

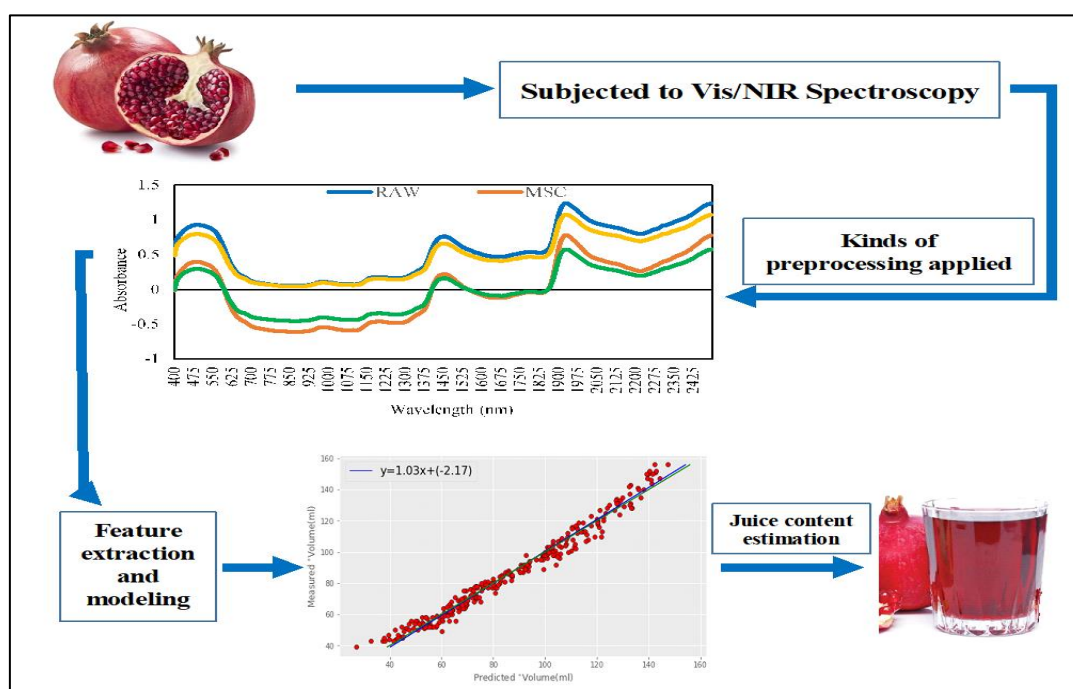
Visible and Near-Infrared Spectroscopy for Determining Pomegranate Juice Volume

Mohammad Jamalifard¹, Seyed Mehdi Nassiri^{1*}, Alimohammad Shirzadifar¹, Alireza Shahsavari²

1. Department of Biosystems Engineering, School of Agriculture, Shiraz University, Shiraz, Iran

4. Department of Horticulture Sciences, School of Agriculture, Shiraz University, Shiraz, Iran

(Received: Nov. 7, 2021- Revised: July 29, 2022- Accepted: Nov. 3, 2022)



Abstract: Pomegranate fruit is one of the essential garden products in Iran, which, in addition to its high economic-commercial value, has many nutritional benefits. Due to the lack of appropriate development of the industry in the processing sector for this product, factories use completely traditional and limited methods for measuring the quality of pomegranate, which reduces the quality of the processed product. In the present study, the feasibility of visible and near-infrared (Vis/NIR) spectroscopy as non-destructive technology was used to determine the pomegranate fruit juice volume, which is one of the crucial quality attributes of the product. Spectral data of the samples obtained by applying waves in the range of 400 to 2500 nm, were evaluated by five methods including multiplicative scatter correlation, standard normal variate (SNV), vector normalization, first derivative, second derivative, as well as non-processed state. Then partial least square regression (PLSR) was used to estimate the juice volume. These processes were implemented according to machine learning algorithms using PYTHON 3.8 software. Lowest modeling error (18.61) was achieved with 36 principal components of extracted features from spectral signature. The results showed that this method estimated the amount of pomegranate juice with a coefficient of determination of 94 %, mean absolute error of 5.3 and distance of 0.98. This outcome resulted from combination of SNV preprocessing with PLSR regression in the range of 400 to 2500 nm. In spite of the effect of preprocessing for juice volume estimation, Vis-NIR spectroscopy possess desired condition for estimation of pomegranate juice volume.

Key words: Machine Learning, Non-Destructive Method, Pomegranate, Spectroscopy

* Corresponding Author: nasiri@shirazu.ac.ir

طیف‌سنجی مرئی و مادون قرمز نزدیک برای تعیین حجم آب میوه انار

محمد جمالی فرد^۱، سید مهدی نصیری^{۱*}، علی محمد شیرزادی فرد^۲، علیرضا شهسوار^۲

۱. بخش مهندسی بیوسیستم، دانشکده کشاورزی، دانشگاه شیراز، شیراز، ایران

۲. بخش علوم باغبانی، دانشکده کشاورزی، دانشگاه شیراز، شیراز، ایران

تاریخ دریافت: ۱۴۰۰/۸/۱۶ - تاریخ بازنگری: ۱۴۰۱/۵/۷ - تاریخ تصویب: ۱۴۰۱/۸/۱۲

چکیده

میوه انار یکی از محصولات باغی مهم در ایران است که علاوه بر ارزش بالای اقتصادی-تجاری از فواید غذایی فراوانی برخوردار است. به دلیل عدم توسعه مناسب صنعت فرآوری این محصول، کارخانه‌ها از روش‌های کاملاً سنتی و محدود در سنجش کیفیت انار استفاده می‌کنند که باعث کاهش مرغوبیت کالای تولیدی می‌شود. در این پژوهش استفاده از فناوری غیرمخرب طیف‌سنجی مرئی-مادون قرمز نزدیک برای تعیین حجم آب میوه انار که از ویژگی‌های اساسی کیفیت این محصول است مورد بررسی قرار گرفت. داده‌های طیفی نمونه‌ها حاصل از اعمال امواج در دامنه ۴۰۰ تا ۲۵۰۰ نانومتر علاوه بر ارزیابی در حالت پردازش نشده، به پنج روش تصحیح پراکندگی ضربی، متغیر نرمال استاندارد (SNV)، نرمال‌سازی بردار، مشتق اول و دوم پیش‌پردازش شدند و مورد ارزیابی قرار گرفتند. سپس به منظور تخمین حجم آب از رگرسیون حداقل مربعات جزئی (PLSR) استفاده شد. پیاده‌سازی این فرآیندها در قالب الگوریتم‌های یادگیری ماشین و با استفاده از نرم افزار PYTHON 3.8 صورت گرفت. برای ایجاد مدل‌ها، تعداد مولفه‌های اصلی مرتبط با ویژگی‌های استخراج شده از امضای طیفی به تعداد ۳۶، کمترین خطا را (۱۸/۶۱) حاصل نمود. نتایج نشان داد که این روش با ضریب تبیین ۹۴ درصد، میانگین خطای مطلق ۵/۳ و ضریب توافق ۰/۹۸ مقدار آب میوه انار را تخمین زد. این نتیجه از ترکیب پیش‌پردازش SNV با رگرسیون PLSR در بازه ۴۰۰ تا ۲۵۰۰ نانومتر حاصل شد. علیرغم تاثیر نوع پیش‌پردازش داده‌ها در تخمین میزان آب انار، طیف-سنجی مرئی-مادون قرمز نزدیک شرایط لازم برای تخمین آب انار را دارا هست.

واژه‌های کلیدی: انار، روش غیرمخرب، طیف‌سنجی، یادگیری ماشین

مقدمه

انار (*Punica granatum* L.) میوه‌ای از خانواده (*Punicaceae*) است که در مناطق گرمسیری و نیمه گرمسیری به عمل می‌آید و به دلیل مرغوبیت بالا، از نظر صادرات، محصولی بی‌همتا بوده و از نظر اقتصادی ارزش بالایی دارد. در آسیای مرکزی درخت انار به صورت وحشی در مناطق ایران، ترکمنستان، و شمال هند رشد می‌کند و به طور کلی مناطقی که دارای آب و هوای مدیترانه‌ای با نور زیاد خورشید، زمستان‌های ملایم با حداقل دمای ۱۲- درجه سلسیوس، تابستان‌های گرم و خشک و بارندگی در مراحل آخر نمو میوه هستند، بهترین اقلیم برای پرورش میوه انار به حساب می‌آیند. در کشور ما مناطق وسیعی که در حاشیه کویر مرکزی قرار دارند دارای آب و هوای نیمه گرمسیری هستند و درختان انار در چنین شرایطی می‌توانند به خوبی رشد کنند. به همین دلیل در تمامی استان‌های کشور که در محدوده کویر لوت و دشت کویر واقع شده‌اند کشت و کار انار از قدیم تا به امروز رونق ویژه‌ای داشته و سطح زیر کشت، تنوع ارقام و میزان باردهی درخت و همچنین کیفیت محصول آن قابل توجه است. علاوه بر این میوه انار در کشورهای زیادی نظیر پاکستان، ارمنستان، افغانستان، ازبکستان، تاجیکستان، آذربایجان، سوریه، لبنان، لیبی، سودان، تونس، بنگلادش، مصر، مراکش، اسپانیا، ایتالیا، آلمان، فرانسه، یونان، چین، روسیه، آمریکا و استرالیا تولید می‌شود که بیانگر پراکندگی بالای این محصول در سطح جهان و در نتیجه اهمیت اقتصادی فراوان آن است (Bashghareh, 2019). ایران با تولید سالیانه ۹۱۷۵۲۸ و ۱۰۹۹۰۳۸ تن انار به ترتیب در سال‌های ۱۳۹۷ و ۱۳۹۸، سومین تولیدکننده برتر این میوه در جهان بوده است که این آمار بیانگر اهمیت ویژه تجاری این محصول در کشورمان ایران است (Ahmadi et al., 2020).

