

تخمین میزان کلروفیل در درختان پسته با استفاده از داده های فراطیفی

داود پناهی بروجنی^{۱*}، علی اسماعیلی^۲، روشنگر درویش زاده^۳، فرزین ناصری^۴

^۱کارشناس ارشد سنجش از دور - دانشگاه تحصیلات تکمیلی صنعتی و فناوری پیشرفته کرمان
panahi.davoud@gmail.com

^۲استادیار گروه سنجش از دور - دانشگاه تحصیلات تکمیلی صنعتی و فناوری پیشرفته کرمان
aliesmaeily@kgut.ac.ir

^۳استادیار گروه کارتوگرافی - دانشکده جغرافیا - دانشگاه تهران
r_darvish@sbu.ac.ir

^۴استادیار گروه اکولوژی و گروه سیستم های اطلاعات مکانی - دانشگاه تحصیلات تکمیلی صنعتی و فناوری پیشرفته کرمان
naseri@kgut.ac.ir

(تاریخ دریافت بهمن ۱۳۹۱، تاریخ تصویب مرداد ۱۳۹۳)

چکیده

تخمین دقیق پارامترهای بیوفیزیکی و بیوشیمیایی گیاهان در زمینه های مختلفی مانند کشاورزی، اکولوژی و هواشناسی اهمیت بسزایی دارد. در میان محصولات کشاورزی ایران، اهمیت استراتژیکی-اقتصادی پسته ضرورت توجه بیشتر به برنامه ریزی جهت افزایش تولید این محصول را آشکار می سازد. هدف اصلی این مطالعه مقایسه روش های آماری تک متغیره شاخص های باریک باند گیاهی و روش های چند متغیره SMLR و PLSR در پیش بینی غلظت کلروفیل برگ درختان پسته با استفاده از داده های انعکاسی فراطیفی می باشد. داده های طیفی از برگ درختان پسته که در مقاطع رشد و شرایط گوناگون جمع آوری گردیده بودند، در اتاق تاریک با دستگاه طیف سنج ASD Field Spectrometer III تهیه شدند. سپس میزان کلروفیل برگ ها با دستگاه کلروفیل متر SPAD قرائت شدند و همچنین آزمایشات شیمیایی جهت استخراج غلظت کلروفیل نمونه ها در آزمایشگاه شیمی انجام شد. در میان شاخص های گیاهی باریک باند جهت پیش بینی غلظت کلروفیل نمونه ها، شاخص RVI با طول موج های بهینه ۶۷۰ و ۷۳۴ نانومتر و استفاده از داده های طیفی تبدیل شده مشتق اول بیشترین دقت را داشت ($R_{cv}^2 = 0.72$ ، $RRMSE_{cv} = 0.25$). نتایج بدست آمده از روش های چند متغیره در پیش بینی غلظت کلروفیل نیز نشان داد که، مدل های PLSR و SMLR با استفاده از داده های طیفی تبدیل شده مشتق اول دقت بالاتری نسبت به روش های تک متغیره داشتند ($R_{cv}^2 = 0.79$ ، $RRMSE_{cv} = 0.21$). به طور کلی نتایج این تحقیق نشان داد، داده های انعکاسی فراطیفی پتانسیل بالایی در پایش سلامت و وضعیت رشد محصولات کشاورزی، مخصوصا برگ درختان پسته دارند و همچنین مدل های آماری چند متغیره نسبت به روش های تک متغیره دقت بالاتری در پیش بینی غلظت کلروفیل برگ درختان پسته داشتند و داده های انعکاسی تبدیل شده نسبت به داده های طیفی اولیه دارای توانایی بالاتری در پیش بینی میزان کلروفیل هستند.

واژگان کلیدی: برگ درختان پسته، کلروفیل، طیف سنجی فراطیفی، شاخص های گیاهی باریک باند، مدل PLSR، مدل SMLR

