

## بررسی اثر پیش‌پردازش‌های مختلف طیفی بر ارزیابی غیر مخرب کیفیت پرتقال با اسپکتروسکوپی فوتومتری فروسرخ نزدیک (NIRS)

بهاره جمشیدی، سعید مینایی\*، عزالدین مهاجرانی و حسن قاسمیان\*\*

\*نگارنده مسئول، نشانی: تهران، دانشگاه تربیت مدرس، دانشکده کشاورزی، ص. پ. ۳۳۶-۱۴۱۱۵، تلفن: ۴۸۲۹۲۴۶۶ (۰۲۱)، پیام‌نگار: minaee@modares.ac.ir  
\*\*به‌ترتیب: دانش‌آموخته دکتری تخصصی دانشگاه تربیت مدرس و استادیار موسسه تحقیقات فنی و مهندسی کشاورزی؛ استاد دانشکده کشاورزی دانشگاه تربیت مدرس؛ استاد پژوهشکده لیزر و پلاسما دانشگاه شهید بهشتی؛ و استاد دانشکده مهندسی برق و کامپیوتر دانشگاه تربیت مدرس  
تاریخ دریافت: ۹۲/۹/۵؛ تاریخ پذیرش: ۹۳/۵/۱۸

### چکیده

اسپکتروسکوپی فروسرخ نزدیک (NIR) در ترکیب با روش‌های شیمی‌سنجی، شامل پیش‌پردازش‌های طیفی و مدل‌سازی‌های چندمتغیره، یکی از پرکاربردترین روش‌های غیر مخرب اندازه‌گیری ویژگی‌های کیفی میوه‌ها و سبزی‌هاست که در سال‌های اخیر بیشتر مد نظر پژوهشگران بوده است. در این پژوهش، توانایی روش اسپکتروسکوپی NIR بازتابی در محدوده طیفی ۱۶۵۰-۹۳۰ نانومتر به منظور پیش‌بینی غیر مخرب مواد جامد حل‌شدنی (SSC) و اسیدیته قابل‌تیر کردن (TA) پرتقال تامسون، همچنین اثر پیش‌پردازش‌های مختلف طیفی بر دقت مدل‌های چندمتغیره پیش‌بینی‌کننده بررسی شد. در این راستا، مدل‌های واسنجی چندمتغیره حداقل مربعات جزئی (PLS) بر پایه اندازه‌گیری‌های مرجع و اطلاعات طیف‌های پیش‌پردازش شده با ترکیب روش‌های مختلف هموارسازی (میانگین‌گیری متحرک (MA)، ساویتزکی-گولای (SG)، تبدیل موجک (WT))؛ نرمال‌سازی (تصحیح پراکنش افزاینده (MSC)، توزیع نرمال استاندارد (SNV))؛ و افزایش قدرت تفکیک طیفی (مشتق‌های اول و دوم ( $D_1, D_2$ )) برای پیش‌بینی SSC و TA پرتقال‌ها تدوین شدند. نتایج نشان داد که اسپکتروسکوپی NIR بازتابی، در ترکیب با روش‌های شیمی‌سنجی، توانایی پیش‌بینی غیر مخرب SSC و TA پرتقال را دارد. همچنین، روش‌های پیش‌پردازش اثر مستقیم بر نتایج مدل‌های PLS تدوین شده برای پیش‌بینی این پارامترهای درونی داشتند و بهترین نتایج پیش‌بینی برای SSC ( $r_p=+0.923$ , RMSEC= $+0.43$ )،  $r_p=+0.936$ , RMSEP= $+0.451$ ،  $r_c=+0.883$ , RMSEC= $+0.133$ ) و TA ( $r_p=+0.863$ , RMSEP= $+0.177$ ،  $r_c=+0.853$ , SDR= $1/853$ ) بر پایه ترکیب روش‌های پیش‌پردازش MA + SNV به دست آمد.

### واژه‌های کلیدی

آنالیز چندمتغیره، اسپکتروسکوپی فروسرخ نزدیک، روش‌های پیش‌پردازش، غیر مخرب

### مقدمه

صنعت کشاورزی و غذایی کشورها با چالش‌های فناوری نوین مواجه است. رویکردهای اخیر برای روبه‌رو شدن با این چالش‌ها، کاهش اعتماد به ارزیابی کیفی ذهنی<sup>۱</sup> و بالا بردن پذیرش روش‌های عینی<sup>۲</sup>، کمی، و غیر مخرب<sup>۳</sup> ارزیابی کیفی است. از این رو، در سال‌های اخیر کاربرد

امروزه، به دلیل اهمیت کیفیت میوه‌ها و سبزی‌ها که به‌طور مستقیم با سلامت مردم در ارتباط است و برای همگام‌شدن با آگاهی‌ها و تقاضاهای رو به افزایش مصرف‌کنندگان در خصوص تضمین کیفی محصول،

گزارش شده است (Camps & Christen, 2009; Mireei *et al.*, 2010; Moghimi *et al.*, 2010; Wang *et al.*, 2011; Pissard *et al.*, 2013; Salguero-Chaparro *et al.*, 2013) ولی دستیابی به نتایج پیش‌بینی قابل اعتماد با این روش برای میوه‌های با پوست ناصاف و ضخیم مانند مرکبات، بسیار سخت‌تر و پیچیده‌تر است (Cayuela, 2008; Jamshidi *et al.*, 2011a, 2014). با این حال، تاکنون پژوهش‌هایی نیز به منظور تعیین و ارزیابی غیرمخرب ویژگی‌های کیفی درونی مرکبات انجام شده است (Magwaza *et al.*, 2012). در این راستا، اگرچه نتایج خوبی برای پیش‌بینی برخی ویژگی‌های کیفی مانند میزان جامد حل‌شدنی<sup>۶</sup> (SSC) به دست آمده است (Cayuela, 2008; Cayuela & Weiland, 2010; Liu *et al.*, 2010a,c, 2012; Jamshidi *et al.*, 2012c, 2013)؛ ولی در مورد پیش‌بینی برخی ویژگی‌های مهم دیگر مانند اسیدیته قابل تیتراژ<sup>۷</sup> (TA)، به دلیل سطوح نسبتاً پایین اسیدهای ارگانیک در مرکبات، دستیابی به نتایج مناسب با روش مذکور به مراتب سخت‌تر است و به پژوهش‌های بیشتر نیاز دارد (Guthrie *et al.*, 2005; Cayuela, 2008; Cayuela & Weiland, 2010; Liu *et al.*, 2010b; Jamshidi *et al.*, 2012c) از این رو و با توجه به پیشرفت‌های اخیر در روش‌های شیمی‌سنجی<sup>۸</sup>، از قبیل روش‌های پیش‌پردازش<sup>۹</sup> و آنالیزهای چندمتغیره برای استخراج اطلاعات مفید از یک طیف NIR، هدف از این پژوهش بررسی امکان تشخیص غیر مخرب دو پارامتر کیفی (SSC و TA) مهم پرتقال تامسون بر پایه اسپکتروسکوپی NIR با به‌کارگیری روش‌های مختلف پیش‌پردازش‌های طیفی به منظور دستیابی به مدل‌های چندمتغیره قابل اعتماد است.

روش‌های غیر مخرب به منظور ارزیابی و پایش کیفی قابل اعتماد و سریع میوه‌ها و سبزی‌ها به شکلی چشمگیر مورد توجه پژوهشگران و صنعتگران قرار گرفته است. برخی از روش‌های غیر مخرب مانند بینایی ماشین<sup>۱</sup>، برای ارزیابی ذهنی وجوه مرئی بیرونی (ظاهر) بر پایه رنگ، شکل، اندازه و عیب‌های بیرونی میوه‌ها و سبزی‌ها هم‌اکنون به صورت تجاری به کار می‌روند؛ ولی ارزیابی غیر مخرب بر اساس ویژگی‌های کیفی درونی مانند: رسیدگی، مقدار قند، اسید، و عیب‌های درونی همچنان مشکل و نیازمند پژوهش‌های بیشتر است (Magwaza *et al.*, 2012).

اسپکتروفتومتری<sup>۲</sup> یا اسپکتروسکوپی<sup>۳</sup> فروسرخ نزدیک<sup>۴</sup> (NIR)، یکی از روش‌های نوین سریع و مقرون به صرفه (بر پایه اصول اپتیک) است که توانمندی آن در کاربردهای کشاورزی و غذایی به‌ویژه به منظور ارزیابی کیفیت درونی میوه‌ها و سبزی‌ها بی‌مانند است (Bokobza, 2006; Cen & He, 2007). بی‌نیاز بودن از آماده‌سازی نمونه، پایین بودن نسبی هزینه‌ها، سازگاری بودن با شرایط زیست-محیطی، و سهولت به‌کارگیری در سامانه‌های درجه‌بندی و کنترل فرایند، از دیگر امتیازهای این روش نسبت به سایر روش‌های غیر مخرب است. ضمن اینکه، پیشرفت‌های اخیر در علم کامپیوتر، ریاضی و روش‌های آماری، محبوبیت اسپکتروسکوپی NIR و تلاش برای هر چه توانمندتر کردن آن در کاربردهای کنترل فرایند برخط<sup>۵</sup> را دو چندان کرده است (Jamshidi *et al.*, 2011b, 2012a,b).

