

## به کارگیری روشهای پیش‌پردازش داده‌های حاصل از طیف سنجی مرئی / مادون قرمز نزدیک در تبیین مدلی برای پیشگویی ماده جامد محلول و اسیدیته لیموترش

امیر جاجرمی<sup>۱</sup>، مسعود تقی زاده<sup>۲\*</sup>، آمنه سازگارنیا<sup>۳</sup>، خلیل بهزاد<sup>۴</sup>

۱- دانش آموخته کارشناسی ارشد، گروه علوم و صنایع غذایی، دانشگاه فردوسی مشهد

۲- استادیار گروه علوم و صنایع غذایی، دانشگاه فردوسی مشهد

۳- دانشیار گروه فیزیک پزشکی، دانشگاه علوم پزشکی مشهد

۴- استادیار گروه علوم و صنایع غذایی، دانشگاه فردوسی مشهد

(تاریخ دریافت: ۹۳/۴/۳۰ تاریخ پذیرش: ۹۳/۶/۲۶)

### چکیده

ارزیابی خصوصیات کیفی از دوره نگهداری تا مصرف از مهمترین فاکتورهای کاهش ضایعات محصولات غذایی می‌باشد. بدین منظور در سالیان اخیر تحقیقات بسیاری برای توسعه روش‌هایی سریع و در عین حال غیر مخرب، توسعه یافته است. در این پژوهش توانمندی روش طیف سنجی مرئی/مادون قرمز نزدیک به عنوان روشی غیر مخرب و در عین حال سریع به منظور تعیین خصوصیات کیفی لیموترش شامل میزان ماده جامد خشک محلول و اسیدیته در حالت انعکاسی و در محدوده ۴۰۰ تا ۱۰۰۰ نانومتر مورد ارزیابی قرار گرفت و کاربرد روش‌های پیش‌پردازش و اصلاح طیف از جمله تبدیل متغیر نرمال استاندارد، تصحیح پخش افزاینده، فیلتر میانه و ساویتزکی گولی، مشتق‌گیری بر روی داده‌های طیفی به منظور تبیین مدلی با صحت و دقت بالا برای پیشگویی خصوصیات کیفی بدون آسیب به ساختار لیموترش مورد بررسی قرار گرفت. مدل پیشگویی با استفاده از روش حداقل مربعات جزئی تبیین گردید. ضریب تعیین و ریشه مربعات خطای پیشگویی برای ماده جامد خشک محلول و اسیدیته به ترتیب ۰/۹۹۴، ۰/۱۰۵ درجه بریکس و ۰/۹۰۹، ۰/۱۱۸ حاصل گردید. نتایج حاصل از این پژوهش بیانگر اهمیت روش‌های پیش‌پردازش داده‌های حاصل از طیف سنجی مرئی/مادون قرمز نزدیک در تبیین مدلی با صحت بالا برای پیش‌بینی خصوصیات کیفی درونی لیموترش می‌باشد.

کلید واژه‌گان: خصوصیات کیفی، طیف سنجی مرئی/مادون قرمز نزدیک، لیموترش، روشهای پیش‌پردازش، مدل‌های پیشگویی.

\* مسئول مکاتبات: mtaghizadeh@um.ac.ir

## ۱- مقدمه

مرکبات یکی از با صرفه‌ترین میوه‌ها برای تولیدکنندگان آن می‌باشد. مرکبات امروزه در جهان جنبه صنعتی بسیار مهمی را دارا بوده، منبع پر درآمدی برای کشورهای تولیدکننده آن می‌باشند. کشت اقتصادی مرکبات در ایران از ۳۰۰ سال پیش در شمال کشور آغاز شده و به جنوب کشور رفته است. در میان مرکبات لیموترش یکی از مهم‌ترین و اقتصادی‌ترین محصولات باغی در جنوب ایران است. لیموترش در ایران در سطح وسیعی معادل ۴۱۸۰۰ هکتار، کشت می‌شود و استان هرمزگان با ۷۰ درصد کل تولید، قطب لیموترش در ایران می‌باشد. بر اساس آمار سازمان خواربار و کشاورزی ملل متحد<sup>۱</sup>، ایران با تولید ۶۱۵۰۰۰ تن در میان ۱۰ کشور برتر تولید لیموترش در جهان قرار دارد. از بین رفتن بخش زیادی از محصولات کشاورزی و اتلاف آن‌ها چرخه تولید تا مصرف امری رایج می‌باشد، به طوری که به طور متوسط ۳۵ درصد از محصولات کشاورزی در ایران و در مراحل گوناگون چرخه تولید تا مصرف ضایع می‌شوند و این مقدار غذای ۱۵ تا ۲۰ میلیون نفر از جمعیت کشور است. ضایعات کشاورزی در سه مرحله پیش از برداشت، برداشت و پس از آن روی می‌دهد اما قسمت عمده ضایعات مربوط به مراحل برداشت و توزیع محصولات می‌باشد. از این رو ارزیابی کیفی محصولات کشاورزی به منظور کنترل شرایط برداشت و نگهداری به منظور کاهش ضایعات امری ضروری می‌باشد. روش‌های متعددی تاکنون برای کیفیت سنجی غیر مخرب محصولات کشاورزی ابداع شده‌اند، که تنها برخی از آن‌ها توانسته از لحاظ فنی و صنعتی توجیه داشته باشند. روش طیف سنجی مرئی/مادون قرمز نزدیک از جمله روش‌هایی می‌باشد که با توجه به سرعت و دقت بالا به منظور ارزیابی خصوصیات کیفی محصولات کشاورزی در چند سال اخیر مورد توجه قرار گرفته است که می‌توان به ارزیابی خصوصیات کیفی در مرکبات [۳، ۴، ۶، ۷، ۸]، کیوی [۹]، سیب [۱۰]، گلابی [۱۱]، گوجه فرنگی [۱۲] و ... اشاره نمود. بر اساس قانون بیر، بین میزان جذب و غلظت رابطه‌ای خطی برقرار است بسیاری از پدیده‌های فیزیکی و شیمیایی منجر به انحراف از رابطه خطی بین میزان غلظت و جذب می‌گردد. به عنوان مثال میزان