علاوه بر پراکندگی و فراوانی انار در ایران و جهان، این میوه از لحاظ خواص و فواید نیز جایگاه ویژه‌ای در

بین میوه‌ها دارد که این خود دلیلی بر اهمیت آن است. یافته‌های اخیر محققان نشان می‌دهد که از انار می‌توان به عنوان یک گیاه دارویی بهره برد. بافت، گل، پوست و برگ میوه انار دارای مواد شیمیایی زیست فعالی است که ضد میکروب بوده و فشار خون را کاهش می‌دهد و همچنین در درمان بیماری‌هایی نظیر دیابت و سرطان مؤثر است (Patel et al., 2018). به علاوه میوه انار با دارا بودن مقادیر قابل توجهی از مواد جامد محلول، اسیدهای آلی، ویتامین‌ها، پلی ساکاریدها، عناصر معدنی، اسید-های چرب و ویتامین C، ارزش غذایی بسیار بالایی دارد (Arendse et al., 2017).

مصرف تازه‌خوری انار به دلیل عمر ماندگاری پایین آن در تمام ماه‌های سال امکان پذیر نیست. بر همین اساس تولید فرآورده‌هایی که قابلیت نگهداری طولانی مدت داشته باشند اهمیت ویژه‌ای دارد. همچنین از آن جا که بخش قابل توجهی از محصول انار برداشت شده علاوه بر داشتن کیفیت مناسب آب و دانه به دلیل آفتاب‌سوختگی، اندازه کوچک و یا رنگ نامطلوب، قابلیت فروش و عرضه مستقیم به بازار را ندارد و ارزش تجاری آن پایین است، اهمیت تولید محصولات فرآوری شده این میوه را دو چندان می‌کند. افزون بر آن، با وجود محدودیت‌هایی نظیر هزینه بالای انبارداری و سردخانه، فاسد شدن و کاهش حجم انار در گذر زمان، خشکی پوست و سایر موارد، ارزش صادراتی این محصول همواره در خطر است که لزوم توجه به فرآوری آن را ضروری می‌سازد. بنابر دلایل ذکر شده امروزه فرآوری میوه انار به شدت مورد توجه کارخانه‌های فرآوری قرار گرفته و بستری بسیار مناسب برای انجام فرآیندهای متفاوت بر روی آن به وجود آمده است. انار از جمله میوه‌هایی است که قابلیت فرآوری و تبدیل بالایی دارد. محصولات بسیار متنوع و زیادی از جمله آبمیوه، رب، معجون، مارمالاد، مربا، ژله، انواع سس و چاشنی، پودر شربت، ترشی، بستنی، نوشابه، پکتین، روغن، اسانس، انار دانه، لواشک، سرکه و خوراک دام، برخی از

فراورده‌های انار هستند.

عدم توسعه و بروزرسانی دستگاه‌های فرآوری میوه انار و همچنین حساسیت میوه انار به آسیب‌های مکانیکی در حین فرآوری، موجب شده است که مراحل گزینش، خرید و ارزیابی کیفیت این محصول توسط کارخانه‌های فرآوری به صورت عینی و در سطح کاملاً محدود و با استفاده از روش‌های سنتی صورت گیرد. اگرچه ارزش تجاری انار وابسته به ویژگی‌های ظاهری آن نظیر اندازه، شکل، و رنگ پوست است اما این ویژگی‌های ظاهری همواره بیانگر کیفیت درونی این میوه نیستند (Munera et al., 2019). این موضوع باعث می‌شود تا میوه‌های نامناسب به راحتی وارد چرخه تولید شوند و در نتیجه کیفیت محصول نهایی به میزان قابل توجهی کاهش یابد. همچنین کارخانه‌ها نیز بابت خرید محصول بی‌کیفیت و یا کم‌کیفیت متحمل ضررهای اقتصادی می‌شوند، از این رو صنعت فرآوری انار نیازمند یک سیستم ارزیابی کیفیت با سرعت و دقت بالا، هزینه کم و غیرمخرب است.

امروزه روش‌های متنوعی برای کنترل کیفیت میوه‌ها در دنیا وجود دارد که در این بین روش‌های غیر-مخرب به دلیل توانایی در تشخیص ویژگی‌های مختلف محصول بدون آسیب رساندن به آن‌ها از اهمیت ویژه‌ای برخوردار هستند. برخی از این روش‌ها شامل طیف سنجی مرئی-مادون قرمز نزدیک (Pourdarbani et al., 2020)، تصویربرداری چند طیفی و فراطیفی^۱ (Yang et al., 2020)، تصویربرداری رزونانس مغناطیسی هسته‌ای^۲ (Baek et al., 2020)، توموگرافی محاسباتی با استفاده از اشعه ایکس (Salmanizadeh et al., 2015). و طیف‌سنجی دی‌الکتریک (Fazayeli et al., 2019) هستند. روش‌های طیف‌سنجی مانند مرئی، مادون قرمز نزدیک، مادون قرمز میانه و طیف‌سنجی فراطیفی غیر مخرب‌تر از روش‌های مکانیکی و شیمیایی هستند و داشتن ابزارهای

ساده، سریع، قابل حمل و فراگیر، از دیگر فواید استفاده از این روش‌ها هستند. بنابر همین دلایل امروزه از این فناوری به شکل گسترده‌ای، در صنایع کشاورزی، غذا، دارو و پتروشیمی استفاده می‌شود (Xiao et al., 2019).

با استفاده از فناوری غیرمخرب اشعه ایکس سعی کردند تا برخی مؤلفه‌های بیوفیزیکی دو رقم میوه انار مانند حجم و وزن میوه، آب میوه، آریل، پوست و هسته و نسبت‌هایی نظیر نسبت حجمی آریل به میوه، نسبت حجمی آب به میوه و نسب حجمی هسته به میوه را اندازه‌گیری کنند (Salmanizadeh et al., 2015). نتایج نشان داد که عامل رقم اثر معناداری بر مؤلفه‌های بیان شده به جز وزن و حجم میوه دارد. همچنین عامل اندازه اثر معناداری بر برخی از مؤلفه‌ها داشته است. بهترین ضرائب تبیین روابط تخمین برای حجم کل میوه بدست آمد که به ترتیب برای ارقام رباب ملس و رباب ترش ۹۷/۸ و ۹۷/۳ درصد محاسبه گردید. نتایج نشان داد روش غیرمخرب سی‌تی اسکن علیرغم هزینه زیاد برای تهیه تصاویر اسکن شده، به خوبی می‌تواند حجم انار و اجزاء درونی آن را تخمین بزند. در پژوهش دیگر نیز مشخص شد توموگرافی محاسباتی اشعه ایکس برآورد قابل قبولی از میزان آب موجود در میوه انار داشته است و اختلاف معناداری در سطح ۵ درصد بین مقدار آب پیش‌بینی شده و واقعی وجود نداشته است (Arendse et al., 2018).

طیف‌سنجی مرئی-مادون قرمز نزدیک ثابت کرده است که قادر به ارائه اطلاعات طیفی بسیار گسترده‌ای در مورد ویژگی‌ها و خواص فیزیکی و شیمیایی از درون و بیرون طیف وسیعی از مواد غذایی از جمله میوه‌ها است چون پرتوهای ساطع شده مادون قرمز نزدیک می‌تواند با پیوندهای ملکولی اساسی و مرتعش اکسیژن-هیدروژن، کربن-هیدروژن و نیتروژن-هیدروژن فعل و انفعال کرده و اطلاعاتی را در مورد پراکندگی ریز

² Nuclear Magnetic Resonance (NMR) Imaging

¹ Multispectral/Hyperspectral Imaging

فرآوری میوه انار نشان داد این مراکز از روش‌های سنتی و تجربی همراه با خطا برای تشخیص محتوای آب نمونه-های خریداری شده استفاده می‌کنند. بنابراین در پژوهش حاضر امکان‌سنجی استفاده از فناوری غیر-مخرب طیف‌سنجی مرئی-مادون قرمز نزدیک برای تخمین محتوای آب میوه انار به عنوان هدف اصلی مورد بررسی قرار گرفت.