* نویسنده رابط

۱- مقدمه

در راستای کشاورزی پایدار، پایش سلامت گیاه، کشف استرس، تعیین ظرفیت فتوسنتز و مقدار محصول اموری اجتناب ناپذیرند [۴،۶]. میزان کلروفیل برگ گیاهان از مهمترین پارامترهای بیوشیمیایی گیاهان می باشند که نقش اساسی در وضعیت سلامت گیاه دارد. کلروفیل برگ مهمترین رنگدانه^۱ در فعالیت های فتوسنتزی گیاه می باشد [۴۲]. یک گیاه سالم نسبت به گیاه بیمار میزان کلروفیل بالاتری دارد. بنابراین ظرفیت کلروفیل برگ معیار مناسبی برای ارزیابی فعالیت فتوسنتزی، کشف استرس گیاه و وضعیت غذایی برگ می باشد و اندازه گیری آن در مدیریت کشاورزی دقیق از اهمیت ویژه ای برخوردار است [۴۲]. اندازه گیری پارامترهای گیاهی به روش های آزمایشگاهی علاوه بر صرف وقت و هزینه زیاد، توانایی مدل کردن این پارامترها را در سطح وسیعی ندارند. با توجه به این که مشخصه های طیفی گیاهان در طول موج های مختلف تحت تاثیر ساختار درون سلولی و برون سلولی برگ و غلظت مواد بیوشیمیایی از جمله کلروفیل، نیترژن و میزان آب موجود در گیاه است [۳۹]، می توان با استفاده از رفتارهای طیفی اندازه گیری شده توسط سنجنده های سنجش از دور میزان هر یک از این پارامترها با سرعت بیشتر و هزینه کمتر و به صورت غیر تخریبی تعیین کرد [۳۴]. سنجش از دور از سنجنده های چند طیفی تا سنجش از دور فرا طیفی توسعه پیدا کرده است. در سنجش از دور چند طیفی به علت قرار گرفتن میانگین اطلاعات چند باند در یک باند، بسیاری از اطلاعات مفید که در باندهای باریک موجود می باشند در این طول موج های پهن از بین می رود [۲۵،۳۴]. اخیرا مطالعات زیادی در رابطه با استخراج کلروفیل برگ با استفاده از داده های سنجش از دوری انجام شده است [۳،۲۵،۲۸،۴۲]. پیشرفت علم و فناوری سنجش از دور و پیدایش سنجش از دور فرا طیفی باعث پیشرفت در شاخص های گیاهی که معروف به شاخص های گیاهی باریک باند^۲ هستند شده است. اهمیت استفاده از شاخص های باریک باند برای پیش بینی پارامترهای بیوفیزیکی در مطالعات [۳،۸،۱۱،۲۹] و برای پارامترهای بیوشیمیایی

گیاهی در مطالعات [۵،۱۶،۱۷،۴۳] نشان داده شده است. این شاخص ها معمولا از اطلاعات محدود دو باند تشکیل می شوند و لذا حجم عظیمی از اطلاعات در سایر باندها نادیده گرفته می شود [۱۱].

بر خلاف روش های تک متغیره، روش چند متغیره SMLR^۳ از چندین باند طیفی در پیش بینی پارامترها استفاده می کند و در مطالعات زیادی برای پیش بینی پارامترهای بیوفیزیکی [۲،۱۱،۱۴] و بیوشیمیایی گیاهان [۱۰،۲۲،۲۶،۲۷] استفاده شده است. این رگرسیون بعضی مواقع هنگام استفاده از داده های فرایطیفی متاثر از مشکل هم خطی است [۱۴]. این مشکل معمولا زمانی ایجاد می شود که تعداد مشاهدات کمتر از تعداد متغیرهای مستقل باشند و داده های ورودی همبستگی بالایی را نشان دهند [۱۱]. روش آماری چند متغیره PLSR^۴ بر خلاف روش SMLR از تمامی اطلاعات سرتاسر طیف بهره می برند و برای آنالیز داده هایی که مشکل هم خطی دارند مناسب می باشد [۱۱]. این روش آماری نیز در پیش بینی پارامترهای گیاهی در مطالعات گذشتگان به طور مکرر استفاده شده است و نتایج قابل قبولی در تخمین پارامترهای گیاهی داشته است [۲،۸،۱۱،۲۵،۲۶،۳۴].

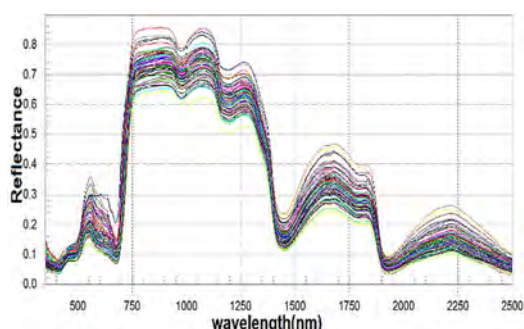
ایران عمده ترین تولید کننده پسته در جهان می باشد [۱]. اهمیت استراتژیکی-اقتصادی محصول پسته در استان کرمان و ایران ضرورت توجه بیشتر به برنامه ریزی جهت افزایش تولید این محصول را آشکار می سازد. مرور تحقیقات داخلی و خارجی نشان داد که ارزیابی وضعیت این محصول با استفاده از سنجش از دور بسیار محدود می باشد. هدف اصلی این مطالعه ارزیابی توانایی سنجش از دور در تخمین پارامتر بیوشیمیایی کلروفیل برگ درختان پسته می باشد. دیگر اهداف این مطالعه مقایسه دقت روش های آماری تک متغیره و چند متغیره SMLR و PLSR با استفاده از داده های انعکاسی اولیه و تبدیل شده با مشتق اول در پیش بینی پارامتر کلروفیل برگ درختان پسته می باشد.

۲- مواد و روش ها

^۳ Stepwise Multiple linear Regression
^۴ Partial Least Square Regression

^۱ pigment
^۲ Narrow vegetation index

از آنجایی که اعداد نشان داده شده توسط SPAD-502 بدون واحد هستند [۲۴]، در مرحله آخر در هر روز نمونه گیری به منظور کالیبره کردن دستگاه کلروفیل متر، کلروفیل نمونه برگ ها به روش آزمایشگاهی Lichtenthaler در آزمایشگاه شیمی اندازه گیری شد [۳۳]. سپس با رگرسیون توانی $y = 1.1e^{(0.04x)}$ و ضریب تعیین ۰/۹ داده های بدون واحد کلروفیل متر کالیبره شدند و از این معادله رگرسیونی برای سایر نمونه ها که در رگرسیون شرکت نداشتند استفاده شد.