پراکنش در طیف‌های حاصل از میوه‌ها و سبزیجات دست نخورده<sup>۲</sup> تحت تاثیر خصوصیات فیزیکی محصول از جمله اندازه، شکل، ضخامت پوسته، جنس پوسته، میزان رطوبت و غیره می‌باشد [۱۳] که منجر به ایجاد نویز و پهن شدگی نواحی طول موج<sup>۳</sup> و همچنین تغییرات از خط مبنا<sup>۴</sup> می‌گردد که در هنگام ایجاد مدل کالیبراسیون موجب ضعف در پیش بینی خصوصیات کیفی محصول می‌گردد [۱۴]. بنابراین پیش پردازش طیف‌ها به منظور کاهش این تاثیرات مهم‌ترین قدم در تبیین مدل کالیبراسیون با قدرت و دقت بالا در پیشگویی خصوصیات کیفی محصولات مورد آزمون می‌باشد [۱۵]. گومزو همکاران (۲۰۰۶)، در مطالعه‌ای به بررسی تاثیر پیش پردازش داده‌های حاصل از طیف سنجی مرئی/مادون قرمز نزدیک و به کاربردن<sup>۲</sup> مدل رگرسیون مولفه اصلی و حداقل مربعات جزئی در تبیین مدلی به منظور پیشگویی اسیدیته و ماده خشک نارنگی گونه ساتسوما پرداختند نتایج حاصل از مدل سازی نشان دهنده برتری روش حداقل مربعات جزئی<sup>۵</sup> به روش رگرسیون مولفه اصلی<sup>۶</sup> می‌باشد [۶]. مقیمی و همکاران ۲۰۱۰ در پژوهشی به ارزیابی روش‌های پیش پردازش داده‌های طیفی در تبیین مدلی کالیبراسیون برای پیشگویی خصوصیات کیفی کیوی رقم هایوارد پرداختند. نتایج نشان دهنده تاثیر بالای انتخاب ترکیب مناسب روش‌های پیش پردازش داده‌های طیفی در تبیین مدلی با دقت و صحت بالا بود [۹].

مطالعات گذشته بیانگر تاثیر بالای روش‌های پیش پردازش بر تبیین مدلی مناسب برای پیشگویی خصوصیات کیفی محصول مورد آزمون دارد. در این پژوهش با توجه به ویژگی‌های به خصوص لیموترش از جمله اندازه کوچک و تراکم بالای دانه در گوشت میوه که باعث افزایش سطح پراکنش نور و ایجاد نویز به علاوه منابع دیگر معمول نویز در داده‌های طیفی ثبت شده می‌شود، از این رو انتخاب و به کاربردن روش‌های مختلف پیش پردازش مناسب داده‌های طیفی به منظور کاهش خطا در مدل پیشگویی اهمیت بیشتری پیدا می‌کند. هدف از این پژوهش ارزیابی تاثیر به کاربردن روش‌های پیش پردازش داده‌های

2. Intact fruit and vegetable  
3. Pick Broadening  
4. Baseline shifts  
5. Partial last square regression  
6. Principle component regression

1. FAO 2007

## ۱-۲- پیش پردازش و خارج کردن نویز از طیف لیموترش

اولین مرحله پس از طیف‌گیری از لیموترش‌ها تبدیل داده‌های انعکاسی به جذبی به منظور ایجاد رابطه خطی بین جذب و غلظت ماده مورد نظر بر اساس قانون بیر<sup>۷</sup> می‌باشد. به این منظور از رابطه ذیل استفاده شد.

$$Abc = \log \left[ \frac{1}{R} \right]$$

در این رابطه  $Abc$  طیف جذبی و  $R$  طیف انعکاسی می‌باشد. داده‌های حاصل از آشکار ساز طیف سنج دارای نویزها و داده‌های غیر مرتبط می‌باشند که برای بدست آوردن مدل کالیبراسیون دقیق و قابل اعتماد، انجام پیش پردازش و خارج نمودن نویزها ضروری است. به علت عمق نفوذ پایین در ناحیه مادون قرمز نزدیک طیف‌های حاصله از وضوح پایینی برخوردار بودند به این منظور داده‌های طیفی قبل از اعمال پیش‌پردازش با استفاده از روش میانگین مرکزی مورد اصلاح قرار گرفتند. میانگین مرکزی روشی نیرومند در بهبود وضوح و تفکیک پذیری طیف‌ها می‌باشد [۱۷]. به منظور پیش‌پردازش داده‌های طیفی از سه روش متفاوت استفاده گردید که شامل روش‌های اصلاح نقاط شروع و پراکندگی نور، هموارسازی و مشتق‌گیری بود. روش‌های فوق‌الذکر هر کدام شامل الگوریتم‌های متفاوتی می‌باشند که بهترین نتیجه با استفاده از ترکیب هر یک از این الگوریتم‌ها با یکدیگر حاصل گردید.