مواد و روش‌ها

تهیه نمونه

میوه‌های انار شامل رقم‌های رباب، اتابکی، شیرین شهوار و خفری استفاده شده در این پژوهش از مجموعه باغ‌های قصردشت واقع در شیراز، استان فارس تهیه شدند. رقم-های تهیه شده از رایج‌ترین رقم‌های تجاری موجود در استان فارس هستند که رقم رباب دارای پوست قرمز با دانه‌های قرمز تیره رنگ، رقم اتابکی دارای پوست قرمز رنگ با دانه‌هایی قرمز، رقم خفری دارای پوست سبز مایل به زرد با دانه‌هایی قرمز رنگ و رقم شیرین شهوار (فاروق، فارس) با پوستی سبز مایل به زرد با دانه‌هایی صورتی رنگ هستند. تعداد میوه‌های تهیه شده در مجموع ۵۱۰ عدد بود که سهم رقم رباب ترش ۱۰۰ عدد، رباب ملس ۱۰۵ عدد، رقم اتابکی ۱۰۵ عدد، رقم شیرین شهوار ۱۰۰ عدد و رقم خفری ۱۰۰ عدد بود. نمونه‌ها به منظور جلوگیری از فساد بلافاصله پس از خریداری در یخچال با دمای ۴ درجه سلسیوس انبار شدند و برای جلوگیری از تغییر در ویژگی‌های فیزیکی مانند کاهش ضخامت پوست، با پلاستیک ضخیم پوشش‌دهی شد (Salmanizadeh *et al.*, 2015). لازم به ذکر است که نمونه‌ها در تاریخ ۲۵ آبان ۱۳۹۸ از باغ تهیه گردید و بلافاصله به سردخانه منتقل شد و تمامی آزمایش‌ها از تاریخ ۲۶ آبان تا ۲ آذر سال ۱۳۹۸ انجام شد. ارقام خریداری شده مطابق استاندارد سازمان خواروبار جهانی در چهار گروه اندازه A، B، C و D دسته-بندی شدند (جدول ۱) (FAO, 2013).

ساختارهای ماده گزارش دهند که به طور غیر مستقیم بیانگر مشخصات فیزیکی و شیمیایی آن ماده هستند (Arendse *et al.*, 2018). در پژوهشی با استفاده از طیف-های بازتابی و عبوری در بازه مرئی و مادون قرمز نزدیک با طول موج ۴۰۰ تا ۱۱۰۰ نانومتر سه شاخص مقدار مواد جامد محلول، اسیدیته و سفتی میوه انار بررسی شد. نتایج نشان داد که استفاده از هر دو حالت طیف-های بازتابی و عبوری برای ارزیابی شاخص‌های بیان شده امکان پذیر است. هر چند که طیف‌های بازتابی نتیجه دقیق‌تری را به دنبال داشت. به علاوه مدل رگرسیون حداقل مربعات جزئی به عنوان روش مناسب برای تخمین شاخص‌های مد نظر معرفی شد (Khodabakhshian *et al.*, 2019).

تخمین میزان آب موجود در محصولات کشاورزی در پژوهش‌های دیگری نیز انجام شده است (Blakey *et al.*, 2009; Neto *et al.*, 2017). روش طیف‌سنجی مادون قرمز نزدیک برای برآورد آب میوه آووکادو در دوره‌های مختلف رسیدگی بکاربرده شد. نتایج نشان داد که روش طیف‌سنجی در بازه ۱۱۰۰ تا ۲۰۰۰ نانومتر در ترکیب با مدل PLSR توانست میزان آب موجود در این میوه را با ضریب تبیین ۹۲ درصد پیش‌بینی کند (Blakey *et al.*, 2009). روش مشابه ترکیب مدل PLSR با روش طیف-سنجی مرئی و مادون قرمز نزدیک در بازه ۵۰۰ تا ۱۰۳۹ نانومتر برای اندازه‌گیری میزان آب موجود در برگ گل آفتاب‌گردان در سه حالت تنش آبی خفیف، میانه و شدید استفاده شد. در این مطالعه میزان آب موجود در برگ گیاه با شاخص میزان آب برگ (نسبت اختلاف وزن خشک و تر برگ بر وزن خشک بر حسب گرم) با استفاده از طیف‌های بازتابی اندازه‌گیری شد. نتایج ارزیابی مدل نشان داد طیف‌سنجی در طول موج‌های ۸۷۷، ۹۹۵ و ۱۰۳۹ نانومتر و مدل PLSR با ضریب تبیین ۸۴ درصد میزان آب موجود در برگ گل آفتاب‌گردان را برآورد نمود (Neto *et al.*, 2017).

بررسی‌های میدانی انجام شده از کارخانه‌های

استفاده شد. در این دستگاه، منبع نور لامپ‌های تنگستن و هالوژن بودند که به ترتیب قادر به انتشار امواج الکترومغناطیس در بازه طول موج مرئی و مادون قرمز نزدیک هستند. نور ساطع شده از منبع به وسیله فیبر نوری به واحد تکفام‌ساز تابیده می‌شد. تکفام‌ساز شامل دو المان برای تجزیه نور ورودی از منبع نور به طول موج‌های متفاوت و یک شکاف است که تنها اجازه عبور یک نور با یک طول-موج خاص را می‌داد. نور عبوری از شکاف تکفام‌ساز مستقیم توسط یک سلول وارد این محفظه می‌شد تا به نمونه‌ای که بر روی سلول قرار گرفته بود، تابیده شود. نور بعد از برخورد با نمونه به سطح شناساگر بازتاب می‌شد و این واحد نیز بر اساس آن میزان نور بازتابی، جذبی و یا عبوری را محاسبه کرده و سپس با فرمت دیجیتال گزارش می‌داد. در درون شناساگرها برای حس کردن پرتوهای نور مرئی از دیودهای نوری سیلیکونی و برای پرتوهای مادون قرمز نزدیک از ابزارهای ترکیبی با دیودهای نوری GaAs استفاده می‌شود. در دستگاه مورد استفاده در این پژوهش گیرنده و فرستنده امواج نور در یک واحد قرار داشتند. آزمایش‌ها در طیف جذبی بازه طول موج ۴۰۰ تا ۲۵۰۰ نانومتر با فواصل ۰/۵ نانومتر انجام شد. تمامی نمونه‌ها کاملاً سالم و فاقد هرگونه فساد و آسیب‌دیدگی بودند. قبل از طیف‌سنجی، دستگاه مطابق روش توصیه شده بدون وجود نمونه واسنجی (کالیبره) شد. رابطه کاربری با دستگاه طیف سنج از نرم‌افزار Vision Air بود. اطلاعات مختلفی نظیر بازه طول موج، نوع شناساگر، نوع سلول استفاده شده، دمای محیط توسط اپراتور دستگاه تعریف شد تا داده‌برداری بر اساس آن انجام شود. برای ایجاد شرایط پایدار دستگاه، نمونه‌برداری ۴۰ الی ۴۵ دقیقه بعد از روشن شدن آن صورت گرفت. داده‌های برداشت شده از نمونه ابتدا به صورت خودکار بر روی نرم‌افزار دستگاه ذخیره شد و سپس به فایل اکسل منتقل گردید.

جدول ۱- استاندارد سازمان خوار و بار جهانی (FAO) برای طبقه‌بندی میوه انار براساس اندازه

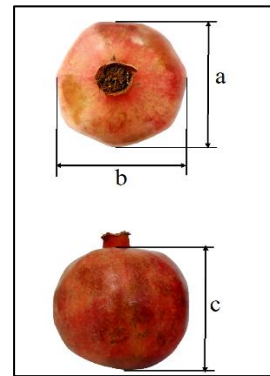
Table 1- FAO standard for classification of pomegranate fruit based on size

کلاس اندازه	قطر هندسی (mm)
A	≥۸۱
B	۷۱-۸۰
C	۶۱-۷۰
D	۵۱-۶۰
E	۴۰-۵۰

برای اندازه‌گیری قطر نمونه‌ها از قطر هندسی مطابق رابطه (۱) استفاده شد (Mohsenin, 1996).

$$D = (abc)^{\frac{1}{3}} \quad (\text{رابطه ۱})$$

در رابطه (۱) قطر اول، b قطر دوم (قطرهای استوایی) و c ارتفاع میوه است که a و b در یک صفحه و عمود بر یکدیگر و c عمود بر آن دو است (شکل ۱). اندازه‌گیری ابعاد با کولیس دیجیتال (Guanglu ساخت کشور چین) با دقت $\pm 0/01$ میلی‌متر انجام شد.



