تشخیص مقدار پروتئین و رطوبت دانه‌های گندم با استفاده از روش غیرمخرب فروسرخ نزدیک (NIR)

[1]بهنام فروزانی، [2*]حسین باقرپور، [3]خلیل زابلی

[1]دانش آموخته کارشناسی ارشد مهندسی مکانیک بیوسیستم، دانشگاه بوعلی سینا، همدان؛ behnamforoozani1992@gmail.com

[2*]استادیار، گروه مهندسی مکانیک بیوسیستم، دانشکده کشاورزی، دانشگاه بوعلی سینا، همدان ؛ h.bagherpour@basu.ac.ir

[3]استادیار ، گروه علوم دامی، دانشکده کشاورزی، دانشگاه بوعلی سینا، همدان؛ zaboli@basu.ac.ir

چکیده

با توجه به فرآوری محصولات متنوع از گندم، بررسی کیفیت دانه‌ی گندم بسیار حائز اهمیت است. پروتئین و رطوبت دانه‌های گندم از شاخص های اصلی کیفی محصول می باشند و نقش به سزایی در فرآوری این محصول ایفا می کنند. به‌منظور اندازه‌گیری این دو پارامتر از روش‌های مخرب و غیر مخرب مختلفی استفاده می‌شود که در مقایسه، روش‌های مخرب وقت‌گیر و هزینه‌بر بوده و روش‌های غیر مخرب نیز بیشتر بر روی آرد محصول انجام شده و بیشتر از روش عبوری می باشد. بنابراین هدف اصلی این پژوهش ارزیابی روش غیرمخرب طیف‌سنجی فروسرخ نزدیک حالت بازتابی جهت پیش‌بینی متغیر پروتئین و رطوبت دانه‌های گندم می باشد. برای این کار از نمونه‌های گندم در ناحیه ۱۶۵۰-۹۵۰ نانومتر طیف گیری شد. اندازه‌گیری مرجع برای پروتئین گندم به‌روش کجلدال صورت گرفت. پس از استفاده از پیش‌پردازش‌های ترکیبی روی طیف‌ها، مدل‌های کالیبراسیون برای ایجاد ارتباط بین داده‌های طیفی و اندازه‌گیری های مرجع به روش PLSR و PCR برای هردو پارامتر تدوین شدند. بهترین مدل با استفاده از روش PLSR و پیش‌پردازش ترکیبی SGD+SNV و MA+D2+SNV به ترتیب برای پروتئین و رطوبت به دست آمد. مقادیر ضریب رگرسیون (R^2)، RMSE و SDR برای داده‌های اعتبار سنجی به ترتیب برابر با ۰/۸۴، ۰/۸۳۵ و ۲/۵۴ برای تخمین پروتئین و ۰/۹۶، ۰/۹۹۴ و ۵/۳۴ برای تخمین رطوبت به دست آمد. نتایج نشان داد که روش طیف‌سنجی فروسرخ نزدیک روشی کارآمد و دارای پتانسیل قوی برای تشخیص سریع پارامترهای پروتئین و رطوبت گندم می‌باشد.

واژه‌های کلیدی: پروتئین، طیف‌سنجی فروسرخ نزدیک، گندم، مدلسازی.

یازدهمین کنگره ملی مهندسی مکانیک بیوسیستم (ماشین‌های کشاورزی) و مکانیزاسیون ایران

Non-destructive determination of protein and moisture content in wheat grains using near infrared spectroscopy

1Behnam Foroozani, 2Hossein Bagherpour , 3Khalil Zaboli [2]

[1]MSc, graduated student, Department of Biosystems Engineering, Bu-Ali Sina University, Hamedan, behnamforoozani1992@gmail.com
[2]Assistant Professor, Department of Biosystems Engineering, Bu-Ali Sina University, Hamedan, h.bagherpour@basu.ac.ir
[2]Assistant Professor, Bu-Ali Sina University- Department of Animal Science, Hamedan, zaboli@basu.ac.ir

Abstract:
Considering the Production of various varieties of wheat, the evaluation of wheat grain quality is very important. Protein as an important ingredient in wheat plays an important role in the production of wheat's products. In order to measure the qualitative and quantitative characteristics of agricultural products several destructive and non-destructive methods are used. The destructive methods are more time-consuming and High-Cost and the non-destructive methods mostly have been used to determine the flour properties of wheat. Therefore, the purpose of this study is to evaluate the use of infrared spectroscopy in reflectance mode to predict protein and moisture content of wheat grain. The spectrum of seed was collected over a wavelength range of 950-1650 nm and the Kjeldahl method was used to measure the protein content of wheat grains. After preprocessing of spectra, the data were modeled using partial least squares (PLS) and Principal component regression (PCR). The best models were obtained using the PLSR method and its preprocessing SGD+SNV and MA+D2+SNV for protein and moisture content, respectively. The correlation coefficient (R^2), root mean square error of prediction (RMSEP) and Standard Deviation Ratio(SDR) were obtained 0.84, 0.835 and 2.54 for protein content, whereas 0.96, 0.994 and 5.34 for moisture content, respectively. The results showed that the infrared spectroscopy method is an efficient method and has a strong potential for rapid detection of protein and moisture content of wheat grains.
Keywords: wheat, Protein, spectroscopy Near infrared.

مقدمه

غلات و حبوبات نقش ویژه‌ای در تأمین غذای مردم، چه در کشورهای پیشرفته و چه در کشورهای در حال توسعه ایفا می‌کنند. همچنین به دلیل سرشار بودن از انواع پروتئین‌ها و مواد معدنی یک منبع غذایی مهم به شمار می‌روند. گندم[†] در دنیا بالاترین سطح زیر کشت و تولید را به خود اختصاص داده است، که به‌طور متوسط ۲۲ درصد کل پروتئین در رژیم غذایی انسان را فراهم می‌کند (FAO, 2013). ترکیبات مهم دانه گندم معمولاً شامل: ۷ تا ۱۸ درصد پروتئین (بستگی به نوع گندم دارد)، ۶۰ تا ۷۰ درصد نشاسته، ۲ تا ۲/۵ درصد سلولز (فیبر خام)، ۱/۵ تا ۲ درصد چربی و مابقی مرکب از رطوبت و مواد کانی است.