شکل ۱- انعکاس های طیفی کلیه نمونه های برگ درختان پسته در محدوده طیفی ۳۵۰ تا ۲۵۰۰ نانومتر

۲-۱- پیش پردازش داده های طیفی

در مرحله اول بدلیل وجود نویز بالا در محدوده های طیفی کمتر از ۴۰۰ نانومتر و بیشتر از ۲۴۰۰ نانومتر، انعکاس های برداشت شده در این محدوده طیفی حذف شدند. سپس فیلتر Savitzky-Golay [۳۷] با اندازه پنجره ۱۷ پیکسل و چند جمله ای درجه دو، برای از بین بردن نویز های احتمالی دستگاه اندازه گیری طیفی بر روی ماتریس بازتاب های طیفی اجرا شد. سپس به منظور کاربردی کردن نتایج حاصل از این مطالعه محدوده های طیفی ۱۴۵۰-۱۹۵۰، ۱۳۵۰-۱۸۵۰ نانومتر که شامل باندهای جذب بخار آب اتمسفر (۱۴۰۰ و ۱۹۰۰ نانومتر) می باشند [۳۲]، از باندهای طیفی همه نمونه ها حذف شدند. علت حذف این باندها را می توان در این دانست که در این محدوده ها تقریباً هیچ انرژی به سنجنده های ماهواره ای نمی رسد و اغلب سنجنده ها در این نواحی تصویربرداری نمی کنند.

در مرحله آخر پیش پردازش داده ها از روش تبدیل مشتق اول استفاده شده است. روش های تبدیل داده عموماً به منظور برجسته کردن ویژگی های خاص جذب

در این مطالعه ۷۵ نمونه از برگ درختان پسته در دوره رشد گیاه و در بازه زمانی ۶ ماهه از باغ پسته مرکز تحقیقات پسته کرمان در استان کرمان با طول جغرافیایی: $30^{\circ} 03' 09''$ تا $57^{\circ} 03' 45''$ شرقی و عرض جغرافیایی: $30^{\circ} 23' 11''$ تا $30^{\circ} 23' 27''$ شمالی تهیه شدند. در هر روز نمونه گیری پس از چیدن حدود ۵۰ برگ از هر درخت، کلروفیل برگ ها در کمترین زمان ممکن با دستگاه کلروفیل متر SPAD-502 قرائت شد. سپس نمونه ها در ظرف حاوی یخ به آزمایشگاه مرکزی دانشگاه تحصیلات تکمیلی صنعتی و فناوری پیشرفته کرمان جهت انجام آزمایشات شیمیایی و طیفی انتقال داده شدند. اندازه گیری های طیفی در آزمایشگاه سنجش از دور انجام شد. در محل اندازه گیری دیوارها و سقف با مواد سیاه رنگ مات به منظور جلوگیری از بازتاب^۱ های زائد پوشانده شده بود. هر گروه برگ از نمونه ها به صورت توده ای^۲ بر روی یک سطح کاملاً مشکی پخش شدند. سپس به منظور اندازه گیری های طیفی، سنسور طیف سنج با زاویه FOV، ۲۵ درجه را در ارتفاع ۱۵ سانتی متری سطح برگ با کمک سه پایه به صورت کاملاً عمود نصب گردید. در این مطالعه دستگاه طیف سنج ASD Field Spectrometer III با محدوده نمونه برداری ۳۵۰ تا ۲۵۰۰ نانومتر، با فاصله نمونه برداری^۳ ۱/۴ نانومتر در محدوده طیفی ۳۵۰ تا ۱۰۰۰ نانومتر و ۲ نانومتر در محدوده طیفی ۱۰۰۰ تا ۲۵۰۰ نانومتر برای اندازه گیری های طیفی استفاده شد. به منظور روشن کردن هدف از یک لامپ هالوژن ۳۰۰ وات که در نزدیکی سنسور قرار داشت استفاده شد. از هر نمونه ۱۲ مرتبه طیف اندازه گیری شد و در انتها از ۱۲ طیف تهیه شده میانگین گرفته شد. البته قبل از هر اندازه گیری دستگاه طیف سنج با صفحه سفید^۴ کالیبره گردید. شکل ۱ انعکاس های طیفی تهیه شده از کلیه نمونه ها را نشان می دهد. طیف های تهیه شده با یکدیگر تفاوت هایی دارند که دلیل این تفاوت ها را می توان دلیل بر تفاوت مقادیر پارامترهای بیوفیزیکی و بیوشیمیایی برگ ها دانست.