شکل ۱- ابعاد میوه انار برای محاسبه قطر هندسی (a و b دو قطر استوایی عمود بر هم، c ارتفاع)

Figure 1- Dimensions of pomegranate fruit for computing geometric diameter (a and b two equatorial diameter, c height)

طیف‌سنجی نمونه‌ها و تعیین حجم آب

برای سنجش پاسخ طیفی نمونه‌های انار، از دستگاه اسپکتروفوتومتر مرئی-مادون قرمز نزدیک (مدل NIR XDS ساخت کمپانی Metrohm کشور سوئیس)

(رابطه ۳)

$$\text{Upper Limit (UL)} = Q_3 + 1.5(IQR)$$

(رابطه ۴)

$$\text{Lower Limit (LL)} = Q_1 - 1.5(IQR)$$

از آنجا که داده‌های ثبت شده علاوه بر داشتن اطلاعات ارزشمند دارای اطلاعات نامربوط و نوفه‌هایی نیز هستند که به دلایلی نظیر پراکندگی نور، بروز تفاوت در سطح تماس نمونه و شناساگر دستگاه طیف‌نگار هنگام نمونه‌برداری‌های متعدد و ویژگی‌های فیزیکی نمونه مثل سطح آن رخ می‌دهند (Shirzadifar et al., 2018)، پنج روش پیش‌پردازش نوفه‌ها بر اساس بررسی منابع و سعی و خطا انتخاب شد (Khodabakhshian et al., 2019, Shirzadifar et al., 2018, Arendse et al., 2018). این روش‌ها عبارت از تصحیح پراکندگی ضربی^۲، متغیر نرمال استاندارد^۳، نرمال‌سازی بردار^۴، مشتق اول^۵ و مشتق دوم^۶ بود. در شکل ۲ امضای طیفی مجموعه داده بدون پیش‌پردازش در مقایسه با سه روش MSC، SNV و VN که بهترین پاسخ‌ها را به همراه داشته‌اند آمده است و در شکل ۳ تاثیر دو پیش‌پردازش مشتق اول و دوم بر امضای طیفی قابل مشاهده است.

پس از اعمال روش‌های پیش‌پردازش، ناهنجاری‌های باقی مانده موجود در داده‌های طیف‌برداری با استفاده از روش شناسایی ناهنجاری^۷ جنگل انزوا^۸ شناسایی و حذف شد (Xu et al., 2017).

کاهش ابعاد و یافتن مؤلفه‌های اصلی

هر کدام از طول‌موج‌هایی که به نمونه‌ها ساطع شده‌اند یک بعد یا یک ویژگی محسوب می‌شوند. نکته حائز اهمیت این است که همه این ویژگی‌ها در ساخت مدل کالیبراسیون به یک اندازه دارای اهمیت نیستند و این اهمیت را می‌توان از روی پراکندگی هر ویژگی از نمونه‌ای به نمونه دیگر تعیین کرد. به این مفهوم که اگر با تغییر در نمونه‌ها مقادیر جذب در یک طول‌موج خاص،

برای این که پاسخ‌های طیفی از اعتبار بیشتری برخوردار باشند از هر نمونه دوبار و در دو جهت مختلف طیف‌برداری شد. با توجه به هماهنگی و روند مشابه در نمودارهای طیفی بدست آمده در دو مرحله تابش، از طیف‌های هر نمونه میانگین‌گیری شد تا به ازای هر نمونه یک دسته پاسخ طیفی به همراه میزان جذب نور حاصل شود.

در طیف‌سنجی، غلظت نمونه با واحد مول بر لیتر و طول مسیر نور با واحد سانتی‌متر سنجیده می‌شود. بنابراین واحدهای جذب مولی لیتر بر مول بر سانتی‌متر هستند. از آنجا که واحد جذب مولی همواره به همین شکل است به طور معمول بدون واحد اعلام می‌شود (Hagen & Sanders, 2007). محتوای آب نمونه‌های طیف‌سنجی شده با استفاده از دستگاه آبمیوه‌گیری دستی استخراج شد و حجم آن با استوانه مدرج با دقت ± 1 میلی‌لیتر اندازه‌گیری شد. دستگاه آب میوه‌گیر دارای مکانیزم فشاری برای خارج کردن کامل آب با فشار یکسان برای همه نمونه‌ها بود (Salmanizadeh et al., 2015).

مدل ریاضی برای برآورد حجم آب

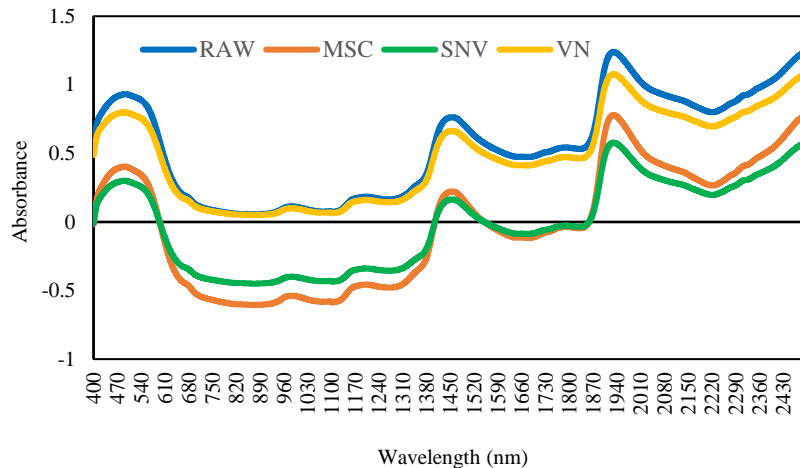
برای ایجاد مدل ریاضی برآورد حجم آب اندازه‌گیری شده بر اساس پاسخ‌های طیفی، الگوریتم‌های یادگیری ماشین و در نرم افزار PYTHON 3.8 پیاده سازی شد. ابتدا پیش‌پردازش داده‌ها صورت گرفت و داده‌های پرت با استفاده از روش دامنه میان چارکی^۱ حذف گردید (روابط ۲ تا ۴) (Saleem et al., 2021). برای این منظور دامنه اعداد مشخص و مقادیر چارک اول (Q_1) و سوم (Q_3) برای حجم آب به طور جداگانه محاسبه گردید. سپس محدوده بالا (UL) و محدوده پایین (LL) محاسبه و اعداد خارج این دامنه به عنوان داده پرت حذف شد.

$$IQR = Q_3 - Q_1 \quad (\text{رابطه ۲})$$

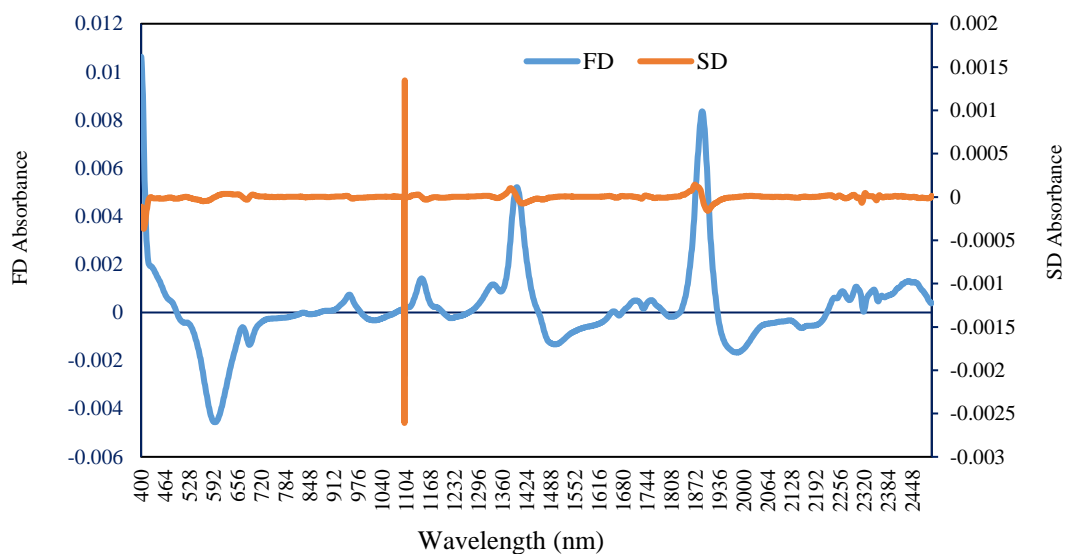
5 First Derivative (FD)
6 Second Derivative (SD)
7 Anomaly Detection
8 Isolation Forest

1 Interquartile Range (IQR)
2 Multiplicative Scatter Correction (MSC)
3 Standard Normal Variate (SNV)
4 Vector Normalization (VN)

دستخوش تغییر نشوند و یا به عبارت دیگر پراکندگی پایینی داشته باشند، می توان نتیجه گرفت که آن طول- موج خاص اهمیت بالایی ندارد و قادر به نشان دادن اختلاف بین نمونه ها نیست به همین رو می توان به منظور کاهش حجم محاسبات، و ساده تر کردن مدل آن را حذف کرد.