وجود کمیت و کیفیت مختلف ترکیبات شیمیایی مثل نشاسته، پروتئین و غیره در گندم باعث می‌گردد تا محصولات مختلفی بتوان از انواع گندم تهیه کرد. پروتئین یکی از اجزای مهم دانه گندم بوده و به دلیل تأثیرگذاری آن بر کیفیت نان از اهمیت ویژه‌ای برخوردار است. پروتئین گندم حاوی اسیدآمینه‌های ضروری بدن است و بخش مهم و قابل توجهی از پروتئین مورد نیاز بدن را تأمین می‌کند. (Bahraei, 2003). از این رو بررسی کمیت و کیفیت آن در فرآیندهای تولید و یا پخت نان از اهمیت ویژه‌ای برخوردار است.

یکی از مهم‌ترین عامل‌های در تعیین زمان برداشت محصول گندم، رطوبت آن است. پایین یا بالا بودن رطوبت می‌تواند باعث افزایش خسارات به محصول در زمان برداشت و انبارداری شده و از طرفی می‌تواند تا حدودی هزینه‌های فرآیندهای پس از برداشت را زیادتر کند. به طوریکه انبار گندم با میزان رطوبت بیشتر از ۱۳ تا ۱۴ درصد، باعث تشدید واکنش‌های شیمیایی و بیوشیمیایی شده و همچنین شرایط مناسبی را برای رشد قارچ و

۲- نویسنده مسئول: حسین باقرپور، گروه مکانیک بیوسیستم، دانشکده کشاورزی، دانشگاه بوعلی سینا. Email:h.bagherpour@basu.ac.ir، تلفن: ۳۴۴۲۵۴۰۰-۰۸۱
† Wheat

حشرات در سیلوها فراهم می‌کند. از طرفی برای برای کاهش تلفات زمان برداشت، معمولاً رطوبت دانه را بین ۱۴ الی ۱۶ درصد در نظر می‌گیرند (Esmaeilzade Moghadam, 2016).

روش‌های اندازه‌گیری کیفیت محصولات کشاورزی را می‌توان به دو بخش عمده‌ی مخرب و غیرمخرب طبقه‌بندی کرد. از مزایای اندازه‌گیری‌های مخرب می‌توان به‌دقت بالا و استاندارد بودن جهانی اشاره نمود؛ اما بدیهی است که روش‌های استاندارد و دقیق، به محیط و شرایط استاندارد و دقیق نیازمندند که این خود معایبی چون مشکل و زمان‌بر شدن آزمایش‌ها، هزینه‌ی زیاد، تلفات اقتصادی ناشی از تخریب نمونه را به همراه دارد (Bagherpour, 2014).

در سال‌های اخیر روش‌هایی همچون ماشین بینایی* فراصوتی، نوری ازجمله طیف‌سنجی فروسرخ نزدیک (NIR)، رزونانس مغناطیسی هسته‌ای (NMR)، رامان و تصویربرداری فرا طیفی و ... به عنوان روش‌های غیرمخرب در حال گسترش و توسعه می‌باشند. طیف‌سنجی NIR در مقایسه با برخی روش‌های غیرمخرب دیگر که تنها قادر به نمایش ساختار درونی مواد هستند، توانایی نمایش جزئیات ترکیبی و غذایی مواد و اندازه‌گیری ترکیبات مواد بیولوژیکی را نیز دارد (Jha and Matsuoka, 2000). همچنین می‌توان میزان ترکیباتی با ۰/۱ غلظت را نیز به کمک این روش به‌صورت غیرمخرب و سریع تعیین نمود. این روش در مقایسه با روش‌های مخرب هزینه کمتری داشته و می‌توان با تدوین یک مدل کالیبراسیون، تعداد زیادی نمونه را بدون نیاز به آماده‌سازی قبلی مورد بررسی و آزمایش قرار داد (Cen and He, 2007).

تحقیقات انجام گرفته در زمینه‌ی کاربرد طیف‌سنجی NIR را می‌توان به دو دسته تقسیم کرد. دسته‌ی اول شامل تحقیقاتی است که در آن پژوهشگر، هدف ارائه‌ی یک مدل رگرسیونی، مبتنی بر داده‌های طیفی NIR برای تخمین و یا پیشگویی میزان یک پارامتر موردنظر را داشته است که در بیشتر این تحقیقات از مدل‌های رگرسیون خطی شامل رگرسیون حداقل مربعات جزئی (PLSR) و رگرسیون مؤلفه‌های اصلی (PCR) استفاده شده است. دسته‌ی دوم از این تحقیقات شامل پژوهش‌هایی است که در آن‌ها هدف، دسته‌بندی و یا طبقه‌بندی محصولات کشاورزی با استفاده از NIR است. طیف NIR یک ترکیب می‌تواند اطلاعات مهمی در مورد ساختار مولکولی و طبیعت شیمیایی ماده ارائه کند؛ با این حال همیشه استخراج داده‌ها به طور مستقیم از طیف ممکن نیست؛ بنابراین، روش‌های آماری چند متغیره برای استخراج اطلاعات مفید از یک طیف NIR نیاز می‌باشد (Golizadeh, 2016).