تاکنون از اسپکتروسکوپی NIR یا Vis/NIR (مرئی/فروسرخ نزدیک) برای تعیین برخی ویژگی‌های کیفی درونی میوه‌ها به روش غیرمخرب استفاده و نتایج خوبی برای محصولات نازک پوست و دارای سطح صاف

1- Machine Vision  
3- Spectroscopy  
5- On-line  
7- Titratable Acidity  
9- Pre-processing Methods

2- Spectrophotometry  
4- Near-infrared  
6- Soluble Solids Content  
8- Chemometrics

## مواد و روش‌ها

### تهیه و انتخاب نمونه‌ها

انجام و در هر موقعیت تعداد ۵ طیف از هر نمونه ذخیره شد. همچنین، میانگین طیف‌های به‌دست‌آمده از هر نمونه به‌عنوان طیف معرف آن نمونه در نظر گرفته شد.

### اندازه‌گیری‌های مرجع

پس از پایان آزمون‌های اسپکتروسکوپی به‌منظور اندازه‌گیری‌های مرجع، نمونه‌ها آب‌گیری شدند. از هر نمونه پرتقال حدود ۱۰۰ میلی‌لیتر آب‌میوه به‌دست آمد؛ آب‌میوه به‌دست آمده صاف شد. برای اندازه‌گیری SSC بر حسب درجه بریکس (°Brix)، از دستگاه رفراکتومتر مدل DR-A1 (ATAGO, Co. Ltd, Japan) با دقت ۰/۱ اعشار استفاده شد. به‌منظور اندازه‌گیری TA بر حسب اسیدسیتریک، تا نقطه پایانی صورتی کم‌رنگ، از روش تیتراسیون با محلول هیدروکسید سدیم (سود) ۰/۱ مولار استفاده شد. درصد اسیددیده آب پرتقال بر حسب اسیدسیتریک از رابطه ۱ به‌دست آمد (James, 1998):

$$\%TA = (t \times 192) / (3 \times 1000) \quad (1)$$

که در آن،

$t$  = میانگین مقدار سود مصرفی (میلی‌لیتر)؛  $192 =$  وزن مولکولی اسیدسیتریک؛ و  $3 =$  ضریب سود در واکنش شیمیایی است.

### آماده‌سازی و پیش‌پردازش طیف‌ها

داده‌های طیفی به‌دست آمده از اسپکترومتر، افزون بر اطلاعات نمونه شامل اطلاعات ناخواسته پس‌زمینه و نویزها نیز هستند. از این رو، برای دستیابی به مدل‌های واسنجی پایدار، دقیق و قابل اعتماد، نیاز به پیش‌پردازش داده‌های طیفی پیش از تدوین مدل‌های رگرسیون است. برای این منظور، روش‌های پیش‌پردازش بسیاری بر پایه شیوه‌های ریاضی و آماری گسترش یافته یا در حال توسعه‌اند که هر کدام برای هدفی استفاده می‌شود. بر این اساس، پس از این‌که طیف‌های بازتابی به‌دست آمده از همه نمونه‌ها به‌منظور ایجاد یک ارتباط خطی با غلظت مولکولی نمونه

نمونه‌ها از پرتقال‌های وارسته تامسون انتخاب شد که یکی از وارسته‌های معروف و بسیار محبوب (از نظر طعم و مزه) در جهان است. تعداد ۱۵۰ نمونه پرتقال تامسون از بازارهای محلی مختلف و در روزهای مختلف خریداری شد. برای ایجاد تنوع بیشتر در نمونه‌ها و جلوگیری از بروز مشکلات نگهداری و انتقال آنها، هر بار ۱۰ پرتقال خریداری و آزمایش‌ها روی آنها اجرا شد. نمونه‌ها پیش از اندازه‌گیری‌های اسپکتروسکوپی به‌مدت یک روز در دمای ۲۵ درجه سلسیوس نگهداری شدند تا دمای آنها با دمای محیط آزمایشگاه متعادل شود. پیش از اسپکتروسکوپی نیز، وزن و قطر استوایی بیشینه نمونه‌ها به‌ترتیب با ترازوی دیجیتال (با دقت ۰/۰۱ اعشار) و کولیس ورنیه اندازه‌گیری شد.

### اندازه‌گیری‌های اسپکتروسکوپی NIR

آزمون‌های اسپکتروسکوپی NIR از نمونه‌های پرتقال تامسون در مُد اندازه‌گیری بازتابی با استفاده از چیدمان مناسب طراحی‌شده (Jamshidi *et al.*, 2012c) و اسپکترومتر<sup>۱</sup> پاشنده مدل EPP2000-NIR دارای قدرت تفکیک ۲/۵ نانومتر و مجهز به آشکارساز نوع آرایه فتودیود<sup>۲</sup> (PDA) اجرا و طیف‌ها در محدوده طیفی ۱۶۵۰-۹۳۰ نانومتر ذخیره شدند. با توجه به اینکه هر بار ۱۰ نمونه مورد آزمون اسپکتروسکوپی قرار می‌گرفت؛ پیش از هر بار اندازه‌گیری NIR، طیف‌های تیره (در حالت منبع نور خاموش) و مرجع (در حالت منبع نور روشن و استفاده از یک دیسک تفلون استاندارد) به‌منظور دستیابی به بازتاب نسبی (Anon, 2005) تعریف شد و شدت مناسب منبع نور با استفاده از دیمِر (کنترل‌کننده شدت نور) و به‌طور یکسان در هر بار اندازه‌گیری تنظیم گردید. اسپکتروسکوپی NIR بازتابی همه نمونه‌های پرتقال در سه موقعیت با زوایای ۱۲۰ درجه حول محور استوایی آنها

1- Spectrometer

2- Photodiode Array

چندجمله‌ای (که بیشتر تجربی است) اهمیت دارد تا ضمن افزایش نسبت سیگنال به نویز، اطلاعات پیک‌ها هم حفظ شود (Heise & Winzen, 2006). بنابراین، الگوریتم SG با در نظر گرفتن تعداد نقطه‌های همسایگی و درجه چندجمله‌ای به ترتیب برابر با ۵ و ۲ اجرا شد.

**تبدیل موجک (WT):** برای پردازش داده‌های طیفی و نیز برای هموارسازی طیف‌ها و مقایسه با دو روش هموارسازی دیگر از تبدیل موجک (WT) یا تجزیه یک تابع بر پایه توابع موجک و مبتنی بر دو پارامتر انتقال<sup>۸</sup> و مقیاس<sup>۹</sup> که کاربردهای گسترده‌ای در هموارسازی، حذف نویز و پس‌زمینه، فشردن داده، و تجزیه سیگنال‌های روی هم افتاده دارد (Fu et al., 2005; Xia et al., 2007; Zhu et al., 2007)؛ استفاده شد. به این ترتیب که ابتدا طیف‌های اصلی با استفاده از تجزیه موجک چندسطحی<sup>۱۰</sup> به ضریب‌های تقریب<sup>۱۱</sup> و ضریب‌های جزئی<sup>۱۲</sup> تجزیه و سپس ضریب‌های جزئی با درصد مشخصی حذف شدند. سرانجام، طیف هموار شده با اجرای تبدیل موجک وارون<sup>۱۳</sup> روی ضریب‌های تقریب و ضریب‌های جزئی کوتاه شده، محاسبه شد. موجک‌ها از نوع دابیچی<sup>۱۴</sup> (Db3) انتخاب و درصد حذف ضریب‌های جزئی ۵۰ در نظر گرفته شد.

**تصحیح پراکنش افزایشنده (MSC):** به منظور نرمال‌سازی طیف‌ها و تصحیح اثرهای جمعی و افزایش ناشی از اثرهای فیزیکی در داده‌های طیفی مانند پخش غیر یکنواخت در کل طیف، اندازه نمونه، و ضریب شکست نور، از روش تصحیح پراکنش افزایشنده (MSC) با خطی‌سازی طیف هر نمونه به طیف میانگین نمونه‌ها استفاده شد (Fu et al., 2007).

به طیف‌های جذبی تبدیل شدند، برای آماده‌سازی و پیش‌پردازش طیف‌ها از برخی روش‌های هموارسازی<sup>۱</sup> و کاهش نویز (میانگین‌گیری متحرک<sup>۲</sup> (MA)، ساویتزکی-گولای<sup>۳</sup> (SG)، تبدیل موجک<sup>۴</sup> (WT))؛ نرمال‌سازی<sup>۵</sup> (تصحیح پراکنش افزایشنده<sup>۶</sup> (MSC)، توزیع نرمال استاندارد<sup>۷</sup> (SNV))؛ و افزایش قدرت تفکیک طیفی (مشق‌های اول و دوم (D<sub>1</sub>, D<sub>2</sub>)) استفاده شد و اثر هشت شیوه ترکیبی SG + MSC؛ MA + SNV؛ MA + MSC؛ SG + SNV؛ WT + MSC؛ WT + SNV؛ WT + D<sub>1</sub>؛ SG + D<sub>1</sub>؛ SG + D<sub>2</sub> بر دقت مدل‌های چندمتغیره تدوین شده برای پیش‌بینی پارامترهای SSC و TA بررسی گردید. برای اجرای این روش‌ها از نرم‌افزارهای Unscrambler X 10.1 (CAMO Software AS, Norway) و OriginPro8 (OriginLab Corporation, USA) استفاده شد.

**میانگین‌گیری متحرک (MA):** در این روش، از یک فیلتر میانگین‌گیر متحرک (MA) استفاده و هر نقطه از طیف با میانگین یک همسایگی از آن جایگزین شد. به منظور افزایش نسبت سیگنال به نویز، بدون از دست دادن اطلاعات پیک‌ها، انتخاب درست عرض پنجره فیلتر (که بیشتر تجربی است) در این مرحله اهمیت ویژه‌ای دارد (Heise & Winzen, 2006)؛ بر این اساس، عرض پنجره ۵ برای فیلتر در نظر گرفته شد.