## ۲-۲- تبیین مدل کالیبراسیون

در آنالیز کمی و کیفی محصولات غذایی ساخت و تدوین مدل کالیبراسیون قابل اعتماد بسیار ضروری است. به کمک مدل می‌توان مشخصات نامعلوم نمونه‌ها را تعیین و پیش‌بینی نمود. در واقع هدف از تدوین مدل در این پژوهش یافتن رابطه‌ای میان خصوصیات نور بازتابشی و خصوصیات کیفی نمونه می‌باشد. در این پژوهش از روش حداقل مربعات جزئی به منظور تبیین مدل کالیبراسیون استفاده گردید. در روش حداقل مربعات جزئی با استفاده از داده‌های پیش‌بینی‌کننده که واریانس را در هر دو متغیر پیش‌بینی‌کننده و متغیر پاسخ توجیه می‌کنند، یکسری

حاصل از طیف سنجی مرئی / مادون قرمز نزدیک در تبیین مدلی با قدرت و دقت بالا در پیشگویی خصوصیات کیفی لیموترش می‌باشد.

## ۲- مواد و روش‌ها

تعداد ۱۰۰ نمونه لیموترش مورد استفاده در این پژوهش از وارپته کی لایم، و به صورت تصادفی از بازار میوه مشهد تهیه گردید. نمونه‌ها پس از شستشو شماره گذاری گردیدند. به منظور طیف‌گیری از نمونه‌ها نیازی به هیچ گونه آماده سازی بر روی نمونه نمی‌باشد که این از مزایای طیف سنجی مرئی / مادون قرمز نزدیک است. طیف مرئی / مادون قرمز نزدیک با استفاده از فیبر دومسیره که قسمت مرکزی برای انتقال نور از نمونه به آشکار ساز و فیبر بیرونی برای انتقال نور از منبع تابش (لامپ هالوژن با توان ۵۰ وات) به نمونه می‌باشد استفاده گردید در وضعیت انعکاسی و در ناحیه ۴۰۰ تا ۱۰۰۰ نانومتر با استفاده از اسپکترومتر مدل AvaSpec-2048TEC ساخت شرکت Avante هلند و با تفکیک پذیری ۱ نانومتر توسط نرم افزار آوا سافت نسخه ۷/۵ ثبت گردید. از هر نمونه در سه جهت با هدف کاهش نویز حرارتی بر روی آشکار ساز و افزایش دقت، تعداد ۱۰ اسکن متوالی در محور ساقه-غلاف گل به منظور پوشش تمامی نواحی میوه و همچنین به دلیل ضخامت کم پوست میوه در این نواحی و افزایش عمق نفوذ گرفته شد و میانگین ۳۰ طیف حاصله، برای هر نمونه گزارش گردید. زمان جمع آوری هر اسکن ۶۵۰ میلی ثانیه تنظیم و فاصله میان فیبر نوری و نمونه ۵ سانتی‌متر تنظیم گردید. بعد از هر طیف‌گیری و در همان روز آزمون‌های مخرب شامل اندازه‌گیری میزان مواد جامد خشک و اسیدیته بر روی هر نمونه صورت پذیرفت. هر نمونه لیموترش به صورت دستی آبیگری و سپس مواد جامد خشک محلول با استفاده از رفراکتومتر دیجیتال مجهز به کنترل دما مدل Abbe و اسیدیته لیموترش‌ها با استفاده از استاندارد AOAC و تیتراسیون با سود ۱/۱، نرمال اندازه‌گیری شد [۱۶].

7. Beer

۲- ریشه میانگین مربعات خطا<sup>۱۳</sup> که به عنوان تقریبی از میزان خطای مدل در پیش بینی مقادیر نسبت به مقادیر واقعی اندازه گیری شده مطرح می‌شود، از رابطه ذیل محاسبه می‌گردد:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (\hat{y}_i - y_i)^2}{N}}$$

در این رابطه N تعداد نمونه،  $\hat{y}_i$  مقدار پیش بینی شده هر پارامتر برای نمونه i ام و  $y_i$  مقدار اندازه گیری شده برای نمونه i ام.

۳- انحراف استاندارد توزیع خطا<sup>۱۴</sup> که به منظور ارزیابی کارایی مدل و از رابطه ذیل محاسبه می‌گردد.

$$SDE = \sqrt{\frac{\sum_{i=1}^N (\hat{y}_i - y_i - ME)^2}{N - 1}}$$

در این رابطه میانگین خطا<sup>۱۵</sup> می‌باشد و از رابطه ذیل محاسبه می‌گردد.

$$ME = \frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)$$

۴- باقیمانده انحراف پیش بینی<sup>۱۶</sup> که از جمله شاخصه‌های مهم در ارزیابی قدرت و انعطاف پذیری مدل به حساب می‌آید و از رابطه ذیل محاسبه می‌گردد.

$$RPD = \frac{SD}{RMSE}$$

تمامی پیش پردازش‌ها و مدل سازی مرتبط با ارزیابی خصوصیات کیفی لیموترش‌ها به صورت غیر مخرب با استفاده از نرم افزار پارلس<sup>۱۷</sup> نسخه ۳/۱ صورت پذیرفت.

### ۳- نتایج و بحث

همان‌گونه که در جدول شماره ۱ قابل مشاهده است میزان ماده خشک محلول لیموترش‌ها از ۸/۱ تا ۱۱/۲ بر مبنای درجه بریکس متغیر است همچنین میانگین اسیدیته لیموترش‌ها ۸/۳۱ بر مبنای اسید سیتریک حاصل گردید.