در پژوهشی با کمک طیف‌سنجی NIR و در ترکیب با شبکه عصبی مصنوعی توانستند با توانایی بالایی پارامترهای کیفی گندم را شامل پروتئین و رطوبت دانه با مقدار R^2 به ترتیب ۰/۹۵۲ و ۰/۹۲ تخمین بزنند (Mutlu et al., 2011)؛ در این پژوهش که بر روی نمونه های حاصل از آرد گندم انجام شد پارامترهای کیفی مختلفی مثل مقدار پروتئین، رطوبت، مقدار جذب آب، زمان آماده سازی خمیر، درجه نرمی خمیر، پایداری و سایر ویژگی‌های فیزیکی با بهره‌گیری از طیف سنج در محدوده طیف ۲۴۹۸ nm - ۴۰۰ مورد بررسی و ارزیابی قرار گرفت. نتایج نشان داد که ترکیب روش NIR و شبکه عصبی مصنوعی قابلیت خوبی در تخمین خصوصیات فیزیکی و کیفی آرد گندم دارد. همچنین در پژوهشی دیگر با طیف‌سنجی بازتابی فروسرخ نزدیک میزان رطوبت، پروتئین، نشاسته و آمیلوز آرد گندم را تعیین شد (Hong et al., 1996). محققین در تحقیق دیگری با همین روش توانستند محتوای پروتئین غلات غذایی مختلف را در محدوده طیف ۲۵۰۰ nm - ۱۱۰۰ پیش‌بینی کنند (Kays et al., 2000). این پژوهش که بر روی چندین محصول مختلف انجام شد، با استفاده از مدل PLS ارتباط مناسبی و با ضریب تعیین ۰/۹۷ و خطای استاندارد ۰/۰۹ درصد بین مقدار نیتروژون اندازه‌گیری شده و خروجی مدل کالیبراسیون بدست آمد. در رابطه با روش‌های مدل‌سازی نیز با طیف‌سنجی عبوری NIR به آنالیز غلظت پروتئین گندم با مدل‌سازی از طریق PLS پرداخته شد که نتایج آن بسیار مطلوب ($0/95 < R^2 < 0/85$ و $0/9 > SEP > 0/4$) بود (Delwiche, 1995). در پژوهش دیگر که بر روی گندم انجام شد توسط طیف‌سنجی فروسرخ نزدیک، خواص پروتئین گندم آسیاب شده و رطوبت دانه گندم مورد ارزیابی قرار گرفت (Osborne and Fearn, 1983). هدف اصلی این پژوهش مقایسه ۱۱ مدل از دستگاه طیف سنج فروسرخ نزدیک در تشخیص مقدار پروتئین و رطوبت گندم بود. نتایج نشان داد که برای مدل‌های مختلف اگر دستگاه ها بخوبی تنظیم شوند، می‌توانند مقدار پروتئین را با دقت ۰/۵٪± نسبت به مقدار واقعی نشان دهند. از طرفی رطوبت بدست آمده با مدل‌های مختلف در مقایسه با روش خشک کردن با آون با میانگین دقت ۰/۳۹٪ گزارش گردید.

---

*Machine vision
Near Infrared
Nuclear Magnetic Resonance
Raman spectroscopy
**Hyperspectral Imaging
Partial Least Squares Regression
Principle Component Regression

با توجه به اهمیت اندازه‌گیری میزان پروتئین دانه‌های گندم در پروسه‌های مختلف تولید و تهیه آرد و همچنین کاهش هزینه آزمایش‌ها و از طرفی کاهش خسارات حاصله از رطوبت گندم، لازم است که این دو متغیر در مراحل مختلف از برداشت محصول تا رسیدن به یک محصول نهایی مورد سنجش و ارزیابی قرار گیرند. هر چند تحقیقاتی در زمینه سنجش پروتئین گندم به روش طیف‌سنجی انجام شده ولی اکثر پژوهش‌ها انجام شده بر روی آرد گندم بوده و از طرفی گزارشی در مورد اندازه‌گیری مستقیم پروتئین دانه‌های گندم در ایران یافت نگردید. سنجش پروتئین گندم علاوه بر کنترل فرآیندهای تولید، اگر در زمان برداشت و مستقیما در روی کمباین انجام گیرد، می‌تواند گام بزرگی برای کشاورزان در جهت مدیریت مزرعه باشد. همچنین تشخیص مقدار رطوبت گندم بر روی کمباین می تواند کمک شایانی در تلفات زمان برداشت و یا مدیریت بعدی فرآوری و انبار داری داشته باشد بنابراین تشخیص بلادرنگ رطوبت دانه ها از اهمیت بسیاری برخوردار است. با اینکه دانش کافی برای تعیین سریع مؤلفه‌های اندازه، وزن و رطوبت وجود دارد، اما هنوز اطلاعات کافی برای تشخیص سریع و غیرمخرب کیفیت محصول خام، به‌ویژه محصولات دانه‌ای، وجود ندارد. لذا در این پژوهش پتانسیل سنجی تشخیص مقدار پروتئین و رطوبت با بهره‌گیری از تکنیک طیف‌سنجی مادون قرمز نزدیک مورد بررسی قرار گرفت.

مواد و روش‌ها

تهیه نمونه

در این تحقیق برای بررسی دو متغیر رطوبت و پروتئین نمونه‌های مورد آزمایش از مزارع روستای کرفس در استان همدان با عرض جغرافیایی "۰۴ '۲۱ °۳۵ و طول "۲۹ '۱۹ °۴۹ تهیه شدند.برای این کار تعداد ۱۰۸ (نمونه از مزارع گندم در مرحله قبل از برداشت نمونه‌برداری شد. پس از آن دانه‌های گندم از خوشه جدا شدند و نمونه‌ها از مواد خارجی و کاه و کلش پاک‌سازی شدند. برای نمونه‌هایی که فقط میزان رطوبتشان محاسبه می‌شود به اندازه یک کیلوگرم گندم از یک رقم تهیه شد. اندازه‌گیری رطوبت با استفاده از استاندارد ملی (۲۷۰۵) صورت گرفت. برای ایجاد نمونه‌هایی با رطوبت مختلف، نمونه گندم به تعداد ۶ بخش مساوی تقسیم شد و با اضافه کردن مقدار مشخصی از آب، ۶ سطح رطوبتی شامل رطوبت‌های ۴، ۸، ۱۱، ۱۴، ۱۸ و ۲۲ درصد به‌دست آمد. برای این کار ابتدا رطوبت اولیه کل نمونه محاسبه گردید که برای نمونه های مخلف بین ۱۰ الی ۱۱٪ بدست آمد. برای ایجاد سطوح رطوبتی بالاتر، به مقدار مشخص آب به هر نمونه اضافه شد تا به میزان رطوبت مورد نظر برسد و رطوبت‌های پایین تر از شاهد نیز با گذاشتن در آون بدست آمد. برای هر سطح رطوبتی نیز چهار تکرار در نظر گرفته شد.