**ساویتزکی-گولای (SG):** هموارسازی ساویتزکی-گولای (SG) بر پایه الگوریتم SG و با برازش یک چندجمله‌ای بر هر بخش متوالی طیف (همسایگی) و جایگزین کردن مقادیر اصلی طیف با مقادیرهای برازش شده انجام شد. در این روش نیز انتخاب تعداد نقطه‌های همسایگی و درجه

1- Smoothing

3- Savitzky-Golay

5- Normalizing

7- Standard Normal Variate

9- Scale Parameter

11- Approximation Coefficients

13- Inverse Wavelet Transform

2- Moving Average

4- Wavelet Transform

6- Multiplicative Scatter Correction

8- Translation Parameter

10- Multilevel Wavelet Decomposition

12- Detail Coefficients

14- Daubechies

روش‌های پیش‌پردازش اعمال شده روی داده‌های طیفی انجام شدند. سپس، مدل‌های PLS تدوین شده اعتبارسنجی شدند. در این راستا، خطای مدل‌ها بر اساس ریشه میانگین مربعات خطای واسنجی<sup>۳</sup> (RMSEC)، ریشه میانگین مربعات خطای پیش‌بینی<sup>۴</sup> (RMSEP)، ضریب همبستگی واسنجی<sup>۵</sup> ( $r_c$ )؛ ضریب همبستگی پیش‌بینی<sup>۶</sup> ( $r_p$ )؛ همچنین شاخص SDR (نسبت انحراف استاندارد (SD) ویژگی موردنظر در دسته آزمون به RMSEP) طبق رابطه‌های ۲ تا ۶ تعیین شدند (Nicolai et al., 2007).

$$RMSEC = \sqrt{\sum_{i=1}^{n_c} (\hat{y}_i - y_i)^2 / n_c} \quad (2)$$

$$RMSEP = \sqrt{\sum_{i=1}^{n_p} (\hat{y}_i - y_i)^2 / n_p} \quad (3)$$

$$r_c = \sqrt{\sum_{i=1}^{n_c} (\hat{y}_i - y_i)^2 / \sum_{i=1}^{n_c} (\hat{y}_i - y_m)^2} \quad (4)$$

$$r_p = \sqrt{\sum_{i=1}^{n_p} (\hat{y}_i - y_i)^2 / \sum_{i=1}^{n_p} (\hat{y}_i - y_m)^2} \quad (5)$$

$$SDR = \sqrt{\sum_{i=1}^{n_p} (y_i - y_m)^2 / (n_p - 1)} / RMSEP = SD / RMSEP \quad (6)$$

که در آن‌ها،

$y_i$  = مقدار اندازه‌گیری شده ویژگی برای نمونه  $i$  ام در دسته آموزش یا دسته آزمون؛  $\hat{y}_i$  = مقدار پیش‌بینی شده ویژگی برای نمونه  $i$  ام در دسته آموزش یا دسته آزمون؛  $n_c$  = تعداد نمونه‌های دسته آموزش؛  $n_p$  = تعداد نمونه‌های دسته آزمون؛ و  $y_m$  = میانگین مقدارهای اندازه‌گیری شده ویژگی در دسته آموزش یا دسته آزمون است.

مدل بهینه از بین مدل‌های تدوین شده برای پیش‌بینی SSC و TA و بررسی اثر پیش‌پردازش‌های مختلف طیفی بر دقت مدل‌های تدوین شده بر پایه داشتن RMSEP کمتر،  $r_p$  و SDR بیشتر انتخاب شد. کمتر بودن

توزیع نرمال استاندارد (SNV): روش توزیع نرمال استاندارد (SNV) نیز برای حذف اثرهای جمعی و افزایش پخش نور، اندازه نمونه، و تغییر فاصله نوری همچنین مقایسه با روش MSC و نرمال‌سازی هر طیف با انحراف استاندارد طیف در طول محدوده طیفی اجرا شد (Fu et al., 2007).

مشق اول و دوم ( $D_1$ ,  $D_2$ ): برای افزایش قدرت تفکیک طیفی، حذف پس‌زمینه و جابه‌جایی خط پایه پس از هموارسازی طیف‌ها با استفاده از الگوریتم ساویتزکی-گولای (SG)، از مشتق‌گیری اول و دوم ( $D_1$  و  $D_2$ ) استفاده شد. در این مرحله نیز انتخاب پارامترهای الگوریتم شامل تعداد نقطه‌های همسایگی و درجه چندجمله‌ای به‌منظور جلوگیری از تقویت نویز طیفی تا آن‌جا اهمیت داشت که تعداد نقطه‌های همسایگی و درجه چندجمله‌ای به ترتیب برابر با ۵ و ۲ در نظر گرفته شد.

### مدل‌سازی رگرسیون چندمتغیره

به‌منظور ایجاد ارتباط بین SSC و TA اندازه‌گیری شده نمونه‌های پرتقال و داده‌های طیفی پردازش شده با برخی شیوه‌های ترکیبی روش‌های پیش‌پردازش مختلف، از مدل‌سازی رگرسیون چندمتغیره خطی حداقل مربعات جزئی<sup>۱</sup> (PLS) بر پایه کاهش تعداد متغیرهای طیفی استفاده شد (Romía & Bernàrdez, 2009). در این راستا، پس از حذف نمونه‌های پرت<sup>۲</sup> از نظر داده‌های طیفی (۸ نمونه) - که ممکن است به دلیل مشکل‌های فنی مربوط به تجهیزات اسپکتروسکوپی، جمع‌آوری داده‌ها، نمونه‌گیری‌های نادرست و غیره به وجود آمده باشند- نمونه‌های باقی‌مانده به صورت تصادفی به دو دسته آموزش (حدود ۷۵ درصد نمونه‌ها) و آزمون (حدود ۲۵ درصد نمونه‌ها) به ترتیب برای تدوین مدل واسنجی و پیش‌بینی تقسیم و مدل‌سازی‌ها برای هر هشت شیوه ترکیبی

1- Partial Least Squares

3- Root Mean Square Error of Calibration

5- Correlation Coefficient of Calibration

2- Outliers

4- Root Mean Square Error of Prediction

6- Correlation Coefficient of Prediction

قطر استوایی بیشینه و وزن نمونه‌ها به ترتیب در محدوده‌های ۶/۲۹۰-۷/۶۴۰ سانتی‌متر و ۲۰۶/۶۸۰-۱۲۷/۸۱۰ گرم برای دسته آموزش و در محدوده‌های ۶/۴۵۰-۷/۶۱۰ سانتی‌متر و ۲۰۵/۳۰۰-۱۳۴/۵۲۰ گرم برای دسته آزمون قرار داشتند؛ به‌طور کلی نمونه‌ها از نظر مشخصات مورفولوژی متنوع بودند.

SSC پرتقال‌ها در دسته‌های آموزش و آزمون به ترتیب بین ۸/۹۰۰ تا ۱۳/۶۰۰ و ۹/۴۰۰ تا ۱۳/۸۰۰ درجه بریکس بود. همچنین، اسیدیت (TA) نمونه‌ها بین ۱/۱۴۶ تا ۲/۵۲۲ درصد در دسته آموزش و بین ۱/۳۱۸ تا ۲/۵۰۲ درصد در دسته آزمون بود.

اختلاف بین RMSEC و RMSEP نیز پس از آن مد نظر قرار گرفت. برای هر مدل نیز تعداد بهینه متغیرهای نهان<sup>۱</sup> (LV) به گونه‌ای انتخاب شد تا مدل کمترین RMSE را در پیش‌بینی پارامتر مورد نظر داشته باشد.

## نتایج و بحث

### مشخصات نمونه‌ها

جدول ۱ داده‌های آماری (کمینه، بیشینه، میانگین، و انحراف استاندارد (SD)) مشخصات مورفولوژی شامل قطر استوایی بیشینه و وزن، همچنین SSC و TA نمونه‌ها را در دسته‌های آموزش و آزمون (پس از حذف داده‌های پرت) نشان می‌دهد.