شکل ۲- مقایسه امضای طیفی حالت بدون پیش پردازش (RAW)، تصحیح پراکندگی ضربی (MSC)، متغیر نرمال استاندارد (SNV)، نرمال سازی بردار (VN)
 Figure 2- Comparison between spectral signatures for: without preprocessing (RAW), Multiplicative Scatter Correlation (MSC), Standard Normal Variate (SNV), Vector Normalization (VN)



شکل ۳- امضای طیفی با پیش پردازش مشتق اول (FD)، مشتق دوم (SD)
 Figure 3- Spectral signature with First Derivative (FD), Second Derivative (SD)

دارند را به خوبی فراهم می کند. ویژگی دیگر این روش این است که نظارت شده^۱ است، یعنی به هنگام یافتن مؤلفه های اصلی هم از مقادیر X (ویژگی ها) و هم از مقادیر Y (برچسب ها) بهره می برد و بارگذاری آن توسط تجزیه مقدار منفرد S که همان ماتریس حاصل شده از

به منظور کاهش ابعاد و تعیین مؤلفه های اصلی از روش حداقل مربعات جزئی استفاده شد. الگوریتم کاهش ابعاد PLS امکان کار با مجموعه داده هایی که در آن ها تعداد مشاهدات از تعداد ویژگی ها کمتر هستند یا به عبارت دیگر چندخطی بوده و تعداد متغیرهای بالایی

(Baek et al., 2020).

$$RSME = \sqrt{\frac{\sum (y_{cal} - y_{act})^2}{n}} \quad (\text{رابطه ۵})$$

$$RPD = \frac{SD}{RMSE} \quad (\text{رابطه ۶})$$

$$R^2 = 1 - \frac{\sum (y_{cal} - y_{act})^2}{\sum (y_{cal} - y_{mean})^2} \quad (\text{رابطه ۷})$$

در روابط بالا y_{cal} ، مقادیر حجم پیش‌بینی شده در مرحله کالیبراسیون، y_{act} مقادیر واقعی حجم آب، y_{mean} مقدار میانگین حجم آب پیش‌بینی شده، SD انحراف معیار مقادیر حجم آب واقعی است و n تعداد نمونه‌ها است.

در ارزیابی بیرونی مدل‌ها علاوه بر ضریب تبیین و ریشه میانگین مربعات خطا، از میانگین خطای سوگیری، میانگین خطای مطلق و شاخص ویلموت^۱ (ضریب فاصله) برای دقت و صحت برازش در مقایسه با خروجی واقعی استفاده شد (Hagen, 2007; Khodabakhshian et al., 2017; Fazayeli et al., 2019).

(رابطه ۸)

$$MBE = \frac{1}{n} \sum (y_{pred} - y_{act}) \quad (\text{رابطه ۹})$$

$$MAE = \frac{1}{n} \sum |y_{pred} - y_{act}| \quad (\text{رابطه ۱۰})$$

$$d = 1 - \frac{\sum (y_{act} - y_{pred})^2}{\sum (|y_{pred} - \bar{y}_{act}| + |y_{act} - \bar{y}_{act}|)^2}$$

تحلیل نتایج

بررسی ویژگی‌های امضای طیفی میوه انار امضای طیفی میوه‌های انار که در دستگاه طیف‌سنج قرار گرفتند به صورت نموداری از میزان مقادیر جذب امواج بر حسب طول موج اعمال شده است. در این نمودار نقاط اوج متعددی وجود دارد که هر کدام از آن‌ها بیانگر یک ویژگی خاص از میوه انار هستند که در این جا به بررسی

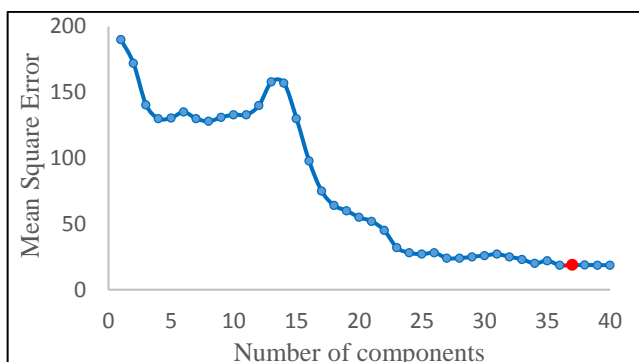
برهمکنش مقادیر X و Y است، انجام می‌شود (Bulut & Alma, 2010). برای ایجاد مدل‌ها از یک تا ۴۰ مؤلفه اساسی استفاده شد و نهایتاً تعداد مؤلفه اساسی که منجر به کمترین خطا (ریشه میانگین مربعات خطا) شد، به عنوان تعداد مطلوب مؤلفه‌های اصلی در نظر گرفته شد (Xiao et al., 2019).

پس از اتمام پیش‌پردازش و آماده‌سازی داده‌ها از روش رگرسیون حداقل مربعات جزئی (PLSR) برای ایجاد الگوریتم تخمین حجم آب میوه استفاده شد (Khodabakhshian et al., 2019; Xiao et al., 2019; Pourdarbani et al., 2020). روش PLSR یک روش چند متغیره است که در تجزیه و تحلیل طیف‌سنجی مرئی-مادون قرمز نزدیک به منظور ایجاد ارتباط بین مجموعه-داده طیف‌ها (X) و شاخص‌های فیزیکی یا شیمیایی اندازه‌گیری شده (Y) کاربرد فراوان دارد (Arendse et al., 2017; Khodabakhshian et al., 2017).

پیش‌بینی حجم آب

عملکرد مدل PLSR از طریق روش ارزیابی تقابل ده‌تایی مورد ارزیابی قرار گرفت. برای این منظور داده‌ها به صورت تصادفی به دو دسته آموزش و ارزیابی تقسیم شدند و این روند ۱۰ مرتبه تکرار گردید. در هر مرتبه مدل PLSR توسط داده‌های آموزشی ساخته شد و توسط داده‌های مختص ارزیابی مورد ارزیابی قرار گرفت. ۷۰ درصد از داده‌ها به شکل تصادفی برای آموزش مدل رگرسیون و ۳۰ درصد باقی مانده برای ارزیابی دقت مدل مورد استفاده قرار گرفت (Shirzadifar et al., 2020). برای بررسی تفاوت پاسخ‌ها در کل بازه، بازه مرئی و بازه مادون قرمز نزدیک، سه بازه ۴۰۰ تا ۲۵۰۰ نانومتر، ۴۰۰ تا ۹۰۰ نانومتر و ۲۵۰۰ تا ۹۰۰ نانومتر به صورت جداگانه مورد تحلیل و بررسی قرار گرفتند. ریشه میانگین مربعات خطا (RMSE)، انحراف باقی‌مانده پیش‌بینی (RPD) و ضریب تبیین (R^2) برای بررسی مدل در مرحله کالیبراسون استفاده شد (Bulut et al., 2010; FAO, 2013).

تعیین شدند و سپس مدل ریاضی با همان تعداد از مؤلفه ساخته شد (Pelliccia, 2018) (شکل ۵).



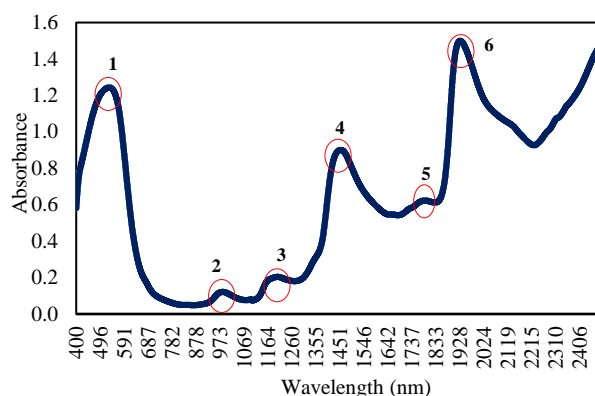
شکل ۵- تعیین تعداد مؤلفه اصلی بهینه بر اساس مقدار میانگین مربع خطا
Figure 5- Determination of number of principle components based on Mean Square Error

به عنوان مثال برای ایجاد یک مدل، از ۱ تا ۴۰ مؤلفه اساسی مورد استفاده قرار گرفت و هنگامی که مدل با ۳۶ مؤلفه بنا شد، کمترین میانگین خطای استاندارد (MSE) با مقدار ۱۸/۶۱ حاصل شد. محدود شدن تعداد مؤلفه‌های مورد بررسی به ۴۰ عدد به این علت بود که بعد از این محدوده مقدار میانگین خطای استاندارد دستخوش تغییرات چندانی نشد. این روند در کلیه حالات استفاده شده در پژوهش پیاده‌سازی شد. در جدول ۲ تعداد مؤلفه اصلی بهینه بر اساس حداقل MSE برای همه مدل‌ها نشان داده شده است.