طیف‌سنجی از نمونه‌ها

طیف‌گیری از نمونه‌ها با استفاده از دستگاه اسپکترومتر مدل Perten DA7200 ساخت آمریکا و به روش بازتابی انجام شد. دامنه کاری دستگاه ۱۶۵۰-۹۵۰ نانومتر و دارای قدرت تفکیک ۵ نانومتر است. به منظور طیف گیری از نمونه‌ها نیاز به هیچ‌گونه آماده‌سازی بر روی نمونه‌ها نبوده و تنها بعد از جداسازی مواد اضافی از آنها، نمونه‌ها برای طیف‌سنجی آماده گردید. برای بررسی مقدار پروتئین با استفاده از روش طیف سنجی، از هر مکان مزرعه دو نمونه انتخاب و پس از مرتب‌سازی، ۱۰۸ نمونه موجود به ترتیب برای طیف‌سنجی شماره‌بندی و در داخل کیسه های پلاستیکی قرار داده شدند. برای هر نمونه دو اسکن از آن گرفته شد که در نهایت میانگین طیف‌های اسکن شده توسط نرم‌افزار دستگاه به عنوان طیف آن نمونه در نظر گرفته شد. پس از انجام مرحله طیف‌گیری، نمونه‌ها بلافاصله به آزمایشگاه تغذیه دام دانشگاه بوعلی سینا منتقل و نسبت به اندازه‌گیری ویژگی‌های کیفی نمونه‌ها اقدام شد. اندازه‌گیری پروتئین گندم به روش کجلدال و با استفاده از استاندارد ملی به شماره ۱۹۰۵۲ انجام شد.

پیش‌پردازش داده‌ها

داده‌های طیفی به دست آمده از دستگاه طیف‌سنج علاوه بر دارا بودن اطلاعات نمونه، شامل اطلاعات ناخواسته* پس‌زمینه‌ها، خواص فیزیکی و نویزها نیز هستند. با توجه به موارد ذکر شده، برای تدوین مدل‌های رگرسیونی قابل اعتماد نیاز به پیش‌پردازش داده‌های طیفی پیش از تدوین مدل‌های رگرسیونی است (Nicolai et al., 2007). کاربرد روش‌های پیش‌پردازش برای کاهش تغییرات طیف در اثر تغییرات فیزیکی ناشناخته بسیار معمول است و در این تحقیق به‌منظور آماده‌سازی و پیش‌پردازش طیف‌ها، از ترکیب برخی روش‌های مختلف هموارسازی و کاهش نویز، مشتق‌گیری†، نرمال‌سازی‡ و افزایش قدرت تفکیک طیفی استفاده شد.

---

* Outliers
† Differentiation
‡ Normalizing

مدل‌سازی خطی

هدف از به‌کارگیری روش‌های مدل‌سازی رگرسیون چندمتغیره، ایجاد یک ارتباط بین ویژگی‌های کیفی اندازه‌گیری شده و داده‌های طیفی است. از این رو، در این پژوهش برای پیش‌گویی پروتئین و رطوبت از دو روش مدل‌سازی رگرسیون چندمتغیره خطی بر پایه کاهش تعداد متغیرهای طیفی شامل رگرسیون مؤلفه‌های اصلی (PCR) و حداقل مربعات جزئی (PLSR) استفاده شد. در این راستا، ابتدا برای کل طیف‌های گرفته شده از نمونه‌ها، نمونه‌های پرت از نظر داده‌های طیفی که ممکن است به دلیل مشکل‌های فنی مربوط به تجهیزات اسپکتروسکوپی، جمع‌آوری داده، نمونه‌گیری نادرست و غیره به وجود آمده باشند، با استفاده از روش تجزیه مؤلفه‌های اصلی (PCA) مشخص و حذف شدند. سپس برای مدل‌سازی، نمونه‌های باقی‌مانده به‌صورت تصادفی به دو دسته آموزش (۷۰٪ نمونه‌ها) و آزمون (۳۰٪ نمونه‌های باقی‌مانده) به ترتیب برای تدوین مدل واسنجی و پیشگویی تقسیم شدند. در این تحقیق روش اعتبارسنجی متقاطع کامل*(حذف تک نمونه) برای ارزیابی مدل‌های واسنجی و پیش‌گویی به کار رفت. مدل‌های تدوین شده توسط ضریب رگرسیون†(R^2) ریشه میانگین مربعات خطا ‡(RMSE) و نسبت انحراف معیار به این خطا §(SDR) ارزیابی شدند. این سه ضریب به ترتیب بر اساس روابط ۱، ۲ و ۳ اعتبارسنجی می‌کنند.

$$R^2=\frac{\sum_{i=1}^{N}(\hat{y}_i-y_i)^2}{\sum_{i=1}^{N}(\hat{y}_i-\bar{y})^2}$$ رابطه (۱)

$$RMSE=\sqrt{\frac{\sum_{i=1}^{N}(\hat{y}_i-y_i)^2}{N}}$$ رابطه (۲)

$$SDR=\frac{SD}{RMSEP}$$ رابطه (۳)

که در این روابط RMSE ریشه میانگین مربعات خطای پیش‌بینی، y_i مقدار اندازه‌گیری شده صفت مورد نظر، $\hat{y}i$ مقدار پیش‌بینی صفت مورد نظر برای نمونه i هنگامی که مدل بدون نمونه i ساخته شده است، N تعداد نمونه آزمون، y_m میانگین مقادیر اندازه‌گیری شده صفت، SDR نسبت انحراف معیار، SD انحراف معیار صفت مورد نظر هستند. یک مقدار SDR بین ۱/۵ و ۲ به این معنی است که مدل می‌تواند مقدارهای پایین متغیر وابسته را از مقدارهای بالای آن تمیز دهد. مقدار بین ۲ تا ۲/۵ برای شاخص SDR نشان می‌دهد که پیش‌گویی‌های کمّی با دقت کم امکان‌پذیر است. مقدار SDR بین ۲/۵ تا ۳ و بیشتر از ۳ به ترتیب بیانگر دقت‌های پیش‌گویی خوب و عالی مدل است.

انتخاب مدل بهینه از بین مدل‌های تدوین شده برای پیش‌گویی پارامتر پروتئین گندم بر پایه داشتن RMSEP کمتر، R_p^2 و SDR بیشتر انجام شد.