جدول ۱- داده‌های آماری نمونه‌های پرتقال تامسون در اسپکتروسکوپی NIR

دسته آزمون (۳۶ نمونه)				دسته آموزش (۱۰۶ نمونه)				
انحراف استاندارد	میانگین	بیشینه	کمینه	انحراف استاندارد	میانگین	بیشینه	کمینه	
۰/۳۳۷	۷/۰۸۰	۷/۶۱۰	۶/۴۵۰	۰/۳۰۷	۶/۹۸۷	۷/۶۴۰	۶/۲۹۰	قطر (سانتی‌متر)
۲۰/۰۶۷	۱۷۲/۲۲۶	۲۰۵/۳۰۰	۱۳۴/۵۲۰	۱۸/۶۲۴	۱۶۳/۱۹۲	۲۰۶/۶۸۰	۱۲۷/۸۱۰	وزن (گرم)
۱/۲۶۲	۱۱/۳۷۲	۱۳/۸۰۰	۹/۴۰۰	۱/۱۲۴	۱۱/۵۳۹	۱۳/۶۰۰	۸/۹۰۰	مواد جامد حل‌شدنی (درجه بریکس)
۰/۳۲۸	۱/۸۷۱	۲/۵۰۲	۱/۳۱۸	۰/۲۸۴	۱/۹۴۵	۲/۵۲۲	۱/۱۴۶	اسیدیت قابل تیتر کردن (درصد)

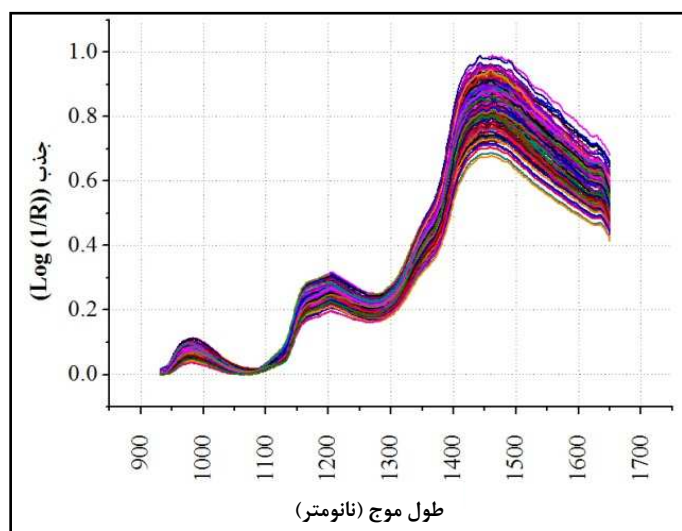
### مشخصات و تفسیر طیف‌ها

چگونگی توزیع اورتون<sup>۲</sup>های پیوندهای اصلی (Cen & He, 2007)، می‌توان آنها را به ترتیب ناشی از اورتون‌های دوم O-H و (یا) NH<sub>2</sub>؛ اورتون‌های دوم C-H، CH<sub>2</sub> و (یا) CH<sub>3</sub>؛ و اورتون‌های اول O-H و (یا) NH<sub>2</sub> دانست (Cen & He, 2007; Jamshidi et al., 2013).

شکل ۱، طیف‌های جذبی NIR همه نمونه‌ها را در محدوده ۹۳۰-۱۶۵۰ نانومتر نشان می‌دهد. مطابق این شکل، طیف‌ها یک روند افزایشی و سه پیک جذبی مشخص در اطراف طول موج‌های ۹۷۰؛ ۱۲۰۰؛ و ۱۴۵۰ نانومتر داشتند که با توجه به

1- Latent Variables

2- Overtone



شکل ۱- طیف‌های جذبی NIR همه نمونه‌های پرتقال تامسون

### پیش پردازش‌های طیفی

به طور کلی، شکل ظاهری طیف‌های جذبی NIR نمونه‌ها پس از هموارسازی با هر سه روش MA، SG، یا WT، مشابه یکدیگر و مشابه طیف‌های جذبی خام نمونه‌ها (شکل ۱) بود. روش‌های نرمال‌سازی MSC و SNV سبب نزدیک‌تر شدن طیف‌ها به یکدیگر و نمایان‌تر کردن تفاوت آنها در محل پیک‌های جذبی شدند. مشتق‌های اول و دوم طیف‌های جذبی (بر پایه الگوریتم SG) نیز ضمن تایید پیک‌های جذبی قبلی (قابل مشاهده در طیف‌های جذبی اصلی)، سبب افزایش قدرت تفکیک طیفی و نمایان‌شدن جزئیات جذبی بیشتر در اطراف طول موج‌های ۱۱۵۰؛ ۱۳۵۰ و ۱۶۴۰ نانومتر شدند که با توجه به چگونگی توزیع اورتون‌های پیوندهای اصلی (Cen & He, 2007)، می‌توان به ترتیب آنها را به اورتون‌های دوم C-H، CH<sub>2</sub> و (یا) CH<sub>3</sub>؛ اورتون‌های اول O-H و (یا) NH<sub>2</sub>؛ و اورتون‌های اول C-H، CH<sub>2</sub> و (یا) CH<sub>3</sub> نسبت داد. با این حال، مشتق‌گیری از طیف‌ها (به‌ویژه مشتق دوم) سبب تیز شدن

نویزها و کاهش نسبت سیگنال به نویز (به‌ویژه در طول موج‌های بالاتر از ۱۴۰۰ نانومتر) می‌شد. شکل ۲، طیف‌های جذبی NIR نمونه‌ها را پس از پردازش با هشت ترکیب مختلف روش‌های هموارسازی، نرمال‌سازی، و مشتق‌گیری شامل MA + MSC؛ MA + SNV؛ WT + MSC؛ SG + SNV؛ SG + MSC؛ MA + SNV؛ WT + SNV؛ SG + D<sub>1</sub>؛ و SG + D<sub>2</sub> نشان می‌دهد. روش‌های مختلف پیش‌پردازش به کار گرفته شده در این پژوهش به منظور ارائه مدل‌های واسنجی بهتر و قابل اعتمادتر برای پیش‌بینی SSC و TA پرتقال‌های تامسون بر پایه طیف‌های NIR، مقایسه شدند؛ این مقایسه پس از اعتبارسنجی مدل‌های رگرسیون چندمتغیره تدوین شده با هشت ترکیب مختلف روش‌های پیش‌پردازش استفاده شده، اجرا شد.

### آنالیزهای چندمتغیره

جدول‌های ۲ و ۳، نتایج واسنجی و پیش‌بینی مدل‌های PLS تدوین شده بر پایه ترکیب روش‌های

می‌دهند تا برای ترکیب روش‌های هموارسازی و نرمال‌سازی MSC. همچنین، روش هموارسازی WT نسبت به دو روش هموارسازی دیگر (MA و SG) و در ترکیب با هر روش نرمال‌سازی، نتایج ضعیف‌تری را در خصوص پیش‌بینی پارترهای SSC و TA نمونه‌ها ارائه کردند. برای پیش‌بینی SSC پرتقال‌ها، مشتق‌گیری اول ( $D_1$ ) در مقایسه با مشتق‌گیری دوم ( $D_2$ )، توانایی بهتری نشان داد. در حالی که نتایج پیش‌بینی TA نمونه‌ها با پیش‌پردازش  $D_2$  بهتر از پیش‌پردازش  $D_1$  بود. با این حال، نتایج پیش‌بینی هر دو پارامتر SSC و TA مدل‌های تدوین شده بر پایه ترکیب روش‌های پیش‌پردازش  $MA + SNV$ ،  $SG + MSC$  و  $SG + SNV$  در مقایسه با سایر مدل‌ها بهتر بود. از بین این سه مدل نیز و با توجه به تفاوت کمتر بین RMSEC و RMSEP، بهترین مدل PLS تدوین شده پیش‌بینی کننده هر دو پارامتر SSC و TA در پرتقال‌های تامسون بر پایه ترکیب روش‌های پیش‌پردازش  $MA + SNV$  در نظر گرفته شد.

مختلف پیش‌پردازش اعمال شده روی طیف‌های NIR را به ترتیب برای SSC و TA پرتقال‌های تامسون نشان می‌دهند.

با توجه به مقدارهای شاخص SDR، تمام مدل‌های رگرسیون چندمتغیره تدوین شده توانایی تشخیص و پیش‌بینی SSC و TA نمونه‌ها را داشتند ( $SDR > 1/5$ ) (Nicolai et al., 2007). با این حال، نتایج بسیار بهتری برای پیش‌بینی پارامتر SSC نمونه‌ها نسبت به TA به دست آمد که دلیل آن پایین بودن غلظت اسیدهای ارگانیک نسبت به مواد جامد حل‌شدنی (SSC) در نمونه‌های پرتقال است و با نتایج بررسی‌های دیگر محققان (Cayuela, 2008; Cayuela & Weiland, 2010; Liu et al., 2010b; Jamshidi et al., 2012c) همخوانی دارد.

اعتبارسنجی مدل‌ها بر اساس  $r_p$ ، RMSEP و SDR نشان داد که مدل‌های تدوین شده بر پایه ترکیب روش‌های هموارسازی و نرمال‌سازی SNV، نتایج پیش‌بینی بهتری برای هر دو پارامتر (TA و SSC) به دست