ترکیبات خطی متعامد از متغیرهای مستقل و وابسته استخراج می‌شود که به آن مولفه<sup>۸</sup> یا متغیر پنهان<sup>۹</sup> می‌گویند و در واقع مدل نهایی بر مبنای این مولفه‌ها تبیین می‌گردد [۱۸]. به طور کلی می‌توان بیان کرد که تکنیک پیاده سازی شده به وسیله روش حداقل مربعات جزئی شامل ایجاد ترکیبی بهینه سازی شده میان میزان واریانس‌های متغیرهای پیش‌بینی کننده و پیش‌بینی شونده است. در این پژوهش متغیر مستقل طول موج و متغیر وابسته میزان خشک محلول و اسیدیته می‌باشد.

### ۲-۳- اعتبار سنجی مدل

پس از تبیین مدل تعیین توانایی آن در پیش بینی و ارزیابی دقت آن بسیار حائز اهمیت می‌باشد. در این پژوهش از روش اعتبار سنجی آزمونی<sup>۱۰</sup> که اعتبار سنجی مدل با استفاده از نمونه‌های مستقل از نمونه‌های به کار برده شده در تبیین مدل، به منظور ارزیابی مدل استفاده گردید.

به منظور مدل‌سازی، نمونه‌ها به صورت تصادفی به دو دسته تقسیم گردیدند، دسته اول که ۷۵ درصد کل نمونه‌ها را شامل می‌گردید برای تبیین مدل پیش بینی و دسته دوم برای ارزیابی مدل ساخته شده به کار گرفته شدند.

### ۲-۴- آماره‌های مورد استفاده به منظور ارزیابی

#### مدل

۱- ضریب تبیین<sup>۱۱</sup> که به منظور ارزیابی میزان همبستگی میان مقادیر پیش بینی و مقادیر واقعی مورد استفاده قرار می‌گیرد و از رابطه زیر بدست می‌آید:

$$R^2 = 1 - \left( \frac{RMSE}{SD} \right)^2$$

در این رابطه SD انحراف معیار<sup>۱۲</sup> می‌باشد که از رابطه ذیل محاسبه می‌گردد.

$$SD = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x}_i)^2}{N - 1}}$$

در رابطه اخیر N تعداد نمونه،  $\bar{x}_i$  مقدار پارامتر اندازه گیری شده و  $x_i$  مقدار میانگین پارامتر اندازه گیری شده می‌باشد

13. Root Mean Square Error

14. Standard deviation of the error distribution

15. Mean Error

16. Residual prediction deviation

17. Parles

8. Component

9. Latent variable

10. Test Validation

11. Coefficient of determination

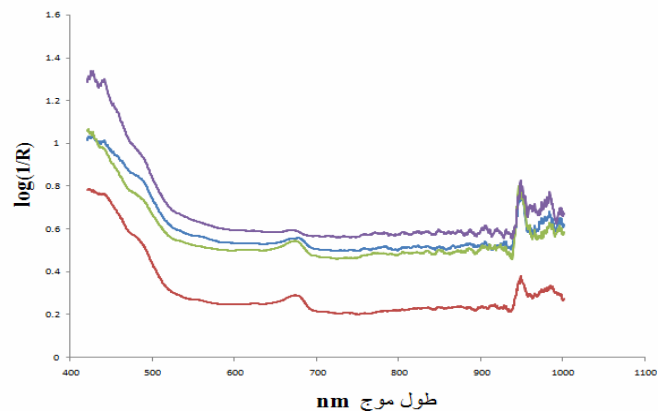
12. Standard deviation

جدول ۱ خصوصیات کیفی اندازه گیری شده با استفاده از روش‌های مخرب

خصوصیات اندازه گیری شده	میانگین	انحراف معیار	میانه	حداکثر مقدار	حداقل مقدار
ماده خشک محلول	۹/۳۷	۰/۶۲	۹/۸۰	۱۱/۲۰	۸/۱
اسیدیته	۸/۳۱	۰/۴۴	۸/۲۷	۹/۵۷	۷/۲۴

### ۳-۱- تاثیر پیش پردازش بر روی داده‌های طیفی

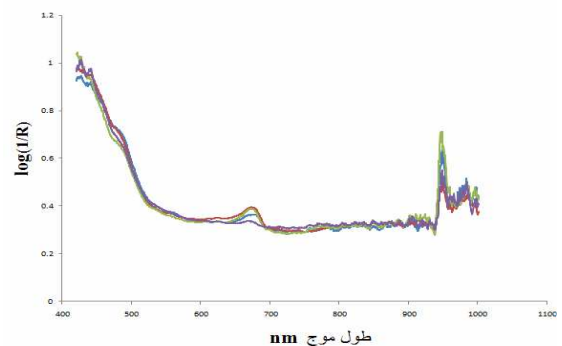
همان‌گونه که در طیف حاصل از ۴ نمونه لیموترش (شکل ۱) مشاهده می‌گردد طیف‌ها دارای نقاط شروع متفاوت به دلیل تفاوت در اندازه و شکل لیموترش‌ها و پراکنش نور حاصل از انعکاس از سطح بیرونی و داخلی به دلیل ناهمواری‌های موجود بر روی سطح و وجود دانه و سایر عوامل درونی می‌باشد. اولین پیش پردازش صورت گرفته بر روی طیف‌ها یکسان سازی نقاط شروع می‌باشد.



شکل ۱ طیف حاصل از ۴ نمونه لیموترش.