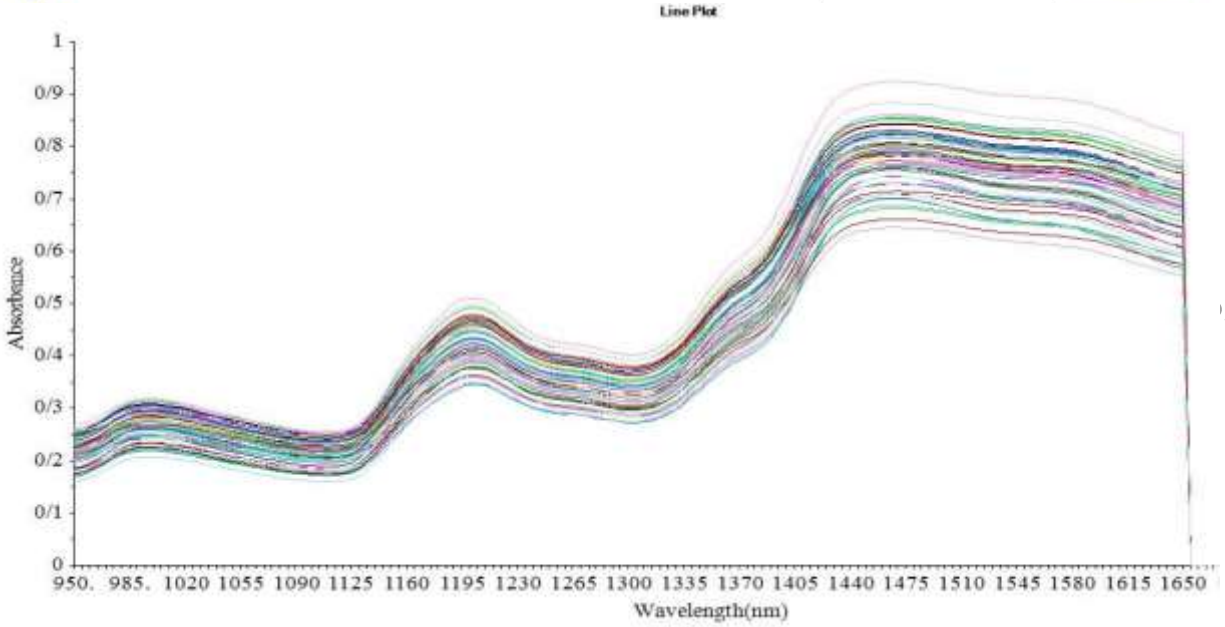
تمام تحلیل‌های آماری مربوط به پیش‌پردازش طیف‌ها و تدوین مدل با استفاده از نرم‌افزار The Unscrambler X 10.4 انجام شد.

بحث و نتایج

شکل (۱) میانگین طیف‌های جذبی خام بخشی از نمونه‌ها را نشان می‌دهد. طیف‌های به دست آمده برای همه نمونه‌ها ساختار مشابهی دارند و فقط مقدار آن‌ها متفاوت است. به دلیل وجود پیک‌های پهن و پیک‌های جذبی کوچک‌تر که هر کدام بیانگر قسمتی از نمونه هستند کسب اطلاعات دقیق به صورت مستقیم از طیف خام دشوار است. همچنین بعضی پیک‌ها در طیف خام به دلیل وجود نویز قابل رؤیت نیستند که پس از اعمال پیش‌پردازش‌ها قابل رؤیت می‌شوند؛ بنابراین برای دستیابی به جزئیات جذبی بیشتر که در این طیف‌ها به طور نامحسوس وجود دارند مشتق‌های اول و دوم و پیش‌پردازش‌های دیگر طیف‌ها بررسی شدند.

یازدهمین کنگره ملی مهندسی مکانیک بیوسیستم (ماشین‌های کشاورزی) و مکانیزاسیون ایران

* Cross Validation
† Correlation Coefficient
‡ Root Mean Squares Error
§ Standard Deviation Ratio



شکل ۱- میانگین طیف جذبی NIR خام به دست آمده از نمونه‌های گندم

Figure 1. The mean absorbance of raw NIR spectra for wheat samples

همان‌طور که پیش از این بیان شد برای به دست آوردن مدل‌های کالیبراسیون قابل اعتماد با بالاترین دقت پیش‌گویی شاخص‌های پروتئین و رطوبت گندم، روش‌های پیش‌پردازش به‌صورت ترکیبی بر روی طیف‌ها اعمال شدند و در نهایت مدلی که بیشترین مقدار R^2 و SDR و پایین‌ترین مقدار RMSE را نتیجه داد به عنوان بهترین مدل گزارش شد. برخی از نتایج مدل‌های تدوین شده برای پروتئین و رطوبت با روش‌های PCR و PLSR بر پایه ترکیب پیش‌پردازش‌های مختلف به ترتیب در جداول (۱) و (۲) آورده شده است. با توجه به جداول طیف‌سنجی فروسرخ نزدیک رگرسیون خوبی بین پروتئین و طیف و همچنین بین رطوبت و طیف را نتیجه داد.

در اعتبار سنجی مدل‌های مربوط به پروتئین گندم با مقایسه خروجی مدل با مقادیر مرجع، بهترین مدل با استفاده از روش PLSR و پیش‌پردازش ترکیبی SGD+SNV به دست آمد. در این مدل ضرایب R^2، SDR و RMSEP برای مدل پیش‌بینی به ترتیب برابر با ۰/۸۴، ۲/۵۴ و ۰/۸۳۵ به دست آمد. اعتبارسنجی مدل‌ها نشان داد که مدل‌هایی که با روش PLSR ایجاد شده‌اند نسبت به روش PCR، به دلیل بالاتر بودن مقادیر R^2 و SDR و همچنین پایین‌تر بودن مقدار RMSEP قابل اعتمادتر هستند. با توجه به مقادیر شاخص SDR، بیشتر مدل‌های رگرسیون چند متغیره تدوین شده به روش PLSR برای نمونه‌های پروتئینی گندم مقدار SDR آن‌ها از ۲ بالاتر بود که این بیان می‌کند امکان تشخیص و پیش‌بینی پروتئین نمونه‌ها با دقت کم وجود دارد. نتایج جدول (۱) نشان می‌دهد که مدل‌هایی که به روش PCR ایجاد شدند به دلیل اینکه مقدار شاخص SDR برای تعداد کمی از مدل‌ها بالاتر از ۲ بود مدل‌هایی با مطلوبیت متوسط می‌باشند و در تخمین مقدار پروتئین نسبت به مدل PLSR ضعیف‌تر عمل می‌کنند.