تخمین حجم آب و نتایج حاصل شده

برای تخمین حجم آب میوه انار ابتدا داده‌های طیفی در شش حالت شامل داده‌های خام و داده‌های پیش‌پردازش شده با روش‌های تصحیح پراکندگی ضربی (MSC)، متغیر نرمال استاندارد (SNV)، نرمال‌سازی بردار (VN)، مشتق اول (FD) و مشتق دوم (SD) آماده شدند و سپس در سه بازه طول موج ۴۰۰ تا ۹۰۰، ۹۰۰ تا ۲۵۰۰ و ۴۰۰ تا ۲۵۰۰ نانومتر، با استفاده از روش رگرسیون حداقل مربعات جزئی تخمین زده شدند. گرچه بازه ۴۰۰ تا ۹۰۰ نانومتر بر اساس توضیحات قبل به دلیل وجود رنگدانه‌های موجود در پوست انار است اما باید همواره

برخی از آن‌ها پرداخته شده است (شکل ۴).



شکل ۴- نمونه امضای طیفی میوه انار (میزان جذب بر حسب طول موج)
Figure 4- Typical spectral signature of pomegranate fruit (amount of absorption in wavelength unit)

نقطه اوج شماره ۱ در بازه ۴۵۰ تا ۶۵۰ نانومتر رخ داد که بیانگر وجود رنگدانه‌های موجود در پوست انار است. این رنگدانه‌ها طول‌موج‌های مذکور را که در محدوده مرئی هستند بیشتر جذب می‌کنند. نقطه اوج شماره ۲ در محدوده ۹۳۰ تا ۹۸۰ نانومتر رویداد که به دلیل وجود پیوندهای بین ملکولی اکسیژن-هیدروژن موجود در دانه‌های انار است (Khodabakhshian *et al.*, 2019). نقاط اوج ۳، ۵ و ۶ موجود در محدوده ۱۱۰۰ تا ۱۲۰۰ نانومتر، ۱۷۵۰ تا ۱۸۰۰ نانومتر و ۱۸۵۰ تا ۱۹۵۰ نانومتر مربوط به پیوندهای ملکولی تغییر شکل یافته اکسیژن-هیدروژن و کربن-هیدروژن موجود در قند محلول در آب انار است و نقطه اوج شماره ۴ ایجاد شده در محدوده ۱۳۵۰ تا ۱۴۵۰ نانومتر نیز مربوط به پیوند-های ملکولی اکسیژن-هیدروژن مرتبط با جذب آب انار است (Arendse *et al.*, 2018). توضیحات قبل بیانگر یک بررسی اجمالی از امضای طیفی میوه انار است که در ادامه در مورد هر بازه به صورت جداگانه بحث و نتیجه‌گیری شده است.

تعیین تعداد مؤلفه اساسی بهینه

هنگام ساخت مدل‌های رگرسیون در ابتدا تعداد مؤلفه‌های اساسی بهینه که کمترین MSE را ایجاد می‌کردند،

مرئی به منظور پیش‌بینی حجم آب میوه انار را نیز توجیه می‌کند. نتایج حاصل شده به شرح جدول‌های ۲ و ۳ است.

این نکته را در نظر داشت که نقاط اوج موجود در امضای طیفی را نمی‌توان به یک عامل مرتبط دانست و عوامل متنوعی در به وجود آمدن این نقاط مؤثر هستند (Arendse et al., 2016)، که این موضوع بررسی بازه

جدول ۲- نتایج مرحله کالیبراسیون مدل رگرسیون حداقل مربعات جزئی (PLSR) برای تخمین حجم آب میوه انار

Table 2- Results of the calibration step of Partial Least Square Regression (PLSR) for estimation of pomegranate juice volume content

RPD	R ²	RMSE	MSE	تعداد مؤلفه بهینه	روش پیش‌پردازش	طول موج (nm)
۰/۴۲	۰/۹۲	۸/۳۹	۷۰/۴۴	۲۱	تصحیح پراکندگی ضربی (MSC)	۹۰۰-۴۰۰
۰/۵۳	۰/۹۳	۷/۳۶	۵۴/۲۹	۲۰	متغیر نرمال استاندارد (SNV)	۹۰۰-۴۰۰
۰/۴۳	۰/۹۳	۸/۴۴	۷۱/۲۶	۲۶	نرمال‌سازی بردار (VN)	
۰/۴۵	۰/۹۳	۸/۱۷	۶۶/۷۹	۱۶	تصحیح پراکندگی ضربی (MSC)	
۰/۵۱	۰/۹۳	۷/۴۶	۵۵/۷۵	۱۸	متغیر نرمال استاندارد (SNV)	۲۵۰۰-۹۰۰
۰/۷۳	۰/۹۵	۶/۱۷	۳۸/۱۲	۲۱	نرمال‌سازی بردار (VN)	
۰/۵۶	۰/۹۴	۷/۱۹	۵۱/۷۲	۱۸	تصحیح پراکندگی ضربی (MSC)	
۰/۵۵	۰/۹۸	۴/۳۱	۱۸/۶۱	۳۶	متغیر نرمال استاندارد (SNV)	۲۵۰۰-۴۰۰
۰/۵۰	۰/۹۳	۷/۶۴	۵۸/۵۱	۸	نرمال‌سازی بردار (VN)	

RMSE، MSE و RPD به ترتیب میانگین مربع خطا، ریشه میانگین مربعات خطا، و انحراف باقی‌مانده پیش‌بینی است.

جدول ۳- نتایج حاصل از ارزیابی مدل رگرسیون حداقل مربعات جزئی (PLSR) برای تخمین حجم آب میوه انار

Table 3- Results of the verification step of Partial Least Square Regression (PLSR) for estimation of pomegranate juice volume content

MBE	MAE	d	R ²	RMSE	روش پیش‌پردازش	طول موج (nm)
-۶/۴۶	۱۳/۰۶	۰/۹۲	۰/۸۲	۲۱/۵۷	تصحیح پراکندگی ضربی (MSC)	۹۰۰-۴۰۰
-۲/۲۴	۱۰/۵۰	۰/۹۵	۰/۸۸	۱۳/۵۳	متغیر نرمال استاندارد (SNV)	۹۰۰-۴۰۰
-۳/۶۰	۱۱/۷۶	۰/۹۴	۰/۸۳	۱۶/۸۱	نرمال‌سازی بردار (VN)	
-۶/۰۷	۱۱/۷۶	۰/۹۳	۰/۸۴	۱۸/۹۶	تصحیح پراکندگی ضربی (MSC)	
-۳/۸۲	۸/۴۲	۰/۹۷	۰/۸۹	۱۲/۳۸	متغیر نرمال استاندارد (SNV)	۲۵۰۰-۹۰۰
-۸/۴۶	۱۲/۹۸	۰/۹۲	۰/۸۵	۲۰/۸۶	نرمال‌سازی بردار (VN)	
-۸/۹۶	۱۴/۰۸	۰/۹۱	۰/۸۲	۲۴/۰۷	تصحیح پراکندگی ضربی (MSC)	
-۰/۳۳	۵/۳۰	۰/۹۸	۰/۹۴	۸/۴۱	متغیر نرمال استاندارد (SNV)	۲۵۰۰-۴۰۰
-۵/۵۸	۱۱/۲۳	۰/۹۴	۰/۸۵	۱۷/۱۲	نرمال‌سازی بردار (VN)	

MBE، RMSE، MAE و d به ترتیب میانگین خطای سوگیری، ریشه میانگین مربعات خطا، میانگین خطای مطلق و شاخص ویلموت است.

