): 113-120.
- Cen, H., and Y. He. 2007. Theory and application of near infrared reflectance spectroscopy in determination of food quality. *Trends in Food Science and Technology* 18: 72-83.
- Flores, K., M. Sánchez, D. Pérez-Marín, J. Guerrero, and A. Garrido-Varo. 2009. Feasibility in NIRS instruments for predicting internal quality in intact tomato. *Journal of Food Engineering* 91: 311-318.
- Fu, X., Y. Ying, H. Lu, H. Xu, and H. Yu. 2007. FT-NIR diffuse reflectance spectroscopy for kiwifruit firmness detection. *Sensing and Instrumentation for Food Quality and Safety* 1: 29-35.
- Jamshidi, B., S. Minaei, E. Mohajerani, and H. Ghassemian. 2011a. Analysis of citrus peel for non-destructive determination of fruit composition by reflectance Vis/NIR spectroscopy. XXXIV CIOSTA CIGR V Conference on Efficient and Safe Production Processes in Sustainable Agriculture and Forestry. Vienna, Austria.
- Jamshidi, B., S. Minaei, E. Mohajerani, and H. Ghassemian. 2012a. Multivariate analysis of reflectance Vis/NIR spectra based on wavelet transform for non-destructive and detection of orange color and pH simultaneously. 7th National Congress on Agricultural Machinery Engineering and Mechanization. Shiraz, Iran. (In Farsi).
- Jamshidi, B., S. Minaei, E. Mohajerani, and H. Ghassemian. 2012b. Vis/NIR spectroscopy for non-destructive classification of orange varieties. 7th National Congress on Agricultural Machinery Engineering and Mechanization. Shiraz, Iran. (In Farsi).



- Jamshidi, B., S. Minaei, E. Mohajerani, and H. Ghassemian. 2012c. Reflectance Vis/NIR spectroscopy for nondestructive taste characterization of valencia oranges. *Computers and Electronics in Agriculture* 85: 64-69.
- Jamshidi, B., S. Minaei, E. Mohajerani, and H. Ghassemian. 2013a. Linear multivariate model based on NIR spectroscopy for non-destructive internal quality prediction of orange. 19th Iranian Conference on Optics and Photonics, and 5th Iranian Conference on Photonics Engineering. Zahedan, Iran. (In Farsi).
- Jamshidi, B., S. Minaei, E. Mohajerani, and H. Ghassemian. 2013b (in press). Prediction of soluble solids in oranges using visible/near-infrared spectroscopy: Effect of peel. *International Journal of Food Properties*.
- Jamshidi, B., S. Minaei, E. Mohajerani, H. Ghassemian, and H. Afkhami Ardakani. 2011b. 17th Iranian Conference on Optics and Photonics, and 3rd Iranian Conference on Photonics Engineering. Kerman, Iran. (In Farsi).
- Jordan, R., R. Seelye, and A. McGlone. 2001. A sensory-based alternative to brix/acid ratio. *Food Technology* 55(6): 36-44.
- Luo, W., S. Huan, H. Fu, G. Wen, H. Cheng, J. Zhou, H. Wu, G. Shen, and R. Yu. 2011. Preliminary study on the application of near infrared spectroscopy and pattern recognition methods to classify different types of apple samples. *Food Chemistry* 128(2): 555-561.
- Luo, X., T. Takahashi, and S. Zhang. 2013. Wavelength selection in visible and near infrared spectra for detection of bruises on apples. *Advanced Science Letters*. 19(9): 2654-2657.
- Magwaza, L., U. Opara, H. Nieuwoudt, P. Cronje, W. Saeys, and B. Nicolaï. 2012. NIR spectroscopy applications for internal and external quality analysis of citrus fruit-A review. *Food and Bioprocess Technology*. 5(2): 425-444.
- Manley, M., J. Elizabeth, M. Lindie, L. Ester, and K. Martin. 2007. Prediction of soluble solids content and post-storage internal quality of Bulida apricots using near infrared spectroscopy. *Journal of Near Infrared Spectroscopy* 15: 179-188.
- Mireei, S. A., S. S. Mohtasebi, R. Massudi, S. Rafiee, and A. S. Arabanian. 2010. Feasibility of near infrared spectroscopy for analysis of date fruits. *International Agrophysics*. 24: 351-356.