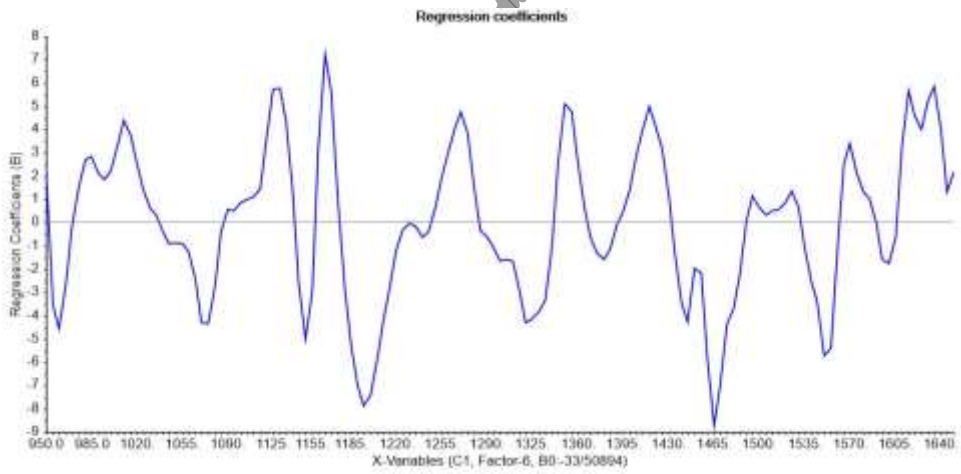
یازدهمین کنگره ملی مهندسی مکانیک بیوسیستم و مکانیزاسیون ایران

جدول ۱- مدل‌های حاصل از طیف‌سنجی NIR جهت تعیین پروتئین گندم

Table 1. Models of NIR spectrometry to determine wheat protein

SDR	RMSEP	R_P^2	RMSEC	R_C^2	PCs	مدل Model	پیش‌پردازش Preprocessing
1.68	1.267	0.727	0.908	0.792	8	PCR	SG+SNV
2.02	0.988	0.768	0.786	0.851	8		MA+D1+SNV
1.59	1.382	0.719	0.936	0.786	8		SG+D2+SNV
2.1	0.978	0.774	0.736	0.861	6		MA+D1+MSC
1.62	1.102	0.758	0.781	0.845	6	PLSR	SG
1.91	1.013	0.774	0.795	0.840	5		SG+D1
2.38	0.874	0.833	0.637	0.897	8		MA+D1+SNV
2.44	0.847	0.839	0.600	0.908	8		D1+SNV
2.13	0.929	0.804	0.593	0.911	9		NOR+BL+D1
2.54	**0.835**	**0.840**	**0.496**	**0.937**	**8**		**SGD+SNV**

شکل (۲) ضریب رگرسیونی*طیف بهترین مدل برای پروتئین گندم را نشان می‌دهد. در مجموع چندین پیک در طیف حاصل از مدل‌سازی مشاهده شد. نسبت دادن برخی از این طول‌موج‌ها به باندهای جذبی مربوط به پروتئین منطقی است. همان گونه که در شکل مشاهده می‌شود محدوده طیفی ۱۴۶۵ نانومتر بالاترین ضریب رگرسیون را با مقادیر داده‌های پروتئین مرجع دارد که می‌تواند مربوط به باند جذبی پروتئین (عموماً در nm ۱۴۷۱-۱۵۳۰) باشد که مطابق با اورتون اول پیوندهای کششی O-H و N-H است. این محدوده با نتایج پژوهش دلوویچ (۱۹۹۸) هم‌خوانی دارد.

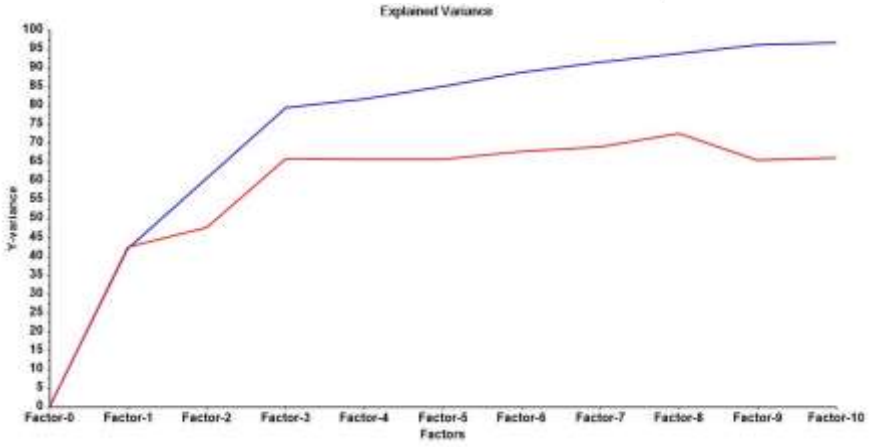


شکل ۲- ضریب رگرسیون مدل PLSR در طول‌موج‌های مختلف برای تخمین مقدار پروتئین

Figure 2. Regression coefficient of PLSR model at different wavelength for prediction of protein

اثر تعداد مؤلفه‌های اصلی بر واریانس تبیین شده برای مدل پروتئین گندم در شکل (۳) نشان داده شده است. با افزایش تعداد مؤلفه‌های اصلی ابتدا مقدار واریانس تبیین شده به میزان قابل توجهی زیاد گردید و پس از آن با بیشتر شدن تعداد فاکتور تغییرات قابل توجهی در مقدار واریانس مشاهده نمی‌شود. همان‌گونه که در شکل (۳) قابل مشاهده است بعد از فاکتور ۸ مقدار واریانس تبیین شده مدل اعتبار سنجی کاهش داشته و همین فاکتور به عنوان بهترین مقدار بهینه در نظر گرفته شد. بدین ترتیب برای تخمین میزان پروتئین از ۸ مؤلفه اصلی در این مدل استفاده شد که همین تعداد فاکتور می‌تواند نزدیک به ۹۰ درصد از واریانس‌های متغیر وابسته را توصیف کند.