ناهمواری‌ها باعث ایجاد نوفه‌هایی در طیف‌های تابیده شده می‌شود. از آنجا که پیش‌پردازش SNV فرایند هموارسازی نوفه‌های طیفی را در سطح محدودی نسبت به سایر روش‌ها انجام می‌دهد و به نوفه‌های موجود در طیف‌های برداشته شده از انار حساس‌تر است، نتایج بهتری را نسبت به سایر روش‌های پیش‌پردازش به دست

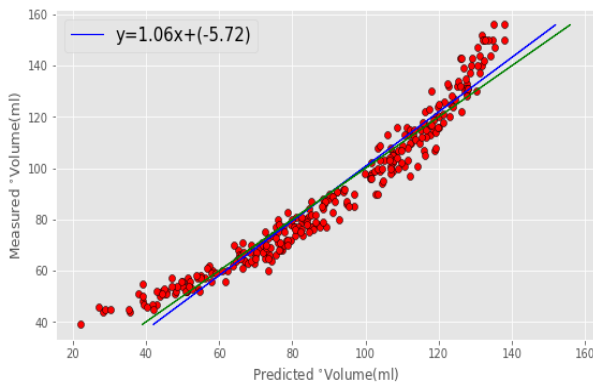
نتایج پیش‌بینی حجم آب در جدول ۳ نشان می‌دهد که در هر سه بازه ۴۰۰ تا ۹۰۰، ۹۰۰ تا ۲۵۰۰ و ۲۵۰۰ تا ۴۰۰ نانومتر خروجی‌های آماری حاصل شده از روش پیش‌پردازش SNV وضعیت بهتری دارند که ممکن است به این علت باشد که سطح پوست میوه انار و اجزاء داخلی آن صاف و یکدست نیستند و این

داده است (Shirzadifar et al., 2018).

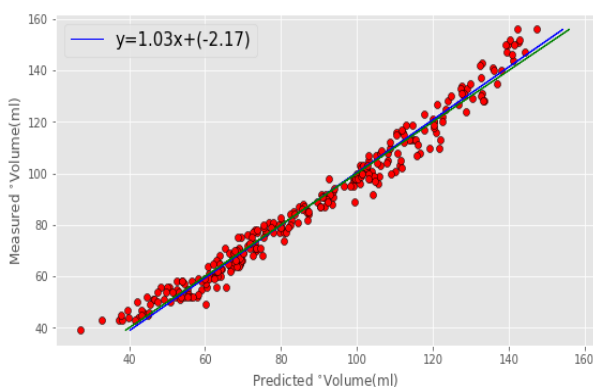
همانطور که قبلا بیان شد وجود پیوندهای ملکولی اکسیژن-هیدروژن موجود در آب دانه‌های انار طول موج‌هایی در بازه ۱۳۵۰ تا ۱۴۵۰ نانومتر را بیشتر به خود جذب می‌کنند، لذا دو بازه ۹۰۰ تا ۲۵۰۰ و ۴۰۰ تا ۲۵۰۰ نانومتر نسبت به بازه ۴۰۰ تا ۹۰۰ نانومتر در تخمین حجم آب بهتر عمل کرده‌اند (Arendse et al., 2017).

با توجه به جداول ۲ و ۳، پیش‌بینی حجم آب با پیش‌پردازش SNV در بازه ۴۰۰ تا ۹۰۰ نانومتر در مرحله آموزش دارای $RMSE=7/36$ و $R^2=0/93$ و در مرحله ارزیابی دارای $RMSE=13/50$ و $R^2=0/88$ بود. بازه ۴۰۰ تا ۲۵۰۰ نیز با پیش‌پردازش SNV در مرحله آموزش دارای $RMSE=4/31$ و $R^2=0/98$ و در مرحله ارزیابی دارای $RMSE=8/41$ و $R^2=0/94$ بود. با توجه به این توضیحات نتایج مرحله آموزش در دو بازه ۴۰۰ تا ۹۰۰ و ۴۰۰ تا ۲۵۰۰ نانومتر نتایج مرحله ارزیابی را تایید می‌کند، اما در بازه ۹۰۰ تا ۲۵۰۰ در مرحله کالیبراسیون پیش‌پردازش VN با $RMSE=6/17$ و $R^2=0/95$ بهترین نتیجه را به خود اختصاص داد در حالی که در مرحله ارزیابی پیش‌پردازش SNV با $RMSE=12/38$ و $R^2=0/89$ بهترین نتیجه بود. در توضیح می‌توان بیان کرد که پیش‌پردازش VN از لحاظ عملکردی شباهت بالایی در تاثیرگذاری بر روی داده‌ها به پیش‌پردازش SNV دارد، از همین رو ممکن است گاهی اوقات پاسخ‌های بهتری را ارائه دهد (Shirzadifar et al., 2018).

نمودارهای آموزش و ارزیابی حجم آب واقعی بر اساس حجم آب پیش‌بینی شده با مدل PLSR برای بهترین (SNV در بازه ۴۰۰ تا ۲۵۰۰) و ضعیف‌ترین (MSC در بازه ۴۰۰ تا ۲۵۰۰) نتیجه در شکل‌های ۶ و ۷ نمایش داده شده‌اند (خط $y=x$ با رنگ سبز مشخص شده است).



(a)



(b)

شکل ۶- نمودارهای آموزش حجم آب پیش‌بینی شده در بازه ۴۰۰ تا ۲۵۰۰ نانومتر با پیش‌پردازش: (a) تصحیح پراکندگی ضربی (MSC)، (b) متغیر نرمال استاندارد (SNV) پومرانیته در بازه ۴۰۰ تا ۲۵۰۰ نانومتر با پیش‌پردازش: (a) تصحیح پراکندگی ضربی (MSC)، (b) Standard Normal Variate (SNV)

در پژوهشی سعی شد با استفاده از توموگرافی محاسباتی اشعه ایکس حجم آب انار به صورت غیر-مخرب تعیین شود. حجم آب اندازه‌گیری شده در روش مخرب $146/1 \pm 16/3$ و در روش غیرمخرب $142/16 \pm 7/4$ گزارش شد. نتایج آزمون t نشان می‌دهد که اختلاف بین مقادیر مشاهده شده و پیش‌بینی شده در سطح $0/05$ معنا دار نیست و روش سی‌تی‌اسکن به طور مناسب قادر به تخمین حجم آب انار است (Arendse et al., 2016). در پژوهش دیگر نیز حجم آب انار با استفاده از پردازش تصاویر حاصل از روش سی‌تی‌اسکن اشعه ایکس تخمین زده شد. بهترین پاسخ در تخمین حجم آب در این روش با ضریب تبیین $91/9$ درصد همراه بود

راحت تری را برای نمونه برداری نسبت به روش سی تی-اسکن فراهم می کند. به علاوه بنابر بررسی های میدانی صورت گرفته و نظر محققان پیشین هزینه نمونه برداری با دستگاه سی تی اسکن بسیار بیشتر از دستگاه طیف-سنجی است که خود دلیلی بر برتری استفاده از این روش است (Arendse *et al.*, 2017).

ارقام استفاده شده در این پژوهش طبق استاندارد سازمان خوار و بار جهانی در سه اندازه A، B و C تهیه شدند تا از نظر حجم آب بایکدیگر تفاوت داشته باشند و نتایج تخمین حجم آب نیز بیانگر قابلیت فناوری طیف سنجی مرئی-مادون قرمز نزدیک در تشخیص این تفاوت ها است. نکته حائز اهمیت این است که الگوریتم ارائه شده در این پژوهش قادر است تا حجم آب میوه انار را فارغ از رقم آن تخمین بزند و می توان نتیجه گرفت که فناوری طیف سنجی قادر به تخمین حجم آب سایر ارقام انار نیز است.

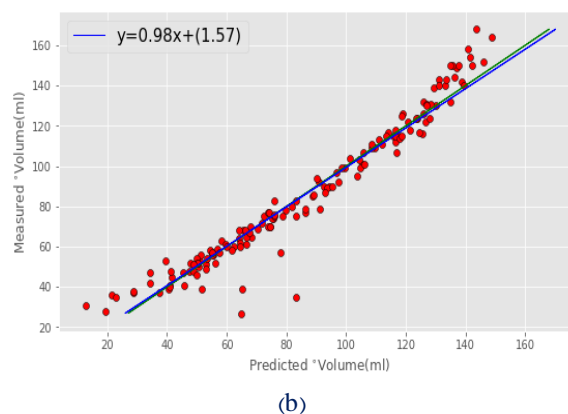
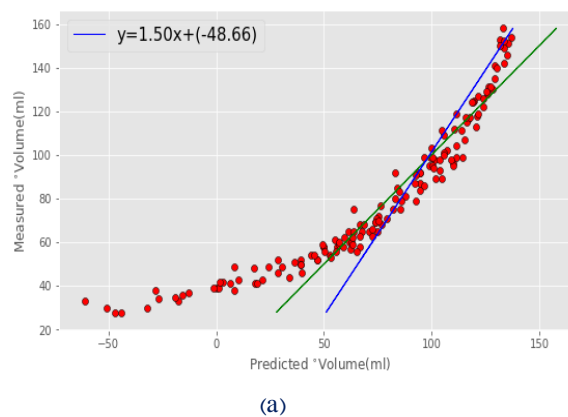
نتیجه گیری

هدف این پژوهش امکان سنجی استفاده از فناوری غیر-مخرب طیف سنجی مرئی-مادون قرمز نزدیک در تعیین حجم آب میوه انار بود که نتایج این امکان سنجی نشان داد که می توان حجم آب را با دقت بسیار بالا تعیین کرد. همچنین تعیین الگوریتم های مناسب به منظور تعیین این ویژگی بررسی شد و می توان در یک نتیجه گیری کلی بیان کرد که ترکیب علم یادگیری ماشین با داده های طیف نگاری و به کارگیری مدل PLSR در تخمین حجم آب با ثبت ضریب تبیین ۹۴ درصد و میانگین خطای مطلق ۵/۳، موفقیت آمیز بوده و راهی جدید را برای سایر پژوهشگران باز کرده است تا این مسیر را ادامه و گسترش دهند.