- ۱ Reflectance
- ۲ Leaf stack
- ۳ sampling interval
- ۴ white reference panel

جدول ۱- شاخص های باریک باند استفاده شده در این مطالعه (R_{λ} نشان دهنده بازتاب در طول موج های λ_1 ، λ_2 می باشد)

شاخص های گیاهی	معادله شاخص	مرجع
شاخص نسبی ساده	$RVI = \frac{R_{\lambda_1}}{R_{\lambda_2}}$	[۳۵]
شاخص گیاهی اختلاف نرمال شده	$NDVI = \frac{R_{\lambda_1} - R_{\lambda_2}}{R_{\lambda_1} + R_{\lambda_2}}$	[۳۶]
شاخص گیاهی تبدیل شده	$TVI = \sqrt{\frac{R_{\lambda_1} - R_{\lambda_2}}{R_{\lambda_2} + R_{\lambda_2}} + 5}$	[۱۳]

۲-۳- رگرسیون خطی چندگانه مرحله ای

هنگامی که تعداد متغیرهای مستقل در یک جامعه آماری برای عمل رگرسیون زیاد است (مثلا تعداد باندهای طیفی در پیش بینی غلظت کلروفیل در این مطالعه)، رگرسیون خطی چندگانه مرحله ای (SMLR) یکی از رایج ترین روش های مورد استفاده برای توسعه یک مدل تجربی می باشد [۴۰]. در الگوریتم روش SMLR، معادله چند متغیره بدون هیچ متغیر اولیه شروع می شود، در هر مرحله موثرترین طول موج از نظر آماری به مدل اضافه می شود. در همین زمان فرایند به محاسبه پارامتر آماری حذفی برای هر یک از طول موج ها می پردازد به طوری که طول موجی که دارای مقدار کمتر از حد آستانه است از مدل حذف می شود. این فرایند تا جایی ادامه می یابد که هیچ متغیری حذف یا اضافه نشود [۱۱]. در این مطالعه روش SMLR برای تخمین غلظت کلروفیل برگ های درختان پسته با استفاده از داده های فراطیفی اولیه و همچنین طیف های تبدیل شده با مشتق اول مورد استفاده قرار گرفت. سپس باندهای طیفی انتخاب شده از الگوریتم در رگرسیون چندگانه خطی برای تخمین کلروفیل قرار گرفتند. تعداد زیاد باندهای انتخاب شده موجب همبستگی درونی مابین بازتاب ها می شود و مشکل چند همخطی ایجاد می کند [۱۵]. این مشکل با میزان ضریب تعیین (R^2) بالا در کالیبراسیون و کاهش آن در تحلیل اعتبار سنجی مشخص می شود [۱۱]. در این مطالعه جهت ارزیابی شدت همخطی باندهای انتخاب شده در مدل ها برای پیش بینی کلروفیل

گیاهان و کم کردن اثرات جانبی مانند زاویه تابش، هندسه تصویربرداری و اثرات خاک زمینه استفاده می شوند. تبدیل مشتق اول (معادله ۱) باعث بهبود ویژگی های جذبی کوچک شده اما مقدار نویز را تقویت می کند، بنابراین این الگوریتم پس از هموارسازی اعمال می شود [۱۲]. استفاده از این روش باعث بر طرف کردن شیفت و حذف کردن تغییرات جزئی در طیف و کم کردن اثر آنها در محاسبات می شود. این روش مناطقی از طیفی که شیب به سرعت عوض می شود را برجسته می کند [۴۲].

$$d^{1st} = \frac{\rho_{n+1} - \rho_n}{\lambda_{n+1} - \lambda_n} \quad (1)$$

در رابطه ۱، d^{1st} مشتق اول، n شماره باند، ρ_{n+1} مقدار انعکاس در طول موج λ_{n+1} و ρ_n مقدار انعکاس در طول λ_n می باشند.

۲-۲- شاخص های باریک باند طیفی

ساختار شاخص های طیفی به نحوی است که کنتراستی بین یک طول موج مبنا با حداقل حساسیت نسبت به رنگدانه مورد نظر و طول موج دیگری (طول موج شاخص) که حداکثر حساسیت را نسبت به آن رنگدانه دارد، ایجاد می کند.

شاخص های باریک باند مورد استفاده در این تحقیق در جدول ۱ نشان داده شده اند. شاخص های پهن باند، معمولا با استفاده از دو طول موج مشخص محاسبه می شوند اما در شاخص های فراطیفی مورد استفاده در این تحقیق با استفاده از ۱۷۹۶ باند باریک موجود، تمامی ترکیبهای دوتایی ممکن که شامل ۱۷۹۶*۱۷۹۶ ترکیب بودند، مورد محاسبه قرار می گیرند. در واقع هر دو زوج طول موج تهیه شده از نمونه ها در آزمایشگاه سنجش از دور برای تشکیل یک شاخص باریک باند جدید استفاده می شود. در نهایت با محاسبه مقدار ضریب تعیین مابین شاخص های ایجاد شده با هر یک از این ترکیبات باندی و غلظت کلروفیل بدست آمده از نمونه ها بهینه ترین دو طول موج در هر شاخص مشخص می شود.

یک فاکتور پنهان باید به مقدار ۲ درصد دقت را در پیش بینی افزایش دهد استفاده شد و تعداد فاکتور های مورد نیاز برای پیش بینی هر مدل با انعکاس های اولیه و تبدیل شده با مشتق اول با این شرایط انتخاب شدند [۸،۱۱].