- Moghimi, A., M. H. Aghkhani, A. Sazgarnia, and M. Sarmad. 2010. Vis/NIR spectroscopy and chemometrics for the prediction of soluble solids content and acidity (pH) of kiwifruit. *Biosystems Engineering*. 106(3): 295-302.
- Nicolai, B. M., K. Beullens, E. Bobelyn, A. Peirs, W. Saeys, K. I. Theron, and J. Lammertyn. 2007. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biology and Technology* 46: 99-118.
- Obenland, D., S. Collin, B. Mackey, J. Sievert, K. Fjeld, and M. L. Arpaia. 2009. Determinants of flavor acceptability during the maturation of navel oranges. *Postharvest Biology and Technology* 52: 156-163.
- Penchaiya, P., E. Bobelyn, B. E. Verlinden, B. M. Nicolai, and W. Saeys. 2009. Non-destructive measurement of firmness and soluble solids content in bell pepper using NIR spectroscopy. *Journal of Food Engineering* 94: 267-273.
- Pissard, A., J. A. Fernández-Pierna, V. Baeten, G. Sinnaeve, G. Lognay, A. Mouteau, P. Dupont, A. Rondia, and M. Lateur. 2013. Non-destructive measurement of vitamin C, total polyphenol and sugar content in apples using near-infrared spectroscopy. *Journal of the Science of Food and Agriculture*. 93(2): 238-244.
- Subedi, P. P., and K. B. Walsh. 2011. Assessment of sugar and starch in intact banana and mango fruit by SWNIR spectroscopy. *Postharvest Biology and Technology* 62(3): 238-245.
- Suphamitmongkol, W., G. Nie, R. Liu, S. Kasemsumran, and Y. Shi. 2013. An alternative approach for the classification of orange varieties based on near infrared spectroscopy. *Computers and Electronics in Agriculture*. 91: 87-93.
- Varmuza, K., and P. Filzmoser. 2009. *Introduction to Multivariate Statistical Analysis in Chemometrics*. CRC Press. Boca Raton.



Pattern recognition of near-infrared spectroscopy for non-destructive discrimination of oranges based on taste index

Bahareh Jamshidi^{1*}, Saeid Minaei², Ezzedin Mohajerani³ and Hassan Ghassemian⁴

1- Ph.D. Graduate, Tarbiat Modares University (Assistant Professor, Agricultural Engineering Research Institute), b.jamshidi@aeri.ir

2- Professor, Faculty of Agriculture, Tarbiat Modares University

3- Professor, Laser and Plasma Research Institute, Shahid Beheshti University

4- Professor, Faculty of Electrical and Computer Engineering, Tarbiat Modares University

Abstract

In recent years, application of near-infrared spectroscopy (NIR) as a non-destructive technique combined with chemometric methods has been widely noticed for quality assessment of food and agricultural products. In chemometric methods, quality analyses are important issues which could be attributed to the problem of pattern recognition. In this research, the feasibility of pattern recognition methods combined with reflectance NIR spectroscopy for non-destructive discrimination of oranges based on their taste was investigated. To this end, both unsupervised and supervised pattern recognition techniques, hierarchical cluster analysis (HCA) and soft independent modeling of class analogies (SIMCA) were used for assessing the feasibility of variety discrimination and classification (according to their taste), respectively based on the spectral information of 930-1650nm range. Qualitative analyses indicated that NIR spectra of orange varieties were correctly clustered using unsupervised pattern recognition of HCA. It was also concluded that supervised pattern recognition of SIMCA for NIR spectra of oranges provided excellent results of variety classification based on BrimA index at 5% significance level (classification accuracy of 98.57%). Moreover, wavelengths of 1047.5nm, 1502nm, and 1475nm contributed more than other wavelengths to discriminate two classes. Samples having the same BrimA index were also correctly classified with high classification accuracy (95.45%) at 5% significance level. The discrimination power of wavelengths of 1475nm, 1583nm, and 1436.75nm were more than those for other wavelengths to achieve this classification. Therefore, reflectance NIR spectroscopy combined with pattern recognition methods can be utilized for determination of other attributes related to taste.

Key Words: Classification, Near-infrared Spectroscopy, Non-destructive, Pattern Recognition, Taste