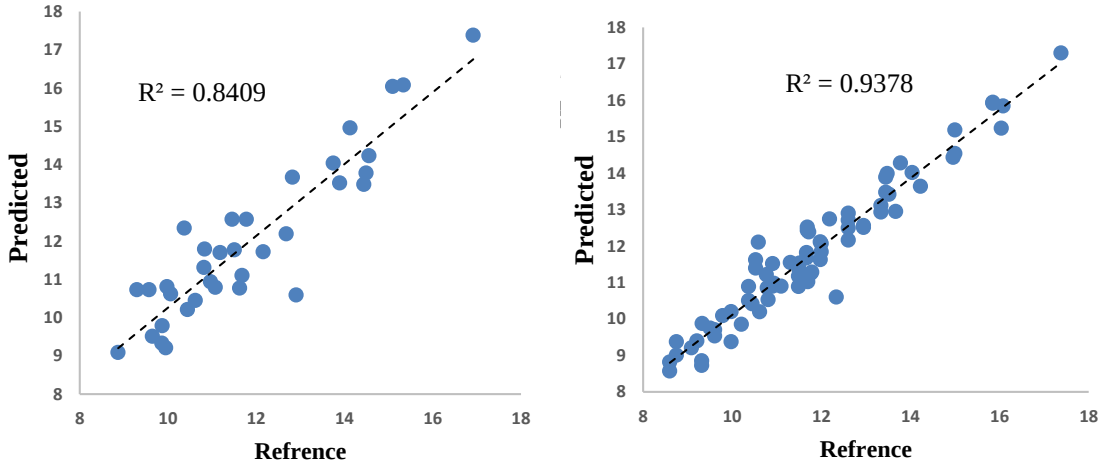
*Regression Coefficients



شکل ۳- واریانس توصیف شده داده‌ها در فاکتورهای مختلف برای پروتئین گندم

Figure 3. Explained variance graph of protein for each PC

در شکل (۴) ارتباط بین مقادیر واسنجی و پیش‌بینی پروتئین گندم برای بهترین مدل انتخاب شده نشان داده شده است. همان‌طور که مشاهده می‌شود مدل مربوطه توانست مقدار پروتئین را با توانایی بالایی تخمین بزند. نتایج این پژوهش با نتایج Mutlu et al., 2011 که با استفاده از مدل شبکه عصبی مصنوعی مقدار پروتئین و رطوبت آرد گندم را به ترتیب برابر با ضریب تعیین ۰/۹۵ و ۰/۹۳ گزارش کرده بودند مطابقت دارد. از طرفی در مقایسه با روش های تصویر برداری فراطیفی که در آن مقدار پروتئین تک دانه‌های گندم را در محدوده طول موج ۹۸۰-۲۵۰۰ nm تصویر برداری کرده بودند ()، تخمین مناسب‌تری بدست آمد. در تصویر برداری حرارتی مقدار R^2 و ریشه میانگین مربعات خطا (RMSE) به ترتیب برابر با ۰/۸۵ و ۰/۸۶ گزارش شده بود.



شکل ۴- ارتباط بین مقادیر اندازه‌گیری شده مرجع و پیش‌بینی شده پروتئین دانه‌های گندم A) دادهای واسنجی B) داده‌های پیش‌بینی

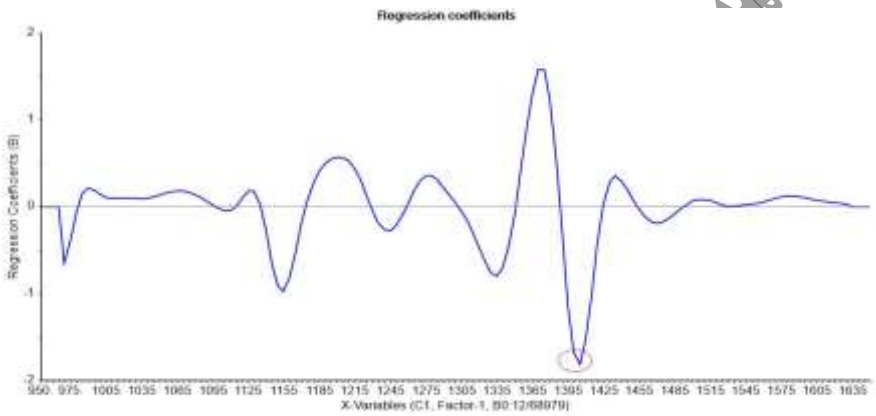
Figure 4. Relationship between predicted and measured value of wheat seed protein A) Calibration data B) prediction data

همان‌گونه که در جدول (۲) نشان داده شده است، با اعتبارسنجی مدل‌های مربوط به رطوبت گندم بهترین مدل به دست آمده با روش PLSR و پیش‌پردازش ترکیبی MA+D2+SNV حاصل شد. ضرایب R^2 و RMSE نیز به ترتیب ۰/۹۶ و ۰/۹۹۴ برای پیش‌بینی رطوبت به دست آمد. ضریب SDR برای مدل انتخابی ۵/۳۴ و همچنین برای تمام مدل‌ها از ۳ بالاتر بود که نشان می‌دهد مدل مربوطه از دقت عالی در پیش‌گویی رطوبت برخوردار است.

جدول ۲- مدل‌های حاصل از طیف‌سنجی NIR جهت تعیین رطوبت گندم

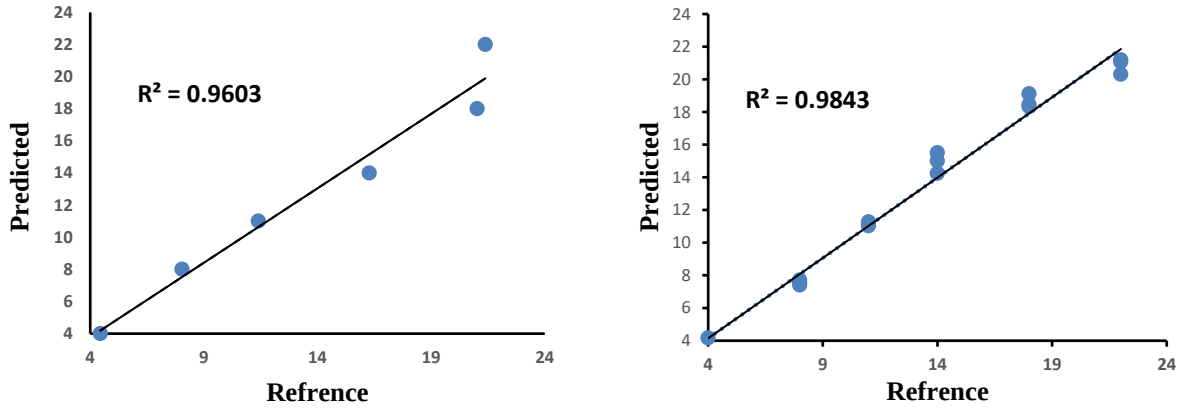
Table 2. Models of NIR spectrometry to determine of the wheat moisture