۲-۵- اعتبارسنجی مدل ها

در این مطالعه برای ارزیابی دقت از دو شاخص ضریب تعیین (R^2) و ریشه میانگین مربعات خطای Cross validation شده RRMSEcv استفاده گردید. اعتبار سنجی مدل با استفاده از روش Cross Validation بدین گونه است که در این روش هر یک از نمونه ها با استفاده از سایر نمونه ها که در مدل رگرسیون استفاده شده اند تخمین زده می شود و میزان خطای آن از اختلاف مابین مقادیر پیش بینی شده و مقدار اندازه گیری شده کلروفیل نمونه ها محاسبه می شود (رابطه ۳). از مزایای این روش توانایی آن در تشخیص outlier ها می باشد و دیگر اینکه تخمینی نا اریب است [۱۱].

$$RMSE_{cv} = \sqrt{\frac{\sum_{i=1}^N (y_i - y'_i)^2}{N}} \quad (3)$$

در این روابط، y_i مقدار اندازه گیری شده مشاهده آم، y'_i مقدار پیش بینی شده مشاهده آم توسط مدل و N تعداد کل مشاهدات در مدل می باشند.

۳- نتایج و بحث

۳-۱- شاخص های باریک باند و ارتباط آنها با مقدار کلروفیل نمونه ها

سه شاخص باریک باند و همبستگی آنها با محتوای کلروفیل برگ ها با استفاده از داده های انعکاسی اولیه و تبدیل شده با مشتق اول محاسبه شدند. در شکل ۲ مقادیر ضریب تعیین R^2 بین شاخص RVI برای هر یک از دو باند طیفی موجود و کلروفیل به صورت پلات دو بعدی نشان داده شده است (سایر پلات های دو بعدی شاخص های دیگر در این مقاله نشان داده نشده اند). در این شکل نواحی طیفی دارای همبستگی بالا با مقادیر کلروفیل، با رنگ تیره تر

نمونه ها از آزمون تورم فاکتور واریانس VIF استفاده شده است (معادله ۲) [۱۱]. برای مقدار VIF مقادیر مختلفی در مطالعات سایرین در نظر گرفته شده است [۷،۳۸]، در این مطالعه مقدار $VIF=10$ به منظور جلوگیری از تورم ضرایب رگرسیون انتخاب شده است.

$$VIF(b_i) = \frac{1}{1-R_i^2} \quad (2)$$

متغیر R_i^2 ضریب تعیین رگرسیون چند گانه ایجاد شده ما بین آمین باند باریک فراطیفی در مقابل سایر باندها می باشد.

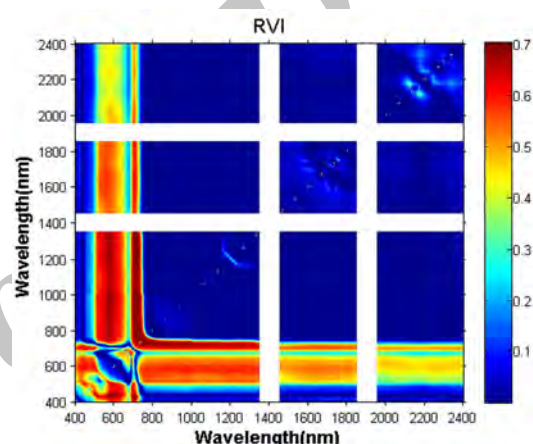
۲-۴- رگرسیون حداقل مربعات جزئی

در هنگام بروز مشکل هم خطی بین متغیرهای مستقل (باندهای طیفی)، رگرسیون حداقل مربعات جزئی (PLSR)، یکی از روش های آماری چند متغیره است که استفاده می شود. در این روش داده های طیفی همبسته با ابعاد بالا به فاکتورهای غیر همبسته به نام متغیر پنهان کاهش می یابند [۱۱،۲۵]. فاکتورهای پنهانی اطلاعات مهمی که در کل طیف برای پیش بینی وجود دارد در بردارند. رگرسیون حداقل مربعات جزئی ترکیبی از روش های چند متغیره تحلیل مولفه های اصلی (PCA) و رگرسیون چندگانه می باشد [۲۶]. تحلیل رگرسیون حداقل مربعات جزئی تا حدی شبیه به رگرسیون تحلیل مولفه های اصلی (PCR) می باشد، با این تفاوت که در روش رگرسیون PLS متغیرهای مستقل نیز همزمان با متغیرهای وابسته به تعدادی مولفه های اصلی (PC) تبدیل می شود به نحوی که بیشترین درصد تغییرات در مولفه های اولیه ارائه می شود [۱۱]. الگوریتم روش PLSR در این مطالعه توضیح داده نمی شود ولی برای توضیحات بیشتر می توان به مقالات [۲۰،۳۰،۴۱] رجوع کرد.