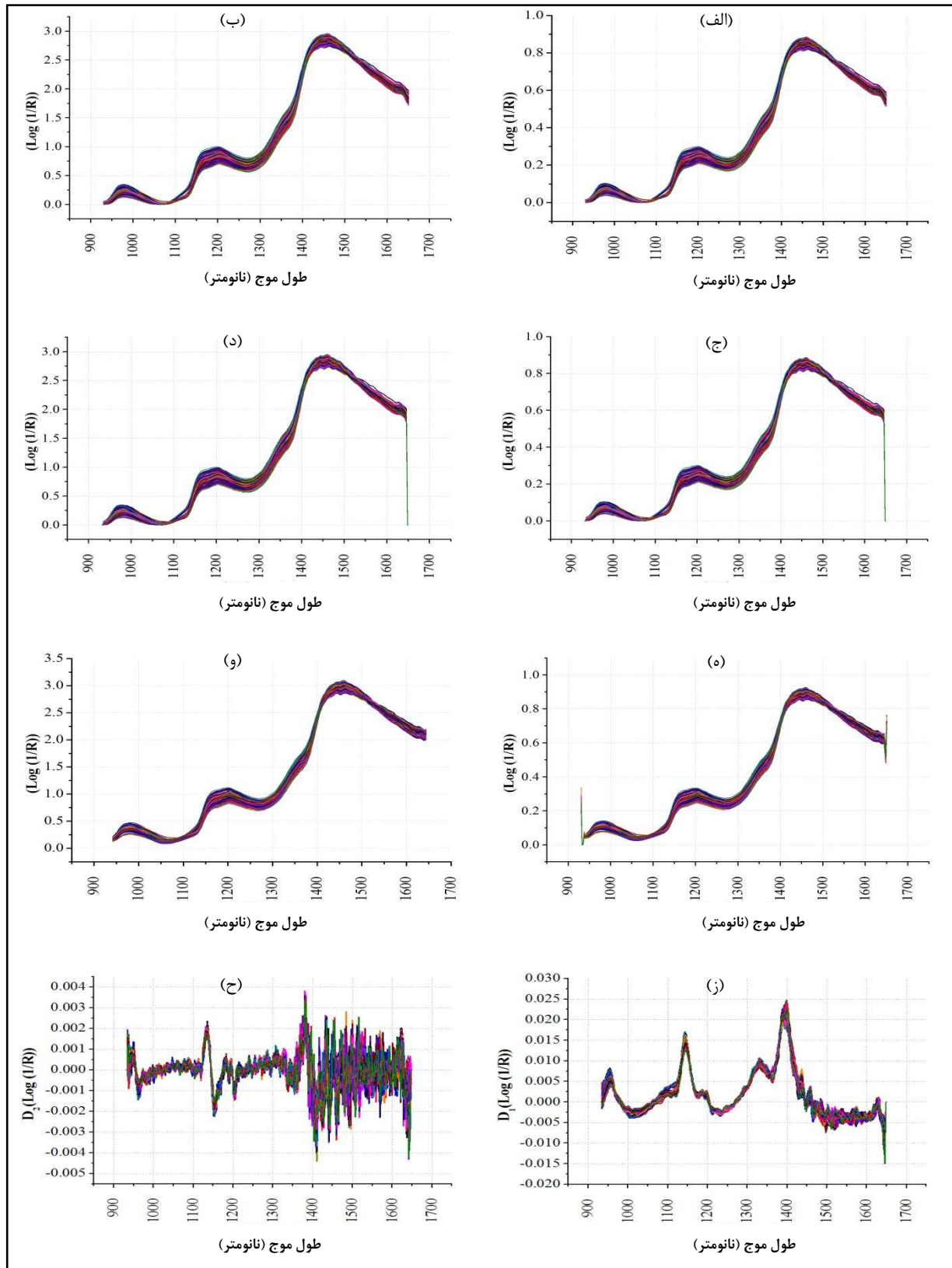
جدول ۲- نتایج واسنجی و پیش‌بینی مدل‌های PLS بر پایه ترکیب روش‌های مختلف پیش‌پردازش برای SSC پرتقال‌ها

پیش‌پردازش	LVs	RMSEC	$r_c$	RMSEP	$r_p$	SDR
MA + MSC	۸	۰/۴۷۲	۰/۹۰۷	۰/۴۶۳	۰/۹۳۱	۲/۷۲۶
MA + SNV	۹	۰/۴۳۰	۰/۹۲۳	۰/۴۵۱	۰/۹۳۶	۲/۷۹۸
SG + MSC	۹	۰/۴۰۳	۰/۹۳۳	۰/۴۶۱	۰/۹۳۳	۲/۷۳۸
SG + SNV	۹	۰/۴۰۰	۰/۹۳۴	۰/۴۵۷	۰/۹۳۴	۲/۷۶۱
SG + $D_1$	۱۱	۰/۲۷۱	۰/۹۷۰	۰/۵۳۹	۰/۹۱۱	۲/۳۴۱
SG + $D_2$	۸	۰/۳۶۵	۰/۹۴۵	۰/۷۲۶	۰/۸۲۰	۱/۷۳۸
WT + MSC	۱۲	۰/۳۵۳	۰/۹۴۹	۰/۴۹۸	۰/۹۲۹	۲/۵۳۴
WT + SNV	۱۲	۰/۳۵۳	۰/۹۴۹	۰/۴۹۶	۰/۹۳۰	۲/۵۴۴

عددهای پرتقال نشان‌دهنده بهترین مدل تدوین شده (با توجه به معیارهای اعتبارسنجی) برای پیش‌بینی SSC در پرتقال تامسون است.



بررسی اثر پیش پردازش‌های مختلف طیفی بر ارزیابی...



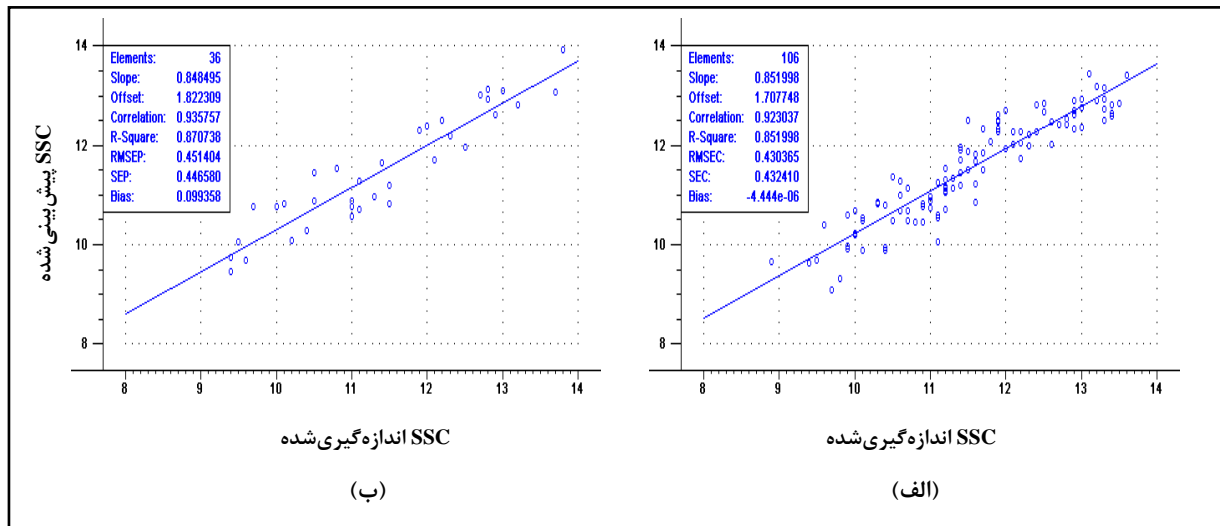
شکل ۲- طیف‌های جذبی NIR پردازش شده با ترکیب روش‌های (الف) MA + MSC، (ب) MA + SNV، (ج) SG + MSC، (د) SG + SNV، (ه) WT + MSC، (و) WT + SNV، (ز) SG + D<sub>1</sub>، (ح) SG + D<sub>2</sub>

جدول ۳- نتایج واسنجی و پیش‌بینی مدل‌های PLS بر پایه ترکیب روش‌های مختلف پیش‌پردازش برای TA پرتقال‌ها

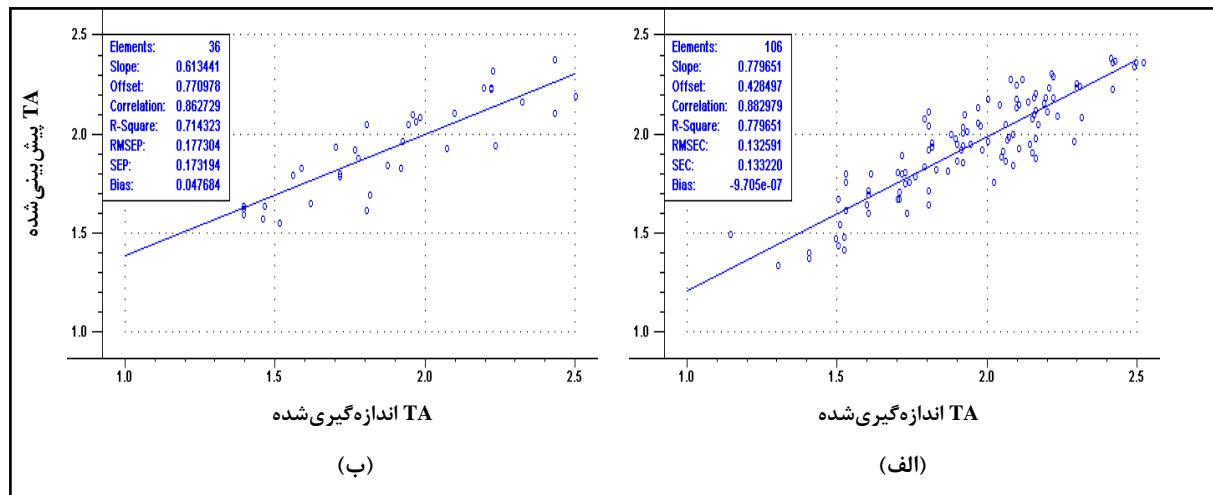
SDR	$r_p$	RMSEP	$r_c$	RMSEC	LVs	پیش‌پردازش
۱/۶۵۷	۰/۸۲۰	۰/۱۹۸	۰/۸۷۰	۰/۱۳۹	۹	MA + MSC
۱/۸۵۳	۰/۸۶۳	۰/۱۷۷	۰/۸۸۳	۰/۱۳۳	۱۰	MA + SNV
۱/۸۴۳	۰/۸۶۴	۰/۱۷۸	۰/۸۹۰	۰/۱۲۹	۱۰	SG + MSC
۱/۸۴۳	۰/۸۶۴	۰/۱۷۸	۰/۸۹۰	۰/۱۲۹	۱۰	SG + SNV
۱/۵۲۶	۰/۷۶۹	۰/۲۱۵	۰/۹۲۰	۰/۱۱۱	۹	SG + D <sub>1</sub>
۱/۷۶۳	۰/۸۲۸	۰/۱۸۶	۰/۹۴۲	۰/۰۹۵	۹	SG + D <sub>2</sub>
۱/۵۹۲	۰/۸۰۰	۰/۲۰۶	۰/۸۶۹	۰/۱۴۰	۱۰	WT + MSC
۱/۶۰۸	۰/۸۰۳	۰/۲۰۴	۰/۸۷۱	۰/۱۳۹	۱۰	WT + SNV

عددهای پرتنگ نشان‌دهنده بهترین مدل تدوین شده (با توجه به معیارهای اعتبارسنجی) برای پیش‌بینی TA در پرتقال تامسون است.