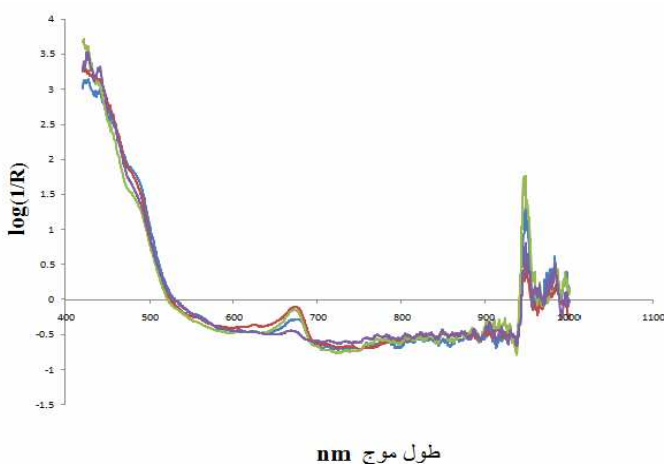
### ۳-۱-۱- تصحیح نقاط شروع و پراکنش نور

برای کاهش پراکنش نور و یکسان سازی نقاط شروع به دلیل عدم یکنواختی اندازه نمونه‌ها روش‌های تصحیح پخش افزایشده و تبدیل متغیر نرمال به کار برده شد.



شکل ۲ تاثیر پیش پردازش تصحیح پخش افزایشده بر روی طیف‌های لیموترش.

هدف از اعمال اولین پیش پردازش تصحیح و یکسان کردن نقاط شروع طیف‌ها می‌باشد. روش تصحیح پخش افزایشده (شکل ۲) بر اساس تصحیح تمامی طیف‌ها بر مبنای طیف ایده‌آل می‌باشد که طیف ایده‌آل در واقع میانگینی از تمامی طیف‌ها می‌باشد و هر طیف به صورت جداگانه بر روی طیف میانگین تطبیق<sup>۱۸</sup> می‌گردد. روش تبدیل متغیر نرمال استاندارد (شکل ۳) بر مبنای تفریق هر طیف از طیف میانگین و کالیبره کردن هر طیف با استفاده از انحراف معیار طیف میانگین باعث یکنواختی نقاط شروع طیف‌ها می‌گردد.



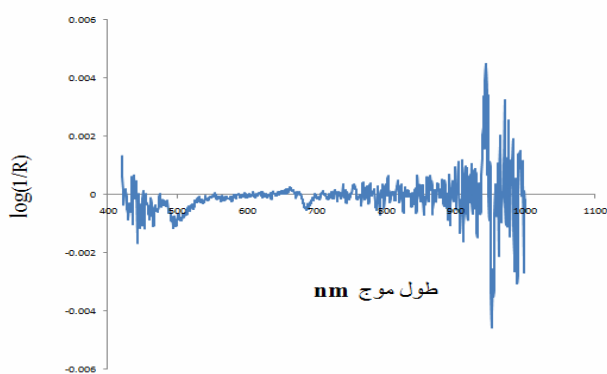
طول موج nm

شکل ۳ تاثیر پیش پردازش تبدیل متغیر نرمال استاندارد بر روی طیف‌های حاصل از ۴ نمونه لیموترش.

### ۳-۱-۲- هموار سازی و کاهش نویز

هدف از اعمال هموار سازی بهبود نسبت سیگنال به نویز می‌باشد. منبع ایجاد نویز در طیف‌ها به عوامل مختلفی بستگی دارد که می‌توان به عدم یکنواختی در بافت، وجود دانه در میوه و همچنین وجود نویز حرارتی ایجاد شده به وسیله منبع حرارتی و آشکارساز دستگاه اشاره نمود. به این منظور روش فیلتر میانه ساویتزکی-گولی مورد استفاده قرار گرفت. روش فیلتر میانه (شکل ۴) برای کاهش نویزهای تناوبی در طیف بر اساس درجه

18. Fitted

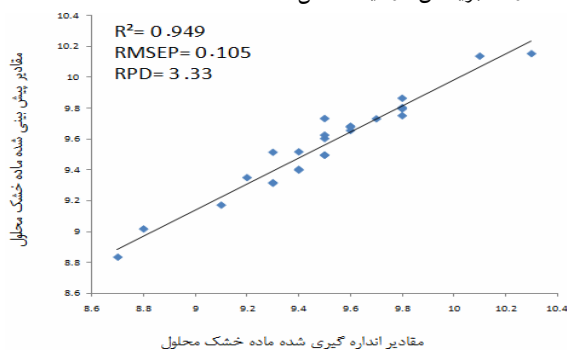


شکل ۶ تاثیر مشتق اول بر روی طیف های حاصل از ۴ لیموترش.

بهترین نتیجه در به کار گیری پیش پردازش ها هنگامی حاصل شد که ترکیبی از روش های مذکور به صورت مناسبی با یکدیگر اعمال گردید.