تعیین تعداد متغیرهای پنهان نقش مهمی در اجرای بهتر رگرسیون ایفا می کند و همچنین انتخاب بهینه تعداد فاکتورها از بیش برآزش در رگرسیون جلوگیری می کند [۲۳]. در این مطالعه تعداد بهینه متغیرهای پنهان از روش اعتبار سنجی Leave one out cross validation تعیین شده است. مینیمم شدن مقدار RMSE با فاکتورهای پنهان، روش معمول در انتخاب تعداد فاکتورها می باشد [۲۰]. در این مطالعه شرط اینکه با اضافه شدن

۱ Coefficient of Determination

(قرمز تیره) کاملا مشخص می‌باشند. مشاهده پلات دو بعدی نشان می‌دهد، مناطق تیره تر بیشتر از ترکیب باندهای مرئی و محدوده مادون قرمز نزدیک تشکیل شده اند که این محدوده ها می‌توانند برای طراحی سنجنده های چند طیفی و کاربرد های خاص مانند کشف افت کلروفیل برگ درختان پسته مورد استفاده قرار گیرند علت انتخاب شدن محدوده مرئی را می‌توان به دلیل جذب شدید رنگدانه های کلروفیل و کاروتنوئید در این منطقه طیفی دانست و علت انتخاب شدن محدوده مادون قرمز نزدیک را در اثر ساختار درونی برگ که عامل اصلی تاثیر گذار بر جذب طیفی در این محدوده می‌باشد، دانست [۱۹].



شکل ۲- نمایش دو بعدی میزان همبستگی (R^2) بین غلظت کلروفیل و شاخص باریک باند RVI برای تمام نمونه ها (دو نوار سفید در شکل مناطق جذب بخار اتمسفر می‌باشند که از محاسبات حذف شده اند).

سپس بر اساس ارزش های R^2 در پلات های همبستگی، بهینه ترین زوج طول موج که بیشترین همبستگی را در آن شاخص ها برای پیش بینی کلروفیل نمونه ها داشتند در جدول ۲ مشخص شد.

جدول ۲- موقعیت بهینه ترین باندها و مقدار R^2 بدست آمده بین غلظت کلروفیل و بهینه ترین شاخص باریک باند

	باندهای طیفی اولیه			باندهای طیفی تبدیل شده		
	λ_1	λ_2	R^2	λ_1	λ_2	R^2
NDVI	۹۹۵	۷۲۵	۰,۶۹	۹۸۳	۶۸۳	۰,۷۳
RVI	۹۹۴	۷۲۲	۰,۷۰	۶۷۰	۷۳۴	۰,۷۴
TVI	۱۳۰۷	۷۲۲	۰,۶۹	۹۸۳	۶۸۳	۰,۷۳

باندهای محدوده مرئی انتخاب شده در شاخص ها نزدیک باندهای جذب کلروفیل (۶۴۰ و ۶۶۰ نانومتر) در برگ گیاهان می‌باشند [۱۰]. البته در مواقعی جذب کلروفیل در طول موج های مرکزی اشباع می‌شود و با طول موج های اطراف بهتر شناسایی می‌شود [۱۷]. مثلا [۲۱] در مطالعه خود نشان داده است که طول موج ۷۰۰ نانومتر به شدت به تغییرات کلروفیل حساس است که همین موضوع می‌تواند دلیلی بر انتخاب شدن طول موج های نزدیک ۷۰۰ نانومتر در شاخص های ارائه شده باشد. البته [۲۸] نیز در مطالعه خود به این نتیجه رسید که شاخص هایی که طول موج های لبه قرمز استفاده می‌کنند به خوبی می‌توانند معرف کلروفیل گیاه باشند. باندهای ۹۸۳، ۹۹۵ و ۱۳۰۷ انتخاب شده در محدوده مادون قرمز نزدیک می‌توانند نشانگرهایی از وضعیت سلامت گیاه باشند و برای پیش بینی تفاوت فتوسنتز و دوره رشد برگ استفاده شوند [۴۳].

پس از انتخاب شاخص ها با باندهای بهینه، مقدار کلروفیل نمونه ها با استفاده از هر یک از این شاخص ها به طور جداگانه به عنوان متغیر مستقل با رگرسیون خطی تخمین زده شد. نتایج حاصل از اعتبارسنجی برای تک تک شاخص ها در جدول ۳ نشان داده شده است.

جدول ۳- مقادیر $RRMSE_{cv}$ و R^2_{cv} بدست آمده با روش Cross validation برای تخمین غلظت کلروفیل نمونه های برگ پسته

	باندهای طیفی تبدیل شده			
	باندهای طیفی اولیه		باندهای طیفی تبدیل شده	
	R^2_{cv}	$RRMSE_{cv}$	R^2_{cv}	$RRMSE_{cv}$
NDVI	۰,۶۷	۰,۲۷	۰,۷۱	۰,۲۵
RVI	۰,۶۸	۰,۲۷	۰,۷۲	۰,۲۵
TVI	۰,۶۷	۰,۲۷	۰,۷۱	۰,۲۶

نتایج جدول ۳ نشان می‌دهد، در همه شاخص ها هنگام استفاده از داده های تبدیل شده با مشتق اول مقادیر ضریب تعیین R^2_{cv} حدود ۰,۰۴ افزایش یافته است که این موضوع اهمیت استفاده از داده های تبدیل شده در پیش بینی پارامترهای گیاهی را تایید می‌کند. در میان تمام شاخص ها با انواع داده های بازتابی، شاخص RVI با استفاده از داده های تبدیل شده مشتق