شکل‌های ۳ و ۴ مقدار پیش‌بینی‌شده پارامتر با مدل‌های واسنجی و پیش‌بینی تدوین شده بر پایه ترکیب روش‌های پیش‌پردازش MA + SNV را در مقابل مقدار اندازه‌گیری شده آن به ترتیب برای SSC و TA نمونه‌ها نشان می‌دهند.



شکل ۳- مقدار اندازه‌گیری شده SSC پرتقال‌های تامسون در مقابل مقدار پیش‌بینی شده آن با مدل‌های PLS (الف) واسنجی، (ب) پیش‌بینی تدوین شده بر پایه ترکیب روش‌های پیش‌پردازش MA + SNV



شکل ۴- مقدار اندازه‌گیری شده TA پرتقال‌های تامسون در مقابل مقدار پیش‌بینی شده آن با مدل‌های PLS (الف) و اسنجی، (ب) پیش‌بینی تدوین شده بر پایه ترکیب روش‌های پیش‌پردازش MA + SNV

از چهار وارسته پرتقال گزارش داده‌اند. از سوی دیگر، نتایج این پژوهش برای پیش‌بینی SSC از نظر ضریب همبستگی  $r_p$  و خطاهای RMSEC، RMSEP ضعیف‌تر و از نظر ضریب همبستگی  $r_c$  و شاخص SDR بهتر از نتایجی است که جمشیدی و همکاران (Jamshidi *et al.*, 2012c) در مد بازتابی و محدوده طیفی ۱۰۰۰-۴۰۰ نانومتر برای پرتقال والنسیا (RMSEC=۰/۳۲۸،  $r_c=۰/۹۱۴$ ، RMSEP=۰/۳۱) و اسنجی (RMSEC=۰/۹۵۸،  $r_p=۰/۹۵۸$ ، SDR=۲/۷۶۱) با مدل PLS بر پایه ترکیب پیش‌پردازش‌های SG + MSC ارائه داده‌اند. همچنین، نتایج پیش‌بینی SSC بهتر از نتایجی است که لیو و همکاران (Liu *et al.*, 2010a) برای پرتقال ناول با مدل غیرخطی شبکه عصبی پس‌انتشار-آنالیز مولفه‌های اصلی<sup>۲</sup> (PCA-BPNN) تدوین شده بر پایه نرمال‌سازی MSC در مد بازتابی و محدوده طیفی ۱۸۰۰-۳۵۰ نانومتر (RMSEC=۰/۴۵۰،  $r_c=۰/۹۴۰$ ، RMSEP=۰/۶۸۰) گزارش داده‌اند، اگرچه ضریب همبستگی  $r_c$  اندکی ضعیف‌تر از نتایجی است که این محققان به دست آورده‌اند. نتایج این پژوهش بسیار بهتر از نتایج پژوهش‌های لیو و همکاران (Liu *et al.*, 2010b, c) است

به‌طور کلی، نتایج پیش‌بینی SSC برای پرتقال تامسون با مدل PLS تدوین شده بر پایه ترکیب روش‌های پیش‌پردازش MA + SNV (RMSEC=۰/۴۳۰،  $r_c=۰/۹۲۳$ ، RMSEP=۰/۴۵۱) و طیف‌های NIR بازتابی در محدوده طیفی ۱۶۵۰-۹۳۰ نانومتر از نظر خطای RMSEP بهتر و از نظر ضریب تبیین  $r_c^2$  اندکی ضعیف‌تر از نتایجی است که کایونلا (Cayuela, 2008) با مدل PLS تدوین شده بر پایه ترکیب روش‌های پیش‌پردازش SNV +  $D_2$  در مد بازتابی و محدوده طیفی ۱۸۵۰-۵۸۰ نانومتر برای پرتقال والنسیا (RMSEP=۰/۵۱۳،  $r_c^2=۰/۹۱۳$ ) ارائه داده است. همچنین نتایج پیش‌بینی SSC به دست آمده در این پژوهش بسیار بهتر از نتایجی است که کایونلا و ویلند (Cayuela & Weiland, 2010) با مدل‌های PLS تدوین شده بر پایه نرمال‌سازی میانگین<sup>۱</sup> (MN) در مد بازتابی و محدوده‌های طیفی ۲۳۰۰-۵۰۰ نانومتر ( $r_c=۰/۹۲۰$ ، RMSEP=۰/۷۴۰، SDR=۲/۱۳۰) و ۱۱۰۰ نانومتر ( $r_c=۰/۹۱۰$ ، RMSEP=۰/۶۸۰، SDR=۲/۳۳۰) با اسپکتروفتومترهای متفاوت برای ترکیبی

که یکی برای نارنگی نانفنگ<sup>۱</sup> در مد عبوری و با مدل PLS تدوین شده روی طول موج های موثر محدوده های ۸۱۷/۵۷-۸۳۶/۱۹، ۹۰۹/۸۵-۹۲۷/۶۰ و ۹۴۵/۵۲-۹۶۲/۷۵ نانومتر بدون اعمال هیچ پیش پردازشی (RMSEP=۰/۶۵۰،  $r_p=0/920$ )، و دیگری برای پرتقال ناول با مدل Spline-PLS تدوین شده بر پایه نرمال سازی MSC در مد بازتابی و محدوده طیفی ۴۵۰-۱۷۵۰ نانومتر (RMSEP=۰/۴۷۰،  $r_p=0/870$ ، SDR=۲/۳۴۰) ارائه داده اند. نتایج به دست آمده در این پژوهش از نظر ضریب همبستگی  $r_p$  بسیار بهتر و از نظر خطای RMSEP ضعیف تر از نتایجی است که لیو و همکاران

(Liu *et al.*, 2012) برای پرتقال ناول با مدل غیرخطی رگرسیون بردار پشتیبان حداقل مربعات<sup>۲</sup> (LSSVR) تدوین شده بر پایه پیش پردازش WT در مد برهم کنش<sup>۳</sup> و محدوده طیفی ۸۲۰-۹۵۰ نانومتر (RMSEP=۰/۳۲۱،  $r_p=0/918$ ) گزارش داده اند.

جدول ۴، نتایج پژوهش حاضر و نتایج گزارش شده دیگر پژوهشگران را درباره پیش بینی SSC در مرکبات نشان می دهد. با این حال، به دلیل متفاوت بودن محدوده های طیفی، نوع اسپکترومترها، تفاوت در مدهای اندازه گیری، همچنین نوع و وارسته محصول مورد بررسی، مقایسه نتایج پیشنهاد نمی شود.

جدول ۴- نتایج پژوهش حاضر و سایر پژوهش های انجام شده برای پیش بینی SSC در مرکبات بر پایه اسپکتروسکوپی NIR و مدل های PLS

مرجع	محصول	محدوده طیفی (نانومتر)	مد اندازه گیری	پیش پردازش	$r_c^2$	RMSEC	$r_c$	RMSEP	$r_p$	SDR
پژوهش حاضر	پرتقال تامسون	۹۳۰-۱۶۵۰	بازتابی	MA + SNV	۰/۸۵۲	۰/۴۳۰	۰/۹۲۳	۰/۴۵۱	۰/۹۳۶	۲/۷۹۸
(Cayuela, 2008)	پرتقال والنسیا	۵۸۰-۱۸۵۰	بازتابی	D <sub>2</sub> + SNV	۰/۹۱۳	-	-	۰/۵۱۳	-	-
(Cayuela & Weiland, 2010)	۴ وارسته مختلف پرتقال	۵۰۰-۲۳۰۰	بازتابی	MN	-	-	۰/۹۲۰	۰/۷۴۰	-	۲/۱۳۰
(Cayuela & Weiland, 2010)	۴ وارسته مختلف پرتقال	۱۱۰۰-۲۳۰۰	بازتابی	MN	-	-	۰/۹۱۰	۰/۶۸۰	-	۲/۳۳۰
(Jamshidi <i>et al.</i> , 2012c)	پرتقال والنسیا	۴۰۰-۱۰۰۰	بازتابی	SG + MSC	-	۰/۳۲۸	۰/۹۱۴	۰/۳۳۱	۰/۹۵۸	۲/۷۶۱
(Liu <i>et al.</i> , 2010a)	پرتقال ناول	۳۵۰-۱۸۰۰	بازتابی	MSC	-	۰/۴۵۰	۰/۹۴۰	۰/۶۸۰	۰/۹۰۰	-
(Liu <i>et al.</i> , 2010b)	نارنگی نانفنگ	۸۱۷/۵۷-۸۳۶/۱۹ ۹۰۹/۸۵-۹۲۷/۶۰ ۹۴۵/۵۲-۹۶۲/۷۵	عبوری	بدون پیش پردازش	-	-	-	۰/۶۵۰	۰/۹۲۰	-
(Liu <i>et al.</i> , 2010c)	پرتقال ناول	۴۵۰-۱۷۵۰	بازتابی	MSC	-	-	-	۰/۴۷۰	۰/۸۷۰	۲/۳۴۰
(Liu <i>et al.</i> , 2012)	پرتقال ناول	۸۲۰-۹۵۰	برهم کنش	WT	-	-	-	۰/۳۲۱	۰/۹۱۸	-

«-» نشان دهنده مواردی است که پژوهشگران گزارش نداده اند.