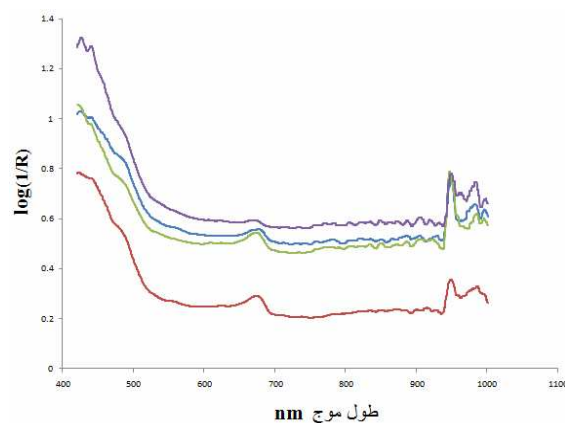
### ۳-۲- مدل سازی پیش بینی ماده خشک محلول لیموترش

در میان مدل های تبیین شده به صورت ترکیب با پیش پردازش های متفاوت بر روی داده های طیفی برای پیش بینی ماده خشک محلول در لیموترش، مدلی با استفاده از ترکیب روش های تصحیح پخش افزاینده، فیلتر میانه با درجه چهارم، مشتق اول و میانگین مرکزی<sup>۱۹</sup> با ده فاکتور، ضریب تبیین به بالاترین مقدار (۰/۹۴۹۲) و به کمترین مقدار ریشه مربعات خطای پیش بینی (۰/۱۰۵ درجه برعکس) رسید (شکل ۷).



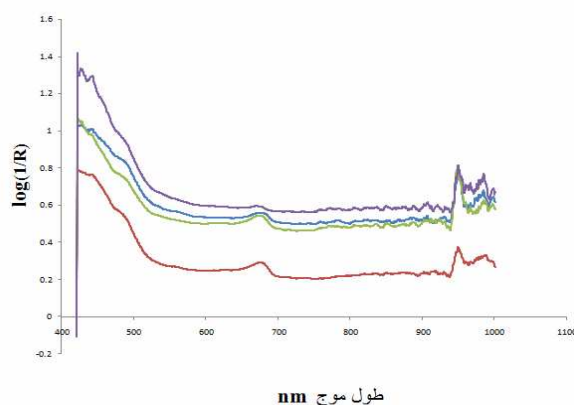
شکل ۷ نتایج حاصل از مدل پیش بینی ماده خشک محلول در لیموترش. در جدول ۲ تاثیر پیش پردازش داده های طیفی بر روی کارایی مدل های پیش بینی کننده ماده خشک محلول آورده شده است. مقایسه این مدل ها نشان می دهد که پیش پردازش تا چه میزان بر روی دقت و کارایی مدل پیش بینی کننده تاثیر گذار می باشد.

انتخابی، یک پنجره بر روی طیف تعبیه، و بر روی تمامی طیف پیمایش می شود (هر نقطه از طیف یکبار در مرکز پنجره قرار می گیرد) و بر اساس مقادیر همسایه ها، آن نقطه از طیف را اصلاح می گرداند.



شکل ۸ تاثیر فیلتر میانه بر کاهش نویز طیف های حاصل از ۴ نمونه لیمو ترش.

روش فیلتر ساویتزکی\_گولی (شکل ۵) بسطی از روش جابه جایی میانگین و مشابه روش فیلتر میانه است و تفاوت آن با روش مذکور بر اساس میانگین گیری از نقاط همسایگی و سپس تطبیق بر روی منحنی چند جمله ای و هموارسازی پیک ها بر روی طیف می باشد.



شکل ۹ تاثیر فیلتر ساویتزکی\_گولی بر کاهش نویز و هموارسازی پیک در طیف های حاصل از ۴ نمونه لیمو ترش.

### ۳-۱-۳- مشتق گیری

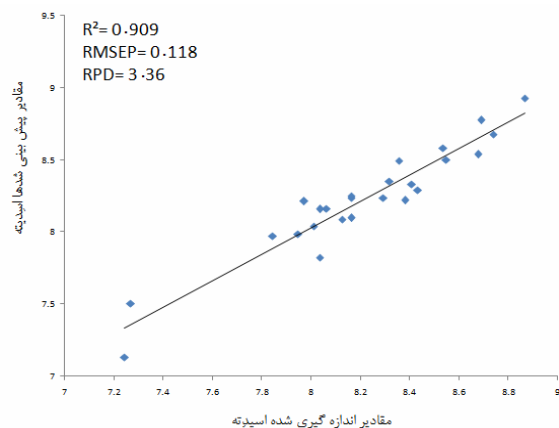
هدف از اعمال مشتق گیری بر روی طیف ها، افزایش تفکیک پذیری پیک ها، کاهش تداخلات طیفی و مشخص شدن پیک های ضعیف می باشد. (شکل ۶).

جدول ۲ تاثیر پیش پردازش داده‌های طیفی بر روی دقت و کارایی مدل پیش بینی کننده ماده خشک محلول لیموترش

RPD	SDE	RMSEP	R <sup>2</sup>	مولفه مدل	پیش پردازش
۱/۸۳	۰/۱۹۰	۰/۱۹۱	۰/۷۱	۱۰	تصحیح پخش افزایشده
۲/۳۱	۰/۱۵۴	۰/۱۵۲	۰/۸۰۷	۷	تصحیح پخش افزایشده، مشتق دوم و فیلتر ساویتزکی-گولی
۲/۶۴	۰/۱۳۵	۰/۱۳۳	۰/۸۶۲	۸	تبدیل متغیر نرمال، مشتق اول، فیلتر میانه
۳/۳۳	۰/۰۸۸	۰/۱۰۵	۰/۹۴۹	۱۰	تصحیح پخش افزایشده، مشتق اول، فیلتر میانه

### ۳-۳- مدل سازی پیش بینی اسیدیته لیموترش

در بین مدل‌های تبیین شده (جدول ۳)، مدلی از ترکیب پیش پردازش‌های تبدیل متغیر نرمال استاندارد، فیلتر ساویتزکی-گولی با دو نقطه هموارسازی و منحنی دو جمله‌ای، مشتق اول و میانگین مرکزی با ۹ فاکتور بدست آمد که در آن ضریب همبستگی به بالاترین میزان (۰/۹۰۹) و ریشه مربعات خطای پیش بینی (۰/۱۱۸) به کمترین مقدار رسید (شکل ۸).