جدول ۴- نتایج حاصل از تخمین Cross validation شده کلروفیل نمونه ها با استفاده از روش SMLR

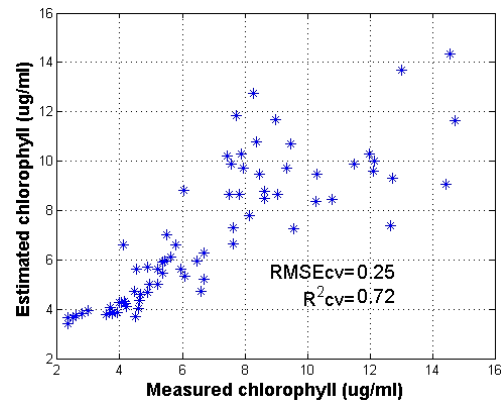
باند های منتخب (nm)	R_{cv}^2	RRMSEcv
باندهای طیفی اولیه	۰,۶۹	۰,۲۶
باندهای طیفی تبدیل شده	۰,۷۸	۰,۲۲

زمانی که از داده های تبدیل شده با مشتق اول جهت پیش بینی استفاده شده است، تعداد باندهای انتخاب شده توسط مدل افزایش یافته است. همزمان با افزایش تعداد باندهای انتخاب شده، R_{cv}^2 نیز از ۰,۶۹ به ۰,۷۹ ارتقا یافته است که این امر را می توان در اهمیت باندهای موجود در محدوده ۱۳۰۰-۲۴۰۰ نانومتر در کنار باندهای موجود در محدوده مرئی و لبه قرمز در پیش بینی مقادیر کلروفیل دانست. باند های ۴۲۷ و ۴۶۵ نزدیک باندهای جذب کلروفیل هستند [۱۰] و همچنین طول موج های ۶۹۸ و ۷۱۸ در محدوده لبه قرمز می باشند که مطالعات گذشته این محدوده را نیز حساس به تغییرات غلظت کلروفیل معرفی کرده اند [۹،۴۳].

نتایج حاصل شده از پیش بینی ها نشان می دهد که استفاده از روش چند متغیره SMLR با داده های تبدیل شده مشتق اول باعث افزایش دقت نسبت به روش های تک متغیره شاخص های گیاهی شده است، زیرا R_{cv}^2 بیشتر و RRMSEcv کاهش یافته است. [۱۱] نیز در مطالعه خود به این نتیجه رسید که روش SMLR در پیش بینی پارامترهای کلروفیل و LAI نسبت به شاخص های باریک باند دقت بالاتری دارد.

سیس روش PLSR با داده های انعکاسی اولیه و تبدیل شده با مشتق اول برای پیش بینی مقادیر کلروفیل استفاده شد. از آنجایی که تعیین تعداد فاکتورهای پنهان بهینه نقش مهمی در اجرای بهتر رگرسیون ایفا می کند [۳۱]، در این مطالعه تعداد فاکتورهای شرکت کننده در مدل از دید بصری بر روی نمودار حاصل از تعداد فاکتورهای پنهان PLSR مقابل مقدار RRMSEcv و همچنین پیاده سازی شرط اینکه با اضافه شدن یک فاکتور پنهان باید به مقدار ۲ درصد RMSEcv کاهش یابد؛ استفاده شد. این نمودار ها در این مقاله نشان داده

اول بالاترین دقت را دارد. طول موج های انتخاب شده در این شاخص ۶۷۰ و ۷۳۰ نانومتر می باشند که باند ۶۷۰ نزدیک طول موج جذب کلروفیل و ۷۳۰ در محدوده لبه قرمز می باشد. [۳،۱۶،۲۸] نیز در مطالعات خود لبه قرمز را در پیش بینی کلروفیل بسیار تاثیرگذار دانستند. شکل ۳ رابطه بین مقدار کلروفیل تخمین زده شده توسط شاخص RVI را در برابر مقدار کلروفیل اندازه گیری شده نمونه ها با داده های طیفی تبدیل شده نشان می دهد.



شکل ۳- نمایش دو بعدی مقادیر کلروفیل اندازه گیری شده در برابر مقادیر تخمین زده شده توسط شاخص باریک باند RVI

همان طور که در جدول ۳ و شکل ۳ مشخص است، رگرسیون خطی با استفاده از هر یک از شاخص های باریک باند دارای توانایی نسبتاً بالایی در تخمین کلروفیل نمونه ها می باشد. مقادیر کلروفیل بالاتر از ۱۰ (um/ml) توسط شاخص های باریک باند به خوبی پیش بینی نشده اند. علت این امر را می توان در اشباع شدن شاخص ها در مقادیر بالای کلروفیل دانست.