از چهار وارسته پرتقال گزارش داده‌اند. از سوی دیگر، نتایج این پژوهش برای پیش‌بینی TA از نظر ضرایب همبستگی  $r_c$  و  $r_p$  بهتر و از نظر خطاهای RMSEC، RMSEP و شاخص SDR اندکی ضعیف‌تر از نتایجی است که جمشیدی و همکاران (Jamshidi et al., 2012c) در مد بازتابی و محدوده طیفی ۱۰۰۰-۴۰۰ نانومتر برای پرتقال والنسیا ( $r_c=0/093$ , RMSEC= $0/0786$ ,  $r_p=0/070$ , RMSEP= $0/070$ ) و با مدل PLS بر پایه ترکیب پیش‌پردازش‌های SG + MSC ارائه داده‌اند. همچنین، نتایج این پژوهش از نظر ضریب همبستگی  $r_p$  بسیار بهتر و از نظر خطای RMSEP اندکی ضعیف‌تر از نتایجی است که لیو و همکاران (Liu et al., 2010b) برای پیش‌بینی TA نارنگی نانفنگ در مد عبوری و با مدل PLS تدوین شده روی طول موج‌های موثر محدوده‌های ۸۳۶/۱۹-۸۱۷/۵۷، ۹۲۷/۶۰-۹۰۹/۸۵ و ۹۶۲/۷۵-۹۴۵/۵۲ نانومتر بدون اعمال هیچ پیش‌پردازشی ( $r_p=0/064$ , RMSEP= $0/09$ ) به دست آورده‌اند. جدول ۵، نتایج پژوهش حاضر و نتایج گزارش‌های دیگر پژوهشگران را در خصوص پیش‌بینی TA در مرکبات نشان می‌دهد. به‌طور کلی و همان‌گونه که پیش از این نیز اشاره شد، به‌دلیل متفاوت بودن محدوده‌های طیفی، نوع اسپکترومترها، تفاوت در مدهای اندازه‌گیری، همچنین نوع و وارسته محصول مورد بررسی، مقایسه نتایج پیشنهاد نمی‌شود.

همچنین نتایج پیش‌بینی TA به‌دست آمده در این پژوهش با مدل PLS تدوین شده بر پایه ترکیب روش‌های پیش‌پردازش MA + SNV ( $r_c=0/883$ , RMSEC= $0/133$ ) و طیف‌های NIR بازتابی در محدوده طیفی ۱۶۵۰-۹۳۰ نانومتر بسیار بهتر از نتایجی است که گوتیه و همکاران (Guthrie et al., 2005) با مدل‌های PLS تدوین شده بر پایه ترکیب روش‌های مختلف مشتق‌گیری ( $D_1$  و  $D_2$ ) و نرمال‌سازی (SNV و detrend) در مد برهم‌کنش و محدوده طیفی ۹۵۰-۷۲۰ نانومتر برای پیش‌بینی TA در نارنگی امپریال<sup>۱</sup> ( $r_c^2=0/3$ , RMSEC= $0/2$ ) به‌دست آورده‌اند. نتایج پیش‌بینی TA در این پژوهش همچنین بهتر از نتایجی است که کایوئلا (Cayuela, 2008) با مدل PLS تدوین شده بر پایه ترکیب روش‌های پیش‌پردازش  $D_2 + SNV$  در مد بازتابی و محدوده طیفی ۱۸۵۰-۵۸۰ نانومتر برای پرتقال والنسیا ( $r_c^2=0/637$ , RMSEP= $0/331$ ) به‌دست آورده‌اند و نیز بهتر از نتایجی است که کایوئلا و ویلند (Cayuela & Weiland, 2010) با مدل‌های PLS تدوین شده بر پایه MN در مد بازتابی و محدوده‌های طیفی ۲۳۰۰-۵۰۰ نانومتر ( $r_c=0/86$ , RMSEP= $0/117$ , SDR= $1/69$ ) و ۲۳۰۰-۱۱۰۰ نانومتر ( $r_c=0/80$ , RMSEP= $0/19$ , SDR= $1/54$ ) با اسپکترومترهای متفاوت برای ترکیبی

جدول ۵- نتایج پژوهش حاضر و سایر پژوهش‌های انجام شده برای پیش‌بینی TA در مرکبات بر پایه اسپکتروسکوپی NIR و مدل‌های PLS

مرجع	محصول	محدوده طیفی (نانومتر)	مد اندازه‌گیری	پیش‌پردازش	$r_c^2$	RMSEC	$r_c$	RMSEP	$r_p$	SDR
پژوهش حاضر	پرتقال تامسون	۹۳۰-۱۶۵۰	بازتابی	MA + SNV	۰/۷۸۰	۰/۱۳۳	۰/۸۸۳	۰/۱۷۷	۰/۸۶۳	۱/۸۵۳
(Guthrie <i>et al.</i> , 2005)	نارنگی امپریال	۷۲۰-۹۵۰	برهم‌کنش	.SNV ,D <sub>2</sub> ,D <sub>1</sub> detrend	۰/۳۰۰	۰/۲۰۰	-	-	-	-
(Cayuela, 2008)	پرتقال والنسیا	۵۸۰-۱۸۵۰	بازتابی	D <sub>2</sub> + SNV	۰/۶۳۷	-	-	۰/۳۳۱	-	-
(Cayuela & Weiland, 2010)	۴ واریته مختلف پرتقال	۵۰۰-۲۳۰۰	بازتابی	MN	-	-	۰/۸۶۰	۰/۱۷۰	-	۱/۶۹۰
(Cayuela & Weiland, 2010)	۴ واریته مختلف پرتقال	۱۱۰۰-۲۳۰۰	بازتابی	MN	-	-	۰/۸۰۰	۰/۱۹۰	-	۱/۵۴۰
(Jamshidi <i>et al.</i> , 2012c)	پرتقال والنسیا	۴۰۰-۱۰۰۰	بازتابی	SG + MSC	-	۰/۰۹۳	۰/۷۸۶	۰/۰۷۰	۰/۸۵۸	۱/۹۲۸
(Liu <i>et al.</i> , 2010b)	نارنگی نانفنگ	۸۱۷/۵۷-۸۳۶/۱۹ ۹۰۹/۸۵-۹۲۷/۶۰ ۹۴۵/۵۲-۹۶۲/۷۵	عبوری	بدون پیش‌پردازش	-	-	-	۰/۰۹۰	۰/۶۴۰	-

«-» نشان‌دهنده مواردی است که پژوهشگران گزارش نداده‌اند.

### نتیجه‌گیری

ترکیب روش‌های مختلف پیش‌پردازش طیفی نیز نشان داد که روش‌های پیش‌پردازش اثر مستقیم بر نتایج پیش‌بینی مدل‌ها دارند و بهترین نتایج پیش‌بینی بر اساس مدل PLS تدوین شده بر پایه ترکیب روش‌های پیش‌پردازش MA + SNV برای هر دو پارامتر SSC (RMSEC=۰/۴۳۰،  $r_c$ =۰/۹۲۳، RMSEP=۰/۴۵۱،  $r_p$ =۰/۹۳۶، SDR=۲/۷۹۸) و TA (RMSEC=۰/۱۳۳) و (SDR=۱/۸۵۳،  $r_p$ =۰/۸۶۳، RMSEP=۰/۱۷۷،  $r_c$ =۰/۸۳) به‌دست آمد.

نتایج آنالیزهای چندمتغیره به‌منظور بررسی امکان تشخیص غیر مخرب SSC و TA پرتقال تامسون با اسپکتروسکوپی NIR نشان داد که از طیف‌های بازتابی NIR (در محدوده طیفی ۹۳۰-۱۶۵۰ نانومتر) و مدل‌سازی رگرسیون چندمتغیره خطی PLS می‌توان برای تشخیص و پیش‌بینی این پارامترهای کیفی استفاده کرد. اعتبارسنجی مدل‌های PLS تدوین شده بر پایه