REFERENCES

- Ahmadi, K., Ebadzadeh H. R., Hatami, F., Hosseinpour, R., & Abdeshah, H. (2020). Agricultural statistics (vol. 3). Communication and Information Technology Center, Planning and Economic Deputy, Ministry of Hihad-e-Agriculture. Retrieved from https://www.maj.ir/Dorsapax/user_files/Sub65/Amarnamehj1-97-98-site.pdf on date (Sept. 2021). (in Farsi)
- Arendse, E., Fawole, O. A., Magwaza, L. S., & Opara, U. L. (2016). Non-destructive characterization

و نتایج آزمون چاو نشان داد که اختلاف بین حجم آب مشاهده شده و پیش بینی شده معنادار نیست (Salmanizadeh *et al.*, 2015).



شکل ۷- نمودارهای ارزیابی حجم آب پیش بینی شده در بازه ۴۰۰ تا ۲۵۰۰ نانومتر با پیش پردازش: (a) تصحیح پراکندگی ضربی (MSC)، (b) متغیر نرمال استاندارد (SNV)

Figure 7- Evaluation curves of estimated pomegranate juice content volume in 400 to 2500 nm interval with preprocessing: (a) Multiplicative Scatter Correction (MSC), (b) Standard Normal Variate (SNV)

در پژوهش حاضر با استفاده از طیف سنجی حجم آب انار با ضریب تبیین ۹۴ درصد تخمین زده شد که از لحاظ دقت اندازه گیری فراتر از روش سی تی اسکن است. همچنین فناوری طیف سنجی مرئی-مادون قرمز نزدیک با داشتن ابزارهای ساده و قابل حمل شرایط به مراتب

- and volume estimation of pomegranate fruit external and internal morphological fractions using X-ray computed tomography. *Journal of Food Engineering*, 186, 42-49.
- Arendse, E., Fawole, O. A., Magwaza, L. S., Nieuwoudt, H. H., & Opara, U. L. (2017). Development of calibration models for the evaluation of pomegranate aril quality by Fourier-transform near infrared spectroscopy combined with chemometrics. *Biosystems Engineering*, 159, 22-32.
- Arendse, E., Fawole, O. A., Magwaza, L. S., Nieuwoudt, H., & Opara, U. L. (2018). Fourier transform near infrared diffuse reflectance spectroscopy and two spectral acquisition modes for evaluation of external and internal quality of intact pomegranate fruit. *Postharvest Biology and Technology*, 138, 91-98.
- Baek, S., Lim, J., Lee, J. G., McCarthy, M. J., & Kim, S. M. (2020). Investigation of the Maturity Changes of Cherry Tomato Using Magnetic Resonance Imaging. *Applied Sciences*, 10(15), 5188.
- Bashghareh, A. (2019). The effect of Pre-harvest chitosan application on quantitative and qualitative characteristics of pomegranate fruit. M.Sc. Thesis on Horticultural Sciences Engineering, Gorgan University. Gorgan, Iran. (in Farsi)
- Blakey, R. J., Bower, J. P., & Bertling, I. (2009). Influence of water and ABA supply on the ripening pattern of avocado (*Persea americana* Mill.) fruit and the prediction of water content using Near Infrared Spectroscopy. *Postharvest Biology and Technology*, 53(1-2), 72-76.
- Bulut, E., & Alma, Ö. G. (2010). Dimensionality Reduction Methods: PCR, PLSR, RRR and health application. *Physical Sciences*, 6(2), 36-47.
- FAO. 2013. Standard for Pomegranate (Codex 310). Retrieved from <http://www.fao.org/home/search/en/?q=standardforpomegranate> (on date Sept. 2021).
- Fazayeli, A., Kamgar, S., Nassiri, S. M., Fazayeli, H., & De La Guardia, M. (2019). Dielectric spectroscopy as a potential technique for prediction of kiwifruit quality indices during storage. *Information Processing in Agriculture*, 6(4), 479-486.
- Hagen, C. L., & Sanders, S. T. (2007). Investigation of multi-species (H₂O₂ and H₂O) sensing and thermometry in an HCCI engine by wavelength-agile absorption spectroscopy. *Measurement Science and Technology*, 18(7), 1992.
- Khodabakhshian, R., Emadi, B., Khojastehpour, M., Golzarian, M. R., & Sazgarnia, A. (2017). Non-destructive evaluation of maturity and quality parameters of pomegranate fruit by visible/near infrared spectroscopy. *International Journal of Food Properties*, 20(1), 41-52.
- Khodabakhshian, R., Emadi, B., Khojastehpour, M., & Golzarian, M. R. (2019). A comparative study of reflectance and transmittance modes of Vis/NIR spectroscopy used in determining internal quality attributes in pomegranate fruits. *Journal of Food Measurement and Characterization*, 13(4), 3130-3139.
- Mohsenin, N. N. (1996). *Physical properties of plant and animal materials* (vol. 1). Gordon Publication. Canada.
- Munera, S., Hernández, F., Aleixos, N., Cubero, S., & Blasco, J. (2019). Maturity monitoring of intact fruit and arils of pomegranate cv. 'Mollar de Elche' using machine vision and chemometrics. *Postharvest Biology and Technology*, 156, 110936.
- Neto, A. J. S., Lopes, D. C., Pinto, F. A., & Zolnier, S. (2017). Vis/NIR spectroscopy and chemometrics for non-destructive estimation of water and chlorophyll status in sunflower leaves. *Biosystems Engineering*, 155, 124-133.
- Patel, M. B., Nath, A., & Mayani, J. M. (2018). A study on physical properties of pomegranate (*Punica granatum* L., *Punicaceae*) fruits. *International Journal of Communication Systems*, 6(5), 1460-1463.
- Pelliccia, D. (2018). Partial Least Square Regression in Python. Retrieved from <https://nirpyresearch.com/partial-least-squares-regression-python>.
- Pourdarbani, R., Sabzi, S., Kalantari, D., & Arribas, J. I. (2020). Non-destructive visible and short-wave near-infrared spectroscopic data estimation of various physicochemical properties of Fuji

- apple (*Malus pumila*) fruits at different maturation stages. *Chemometrics and Intelligent Laboratory Systems*, 206, 104147.
- Saleem, S., Aslam, M., & Shaukat, M. R. (2021). A review and empirical comparison of univariate outlier detection methods. *Pakistan Journal of Statistics*, 37(4), 447-462.
- Salmanizadeh, F., Nassiri, S. M., Jafari, A., & Bagheri, M. H. (2015). Volume estimation of two local pomegranate fruit (*Punica granatum* L.) cultivars and their components using non-destructive X-ray computed tomography technique. *International Journal of Food Properties*, 18(2), 439-455.
- Shirzadifar, A., Bajwa, S., Mireei, S. A., Howatt, K., & Nowatzki, J. (2018). Weed species discrimination based on SIMCA analysis of plant canopy spectral data. *Biosystems Engineering*, 171, 143-154.
- Shirzadifar, A., Bajwa, S., Nowatzki, J., & Shojaeiarani, J. (2020). Development of spectral indices for identifying glyphosate-resistant weeds. *Computers and Electronics in Agriculture*, 170, 105276.
- Xiao, H., Feng, L., Song, D., Tu, K., Peng, J., & Pan, L. (2019). Grading and sorting of grape berries using visible-near infrared spectroscopy on the basis of multiple inner quality parameters. *Sensors*, 19(11), 2600.
- Xu, D., Wang, Y., Meng, Y., & Zhang, Z. (2017, December). An improved data anomaly detection method based on isolation forest. *10th International Symposium on Computational Intelligence and Design (ISCID)* (Vol. 2, pp. 287-291). IEEE.
- Yang, L., Gao, H., Meng, L., Fu, X., Du, X., Wu, D., & Huang, L. (2020). Nondestructive measurement of pectin polysaccharides using hyperspectral imaging in mulberry fruit. *Food Chemistry*, 334, 127614.