شکل ۸ نتایج حاصل از مدل پیش بینی اسیدیته لیموترش.

نتایج بدست آمده در خصوص مدل پیش بینی کننده ماده خشک محلول در لیموترش در مقایسه با نتایج گزارش شده توسط گومز و همکاران (۲۰۰۶) برای نارنگی گونه سوتساما با به کار بردن پیش پردازش تصحیح پخش افزایشده و فیلتر جابه جایی میانگین [۵]، لیو و همکاران (۲۰۱۰) برای پرتقال گونه ناول با استفاده از پیش پردازش تصحیح پخش افزایشده [۷] و کایولا (۲۰۰۸) برای پرتقال گونه والنسیا با استفاده از پیش پردازش مشتق اول، متغیر نرمال استاندارد و متعامد سازی [۶] دارای ضریب تبیین بالاتر و ریشه مربعات خطای پایین تری بدست آمد. پیش پردازش‌های به کار برده شده در تبیین مدل پیشگویی ماده خشک محلول لیموترش در مقایسه با نارنگی گونه ساتسوما و پرتقال گونه ناول دارای نیاز به پیش‌پردازش بیشتری به دلیل وجود تراکم دانه بیشتر در گوشت میوه و ایجاد نویز در داده‌های طیفی بود و میزان پیش پردازش صورت گرفته بر روی طیف‌های لیموترش مشابه با پرتقال گونه والنسیا، به دلیل مشابهت در خصوصیات ساختاری به دلیل وجود تراکم دانه در گوشت هر دو میوه می‌باشد.

جدول ۳ تاثیر پیش پردازش داده‌های طیفی بر روی دقت و کارایی مدل پیش بینی کننده اسیدیته لیموترش

RPD	SDE	RMSEP	R <sup>2</sup>	مولفه مدل	پیش پردازش
۱/۷۲	۰/۲۳۳	۰/۲۳	۰/۶۵۷	۱۰	تبدیل متغیر نرمال
۲/۱۹	۰/۱۶۴	۰/۱۸۱	۰/۸۴	۶	تصحیح پخش افزایشنده ، مشتق دوم، فیلتر میانه
۲/۸۶	۰/۱۴۱	۰/۱۳۸	۰/۸۷۴	۷	تبدیل متغیر نرمال ، مشتق اول، فیلتر میانه
۳/۳۶	۰/۱۲	۰/۱۱۸	۰/۹۰۹	۹	تبدیل متغیر نرمال ، فیلتر ساویتزکی_گولی، مشتق اول

Raman analysis of plants. *Journal of Molecular Structure*, 509: 201–212.

- [3] Kawano, S., Watanabe, H., and Iwamoto, M., 1992. Determination of sugar content in intact peaches by near infrared spectroscopy with fiber optics in interactance mode. *Journal of the Japanese Society for Horticultural Science*, 61: 445–451.
- [4] Greensill, C.V., and Walsh, K.B., 2002. Calibration transfer between miniature photodiode array-based spectrometers in the near infrared assessment of mandarin soluble solids content. *Journal of Near Infrared Spectroscopy*, 10; 27–35.
- [5] Gómez, A. H., He, Y., and Pereira, A. G. 2006. Non-destructive measurement of acidity, soluble solids and firmness of Satsuma mandarin using vis-NIR spectroscopy techniques. *Journal of Food Engineering*, 77: 313–319.
- [6] Cayuela, J. A. 2008. Vis-NIR soluble solids prediction in intact oranges (*Citrus sinensis* L.) cv. Valencia Late by reflectance. *Postharvest Biology and Technology*, 47: 75–80.
- [7] Liu, Y., Sun, X., and Ouyang, A. 2010. Non-destructive measurements of soluble solid content of navel orange fruit by visible-NIR spectrometric technique with PLS and PCA-BPNN. *LWT-Food Science and Technology*, 43: 602–607.
- [8] Jamshidi, B., Minaei, S., Mohajerani, E., and Ghassemian, H. 2012. Reflectance Vis/NIR spectroscopy for nondestructive taste characterization of Valencia oranges. *Computers and Electronics in Agriculture*, 85: 64–69.
- [9] Moghimi, A., Aghkhani, M. H., Sazgarnia, A., & Sarmad, M. 2010. Vis/NIR spectroscopy and chemometrics for the prediction of soluble

نتایج بدست آمده در خصوص مدل پیش بینی کننده اسیدیته در لیموترش در مقایسه با نتایج بدست آمده توسط کایولا (۲۰۰۸) برای پرتقال گونه والنسیا [۶]، مک گلون و همکاران (۲۰۰۳) برای نارنگی گونه سانتسوما [۱۹] و گوسری و همکاران (۲۰۰۵) برای نارنگی گونه امپریال [۲۰] دارای ضریب تبیین بالاتر و ریشه خطای پیش بینی پایین تری بود. ضریب بالاتر مدل پیشگویی حاصل از این پژوهش در مقایسه با نتایج حاصل از سایر پژوهش‌های ذکر شده را می‌توان به علت سطح بالای اسید سیتریک موجود در لیموترش نسبت به سایر مرکبات ذکر نمود.