## مراجع

- Anon, 2005. OOIBase32 Spectrometer Operating Software, Instalation and Operation Manual. Ocean Optics, Inc. USA.
- Bokobza, L. 2006. Origin of Near-Infrared Absorption Bands. In: Siesler, H. W., Ozaki, Y., Kawata, S. and Heise, H. M. (Eds.) Near-Infrared Spectroscopy: Principles, Instruments, Applications. 3<sup>rd</sup> Reprint. Wiley-VCH. Germany.
- Camps, C. and Christen, D. 2009. Non-destructive assessment of apricot fruit quality by portable visible-near infrared spectroscopy. LWT-Food Sci. Technol. 42, 1125-1131.
- Cayuela, J. A. 2008. VIS/NIR soluble solids prediction in intact oranges (*Citrus sinensis* L.) cv. Valencia Late by reflectance. Postharvest Biol. Technol. 47, 75-80.
- Cayuela, J. A. and Weiland, C. 2010. Intact orange quality prediction with two portable NIR spectrometers. Postharvest Biol. Technol. 58(2): 113-120.
- Cen, H. and He, Y. 2007. Theory and application of near infrared reflectance spectroscopy in determination of food quality. Trends Food Sci. Technol. 18, 72-83.
- Fu, X., Yan, G., Chen, B. and Li, H. 2005. Application of wavelet transforms to improve prediction precision of near infrared spectra. J. Food Eng. 69, 461-466.
- Fu, X., Ying, Y., Lu, H., Xu, H. and Yu, H. 2007. FT-NIR diffuse reflectance spectroscopy for kiwifruit firmness detection. Sens. Instrumen. Food Qual. 1, 29-35.
- Guthrie, J. A., Walsh, K., Reid, D. J. and Liebenseberg, C. J. 2005. Assessment of internal quality attributes of mandarin fruit. 1. NIR calibration model development. Aust. J. Agric. Res. 56, 405-416.
- Heise, H. M. and Winzen, R. 2006. Chemometrics in Near-Infrared Spectroscopy. In: Siesler, H. W., Ozaki, Y., Kawata, S. and Heise, H. M. (Eds.) Near-Infrared Spectroscopy: Principles, Instruments, Applications. 3<sup>rd</sup> Reprint. Wiley-VCH. Germany.
- James, C. S. 1998. Analytical Chemistry of Foods. Springer. USA.
- Jamshidi, B., Minaei, S., Mohajerani, E. and Ghassemian, H. 2011a. Analysis of citrus peel for non-destructive determination of fruit composition by reflectance Vis/NIR spectroscopy. Proceeding of the XXXIV CIOSTA CIGR V Conference on Efficient and Safe Production Processes in Sustainable Agriculture and Forestry. June 29 - July 1. Vienna. Austria.
- Jamshidi, B., Minaei, S., Mohajerani, E. and Ghassemian, H. 2012a. Multivariate analysis of reflectance Vis/NIR spectra based on wavelet transform for non-destructive and detection of orange color and pH simultaneously. Proceeding of the 7<sup>th</sup> National Congress on Agricultural Machinery Engineering and Mechanization. Sep. 4-6. Shiraz. Iran. (in Farsi)
- Jamshidi, B., Minaei, S., Mohajerani, E. and Ghassemian, H. 2012b. Vis/NIR spectroscopy for non-destructive classification of orange varieties. Proceeding of the 7<sup>th</sup> National Congress on Agricultural Machinery Engineering and Mechanization. Sep. 4-6. Shiraz. Iran. (in Farsi)

- Jamshidi, B., Minaei, S., Mohajerani, E. and Ghassemian, H. 2012c. Reflectance Vis/NIR spectroscopy for nondestructive taste characterization of valencia oranges. *Comput. Electron. Agric.* 85, 64-69.
- Jamshidi, B., Minaei, S., Mohajerani, E. and Ghassemian, H. 2013. Linear multivariate model based on NIR spectroscopy for non-destructive internal quality prediction of orange. *Proceeding of the 19<sup>th</sup> Iranian Conference on Optics and Photonics, and 5<sup>th</sup> Iranian Conference on Photonics Engineering.* Jan. 22-24. Zahedan. Iran. (in Farsi)
- Jamshidi, B., Minaei, S., Mohajerani, E. and Ghassemian, H. 2014. Prediction of soluble solids in oranges using visible/near-infrared spectroscopy: Effect of peel. *Int. J. Food Prop.* 17, 1460-1468.
- Jamshidi, B., Minaei, S., Mohajerani, E., Ghassemian, H. and Afkhami Ardakani, H. 2011b. Reflectance spectra analysis of citrus by Vis/NIR spectroscopy for nondestructive determining of inner chemical compositions. *Proceeding of the 17<sup>th</sup> Iranian Conference on Optics and Photonics, and 3<sup>rd</sup> Iranian Conference on Photonics Engineering.* Feb. 8-10. Kerman. Iran. (in Farsi)
- Liu, Y., Sun, X. and Ouyang, A. 2010a. Nondestructive measurement of soluble solids content of navel orange fruit by visible-NIR spectrometric technique with PLSR and PCA-BPNN. *LWT-Food Sci. Technol.* 43(4): 602-607.
- Liu, Y., Sun, X., Zhang, H. and Ouyang, A. 2010b. Nondestructive measurement of internal quality of Nanfeng mandarin fruit by charge coupled device near infrared spectroscopy. *Comput. Electron. Agric.* 71, s10-s14.
- Liu, Y., Sun, X., Zhou, J., Zhang, H. and Yang, C. 2010c. Linear and nonlinear multivariate regressions for determination sugar content of intact Gannan navel orange by Vis-NIR diffuse reflectance spectroscopy. *Math. Comput. Model.* 51(11-12): 1438-1443.
- Liu, Y., Gao, R., Hao, Y., Sun, X. and Ouyang, A. 2012. Improvement of near-infrared spectral calibration models for brix prediction in 'gannan' navel oranges by a portable near-infrared device. *Food Bioprocess Technol.* 5(3): 1106-1112.
- Magwaza, L., Opara, U., Nieuwoudt, H., Cronje, P., Saeys, W. and Nicolai, B. 2012. NIR spectroscopy applications for internal and external quality analysis of citrus fruit-A review. *Food Bioprocess Technol.* 5(2): 425-444.
- Mireei, S. A., Mohtasebi, S. S., Massudi, R., Rafiee, S. and Arabanian, A. S. 2010. Feasibility of near infrared spectroscopy for analysis of date fruits. *Int. Agrophysics.* 24, 351-356.
- Moghimi, A., Aghkhani, M. H., Sazgarnia, A. and Sarmad, M. 2010. Vis/NIR spectroscopy and chemometrics for the prediction of soluble solids content and acidity (pH) of kiwifruit. *Biosyst. Eng.* 106(3): 295-302.
- Nicolai, B. M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K. I. and Lammertyn, J. 2007. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biol. Technol.* 46, 99-118.



- Pissard, A., Fernández-Pierna, J. A., Baeten, V., Sinnaeve, G., Lognay, G., Mouteau, A., Dupont, P., Rondia, A. and Lateur, M. 2013. Non-destructive measurement of vitamin C, total polyphenol and sugar content in apples using near-infrared spectroscopy. *J. Sci. Food Agric.* 93(2): 238-244.
- Romía, M. B. and Bernàrdez, M. A. 2009. Multivariate Calibration for Quantitative Analysis. In: Sun, D. W. (Ed.) *Infrared Spectroscopy for Food Quality Analysis and Control*. 1<sup>st</sup> Ed. Academic Press.
- Salguero-Chaparro, L., Baeten, V., Fernández-Pierna, J. A. and Peña-Rodríguez, F. 2013. Near infrared spectroscopy (NIRS) for on-line determination of quality parameters in intact olives. *Food Chem.* 139(1-4): 1121-1126.
- Wang, J., Nakano, K. and Ohashi, S. 2011. Nondestructive evaluation of jujube quality by visible and near-infrared spectroscopy. *LWT-Food Sci. Technol.* 44, 1119-1125.
- Xia, J., Li, X., Li, P., Ma, Q. and Ding, X. 2007. Application of wavelet transform in the prediction of navel orange vitamin C content by near-infrared spectroscopy. *Agric. Sci. Chin.* 6(9): 1067-1073.
- Zhu, D., Ji, B., Meng, C., Shi, B., Tu, Z. and Qing, Z. 2007. Study of wavelet denoising in apple's charge-coupled device near-infrared spectroscopy. *J. Agric. Food Chem.* 55(14): 5423-5428.

## **Effect of Spectral Pre-Processing Methods on Non-Destructive Quality Assessment of Oranges Using NIRS**

**B. Jamshidi, S. Minaei\* , E. Mohajerani and H. Ghassemian**

\* Corresponding Author: Professor, Faculty of Agriculture, Tarbiat Modares University, Tehran, Iran. P. O. Box: 14115-336.  
E-mail: minae@modares.ac.ir

Received: 26 November 2013, Accepted: 9 August 2014

Near-infrared spectroscopy (NIRS) combined with chemometric methods of spectral pre-processing and multivariate modeling is popular non-destructive method for measuring quality attributes of fruits and vegetables. The present study investigated the feasibility of reflectance NIRS in a spectral range of 930-1650 nm for non-destructive prediction of soluble solids content (SSC) and titratable acidity (TA) in Thomson oranges. The effect of spectral pre-processing methods on the accuracy of multivariate predictor models was also assessed. Partial least squares (PLS) multivariate calibration models were developed using the reference measurements and pre-processed spectra. The following methods were used: smoothing (moving average (MA), Savitzky-Golay (SG), wavelet transform (WT)); normalizing (multiplicative scatter correction (MSC), standard normal variate (SNV)); and increasing the spectral resolution (the first and second derivatives ( $D_1$ ,  $D_2$ )) to predict the SSC and TA of oranges. The results indicate that reflectance NIRS plus chemometrics gives the potential for non-destructive prediction of SSC and TA in oranges. Pre-processing methods directly affected the results of the PLS models. The best prediction results for SSC (RMSEC = 0.430,  $r_c$  = 0.923, RMSEP = 0.451,  $r_p$  = 0.936, SDR = 2.798) and TA (RMSEC = 0.133,  $r_c$  = 0.883, RMSEP = 0.177,  $r_p$  = 0.863, SDR = 1.853) were achieved based on a combination of the MA + SNV pre-processing methods.

**Keywords:** Multivariate analysis, Near-infrared spectroscopy, Non-destructive, Pre-processing methods