RMSEP	R_p^2	RMSEC	R_c^2	Factor	مدل Model	پیش‌پردازش Preprocessing
0.994	**0.960**	**1.024**	**0.984**	**1**	**PLSR**	**MA+D2+SNV**
1.12	0.960	1.14	0.981	3		MF+D1+SNV
1.54	0.954	1.43	0.977	2		MA+D2
1.284	0.955	1.2	0.972	1	PCR	MA+D1+SNV
1.33	0.940	1.274	0.981	1		MA+D2+SNV
1.333	0.952	1.286	0.979	2		MF+D2+MSC



شکل ۵- ضریب رگرسیونی مدل PLSR در طول‌موج‌های مختلف برای تخمین رطوبت

Figure 5. Regression coefficient of PLSR model at different wavelength for prediction of moisture



A **B**

شکل ۶- ارتباط بین مقادیر اندازه‌گیری شده مرجع و پیش‌بینی شده رطوبت دانه‌های گندم A) داده‌های واسنجی B) داده‌های پیش‌بینی

Figure 6. Relationship between predicted and measured value of wheat seed moisture A) Calibration data B) prediction data

نمودار ضریب رگرسیون (شکل۵) نشان می‌دهد که بهترین طول‌موج برای تخمین میزان رطوبت در محدوده طول موج ۱۴۰۵ نانومتر که مربوط به اولین اورتون پیوند O-H می‌باشد قرار داشته و دارای بیشترین ضریب رگرسیونی نسبت به بقیه طول‌موج‌ها است. این نتایج با گزارش مربوط به جدول استاندارد مربوط به پیوندهای اتمی کاملا هم خوانی دارد (Anonymous, 2005). نتایج حاصل از مدل رگرسیونی مربوط به نمودارهای کالیبراسیون و تخمین رطوبیت در شکل ۶ نشان داده شد است.

نتیجه‌گیری

هرچند روش طیف‌سنجی بیشتر برای تشخیص پروتئین گندم فرآوری شده مثل آرد بکار رفته شده است ولی نتایج تحلیل روش‌های رگرسیونی به منظور تشخیص غیرمخرب شاخص میزان پروتئین گندم حاکی از آن است که از طیف‌سنجی فروسرخ نزدیک در محدوده ۱۶۵۰-۹۵۰ نانومتر می‌توان برای تشخیص و پیش‌بینی دانه‌های گندم نیز استفاده نمود. اعتبارسنجی مدل‌های رگرسیون چند متغیره خطی بر پایه ترکیب روش‌های مختلف پیش‌پردازش طیف‌ها نشان داد که بهترین مدل پیشگو برای پروتئین، مدل‌های PLSR می‌باشد و ترکیب مختلف پیش‌پردازش‌ها اثر مستقیم بر نتایج پیش‌گویی مدل‌ها داشتند. در نهایت با انتخاب پیش‌پردازش ترکیبی SGD+SNV دقت مدل در تخمین پروتئین به بالاترین میزان ضریب رگرسیون (۰/۸۴) و SDR (۲/۵۴) و میزان پایین RMSE (۰/۸۳۵) رسید. روش طیف‌سنجی NIR برای تشخیص میزان رطوبت دانه گندم نیز موفق عمل کرد؛ به گونه‌ای که با اعتبارسنجی مدل‌های رگرسیون چند متغیره خطی و ترکیب پیش‌پردازش‌ها، مدل‌های تدوین شده توانستند شش سطح رطوبتی را با ضریب رگرسیون بالایی (۰/۹۶) تشخیص دهند.

Reference

Anonymous (2005). Near-IR Absorption Bands. Available from: www.asdi.com.
Anonymous (2018). FAO statistical databases. Food and Agricultural Organization of United Nation. Available from: www.fao.org, Accessed on 2018/01/30.
Bagherpour, H., S. Minaei & M. Abdollahian Noghabi. 2014. Non-destructive determination of sugar content in root beet by near infrared spectroscopy (NIRS). Journal of Food Science & Technology (2008-8787) 12.
Bahraei, S. 2003. Bread wheat quality evaluation based on the high molecular weight glutenin subunits.
Caporaso, N., Whitworth, M. B., & Fisk, I. D. (2018). Protein content prediction in single wheat kernels using hyperspectral imaging. *Food chemistry*, *240*, 32-42.
Cen, H. & Y. He. 2007. Theory and application of near infrared reflectance spectroscopy in determination of food quality. Trends in Food Science & Technology 18: 72-83.
Delwiche, S. 1995. Main content area Single wheat kernel analysis by near-infrared transmittance: protein content. Cereal chemistry 72: 11-16.
Esmaeilzade Moghadam, M. (2016). Wheat guide (Planting to harvest). Agricultural Education Publishing. (In Farsi).
Golizadeh, M. (2016). Using of Near Infrared Spectroscopy (NIR) for Determination of Quince Fruit Maturity.
Hong, J.-H., K. Ikeda, I. Kreft & K. Yasumoto. (1996). Near-infrared diffuse reflectance spectroscopic analysis of the amounts of moisture, protein, starch, amylose, and tannin in buckwheat flours. Journal of nutritional science and vitaminology 42: 359-366.
Jha, S. N. & T. Matsuoka. (2000). Non-destructive techniques for quality evaluation of intact fruits and vegetables. Food Science and Technology Research 6: 248-251.
Kays, S. E., F. E., Barton & W. R., Windham. (2000). Predicting protein content by near infrared reflectance spectroscopy in diverse cereal food products. Journal of Near Infrared Spectroscopy 8: 35-43.
Mutlu, A.C., Boyaci, I.H., Genis, H.E., Ozturk, R., Basaran-Akgul, N., Sanal, T. & Evlice, A.K. (2011). Prediction of wheat quality parameters using near-infrared spectroscopy and artificial neural networks. European food research and technology 233(2), 267-274.
Osborne, B. G. & T. Fearn. 1983. Collaborative evaluation of near infrared reflectance analysis for thedetermination of protein, moisture and hardness in wheat. Journal of the Science of Food and Agriculture34:1011-1017.

یازدهمین کنگره ملی مهندسی مکانیک بیوسیستم (ماشین‌های کشاورزی) و مکانیزاسیون ایران