#### ۴- نتیجه گیری

با توجه به لزوم طراحی و استفاده از دستگاه‌های خودکار و بهره گیری از روش‌های سریع و مطمئن در بررسی کیفیت محصول (لیموترش) طی برداشت، فرآوری و نگهداری، ارزیابی خصوصیات کیفی درونی لیموترش با استفاده از روشی غیر مخرب و با سهولت و دقت بالا امری ضروری به نظر می‌آید. نتایج حاصل از این تحقیق نشان دهنده قابلیت بالای روش طیف‌سنجی مرئی/مادون قرمز نزدیک در اندازه گیری هم زمان پارامترهای کیفی لیموترش می‌باشد. همچنین نتایج نشان‌دهنده تاثیر بالای پیش پردازش داده‌های طیفی در تبیین مدلی با صحت و دقت بالا به منظور پیش‌بینی خصوصیات کیفی محصول است.

#### ۵- منابع

- [1] Food and Agriculture Organization (FAO), 2010., Available at: [www.fao.org/statistics.htm](http://www.fao.org/statistics.htm)
- [2] Schrader, B., Klumpb, H.H., Schenzelc, K. and Schulz, H. 1999. Non-destructive NIR FT



- [15] Cen, H., He, Y., and Huang, M. 2006. Measurements of soluble solids contents and pH in orange juice using chemometrics and Vis-NIRS. *Journal of Agricultural Food Chemistry*, 54:7437-7443.
- [16] Official Methods of Analysis. 1999. 16th Edition, 5th Reversion, AOAC International, Gaithersburg, MD, method 936.16.
- [17] Ozaki, Y., McClure, W. F., & Christy, A. A. 2006. Near-infrared spectroscopy in food science and technology: John Wiley & Sons. 58-60
- [18] Wold, S., Sjöström, M., and Eriksson, L. 2001. PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, 58: 109-130.
- [19] McGlone, V.A., Jordan, R.B., and Martinsen, P.J., 2002. Vis/NIR estimation at harvest of pre- and post-storage quality indices for 'Royal Gala' apple. *Postharvest Biology and Technol*, 25: 135-144.
- [20] Guthrie, J. A., Walsh, K. B., Reid, D. J., and Liebenberg, C. J. 2005. Assessment of internal quality attributes of mandarin fruit. 1. NIR calibration model development. *Crop and Pasture Science*, 56: 405-416.
- solids content and acidity (pH) of kiwifruit. *biosystems engineering*, 106, 295-302.
- [10] Liu, Y. d., Ying, Y.-b., Fu, X., & Lu, H. 2007. Experiments on predicting sugar content in apples by FT-NIR technique. *Journal of Food Engineering*, 80: 986-989.
- [11] Paz, P., Sánchez, M.-T., Pérez-Marín, D., Guerrero, J. E., and Garrido-Varo, A. 2009. Instantaneous quantitative and qualitative assessment of pear quality using near infrared spectroscopy. *Computers and electronics in agriculture*, 69: 24-32.
- [12] Flores, K., Sánchez, M.-T., Pérez-Marín, D., Guerrero, J.-E., and Garrido-Varo, A. 2009. Feasibility in NIRS instruments for predicting internal quality in intact tomato. *Journal of Food Engineering*, 91: 311-318.
- [13] Leonardi, L., & Burns, D. H. 1999. Quantitative multi wavelength consistent measurements using single-wavelength photo time of flight correction. *Applied Spectroscopym*, 53: 637-646.
- [14] Blanco, M., & Villarroya, I. 2002. NIR spectroscopy: a rapid-response analytical tool. *Trends in Analytical Chemistry*, 21: 240-250.

## Application of preprocessing techniques for visible/near infrared spectroscopy data in developing a model for the prediction of soluble solid and acidity of lime

Jajarmi, A. <sup>1</sup>, Taghizadeh, M. <sup>2\*</sup>, Sazgarnia, A. <sup>3</sup>, Behzad, Kh. <sup>2</sup>

1. MSc., Dept. of Food Science and Technology, Ferdowsi University of Mashhad.

2. Assistant Professor, Dept. of Food Science and Technology, Ferdowsi University of Mashhad.

3. Associate Professor, Dept. of Medical Physics, Mashhad University of Medical Sciences, Iran.

(Received: 93/4/30 Accepted: 93/6/26)

Investigation of quality characteristics of food products during different manufacturing stages such as storage, processing and consumption is important to reduce food loss. In recent years many researches have established for developing rapid and non-destructive techniques for quality control. In this study the potential of visible and near infrared spectroscopy (Vis/NIRS) in determining the quality parameters of lime including total soluble solid and acidity in reflection mode was investigated in the wavelength range of 400 to 1000 nm. The effects of different pre-processing techniques and spectral treatments, such as standard normal variable transformation (SNV), multiplicative scatter correction (MSC), median filter, Savitzky & Golay and the derivatives were evaluated. The model was developed based on partial least squares (PLS) regression. The correlation coefficient ( $R^2$ ) and root mean square error of prediction (RMSEP) for predictive model of soluble solids content was 0.949, 0.105 °Brix respectively. These parameters of the model predicting acidity was found to be 0.909 and 0.118 respectively. These results showed the high potential of Vis/NIRS and the important role of preprocessing techniques in developing precise models for the prediction of lime internal quality characteristics.

**Key words:** Quality characteristics, Visible and near infrared spectroscopy, Preprocessing techniques, Predictive models.

---

\* Corresponding Author E-Mail Address: mtaghizadeh@um.ac.ir