۳-۲- روش های آماری چند متغیره و پیش بینی مقادیر کلروفیل

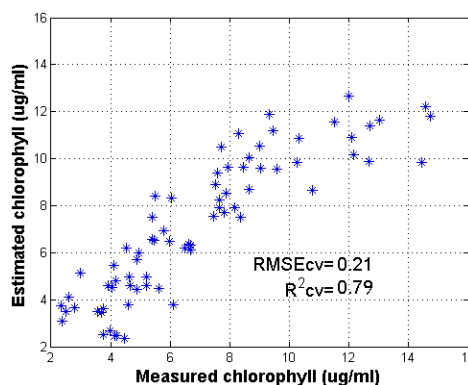
در ابتدا با استفاده از مدل رگرسیون خطی چند گانه مرحله ای موثرترین باندها از نظر آماری برای پیش بینی مقادیر کلروفیل نمونه ها با استفاده از داده های طیفی اولیه و داده های طیفی تبدیل شده با مشتق اول انتخاب شدند. در مرحله بعد باندهای انتخاب شده در رگرسیون چندگانه خطی قرار گرفتند و دقت مدل ها در تخمین کلروفیل با استفاده از روش Cross validation تعیین شد (جدول ۴).

نشده است. نتایج حاصل از تخمین کلروفیل با استفاده از مدل PLSR در جدول ۵ نشان داده شده است.

جدول ۵- نتایج حاصل از تخمین Cross validation شده کلروفیل نمونه ها با استفاده از روش PLSR

تعداد فاکتورها	R_{cv}^2	RRMSEcv
باندهای طیفی اولیه	۰,۶۷	۰,۲۵
باندهای طیفی تبدیل شده	۰,۷۹	۰,۲۱

در تخمین پارامتر کلروفیل هنگام استفاده از داده های بازتابی مشتق اول تعداد فاکتورهای پنهان از ۳ به ۵ افزایش یافته و دقت پیش بینی R_{cv}^2 از ۰,۶۷ به ۰,۷۹ افزایش پیدا کرده است. شکل ۴ رابطه بین مقادیر کلروفیل تخمین زده شده با مدل PLSR را در برابر مقدار اندازه گیری شده نمونه ها با داده های طیفی تبدیل شده نشان می دهد. بر خلاف شاخص های باریک باند، مقادیر بالای کلروفیل در مدل های چند متغیره به خوبی پیش بینی شدند و در مقادیر پایین اشباع نشدند.



شکل ۴- نمایش دو بعدی مقادیر کلروفیل اندازه گیری شده در برابر مقادیر تخمین زده شده توسط مدل PLSR

نتایج بدست آمده اهمیت استفاده از داده های تبدیل شده در پیش بینی پارامترهای بیوشیمیایی گیاهان در هنگام استفاده از مدل PLSR را نشان می دهد. [۱۸] نیز در مطالعه خود هنگام استفاده از داده های تبدیل شده با مشتق اول در مدل PLSR به دقت بالایی نسبت به داده های انعکاسی اولیه رسید. [۸] نیز در مطالعه خود اشاره

داشت که انتخاب ورودی طیفی مناسب در بهبود مدل PLSR در پیش بینی پارامترهای گیاهی تاثیرگذار است. در این تحقیق ارتباط بین غلظت کلروفیل و طیف های اندازه گیری شده با روش های چند متغیره SMLR و PLSR دقت بهتری را نسبت به روش های تک متغیره ارائه داده اند. علت آنرا می توان در این دانست که شاخص ها تنها از تعداد محدودی از اطلاعات طیفی موجود در داده های فراطیفی استفاده کرده است [۱۱]. در مطالعات [۸,۱۱,۲۵,۳۴] نیز استفاده از روش های چند متغیره در پیش بینی پارامترهای گیاهی باعث بهبود دقت تخمین شده است.

۴- جمع بندی و نتیجه گیری

توسعه در زمینه سنجش از دور فراطیفی و تصویربرداری اسپکترومتری دریچه ای جدید به سمت پایش و تخمین پارامترهای گیاهی گشوده است. در این مطالعه روش های آماری چند متغیره SMLR و PLSR و روش های تک متغیره شاخص های باریک باند RVI, NDVI, و TVI در پیش بینی کلروفیل برگ درختان پسته با استفاده از داده های فرا طیفی استفاده شده است. همچنین توانایی داده های انعکاسی تبدیل شده با مشتق اول در پیش بینی کلروفیل مورد ارزیابی قرار گرفت. مهمترین نتایج حاصل از این مطالعه به شرح زیر می باشد: هنگام استفاده از شاخص های باریک باند RVI, NDVI و TVI، کلروفیل نمونه های برگ پسته با دقت قابل قبولی پیش بینی شدند ($R_{cv}^2 > 0.65$).

بهترین شاخص باریک باند برای پیش بینی کلروفیل، شاخص RVI با داده های تبدیل شده و طول موج های ۶۷۰ نانومتر در محدوده مرئی قرمز و نزدیک نقطه جذب کلروفیل و ۷۳۴ نانومتر در محدوده لبه قرمز می باشد.

در اکثر روش ها حداقل یکی از بهینه ترین باندهای انتخاب شده در محدوده مرئی یا لبه قرمز می باشند که علت این امر را حضور رنگدانه های کلروفیل دانست. در این تحقیق روش های چند متغیره SMLR و PLSR دقت بالاتری نسبت به روش های تک متغیره در پیش بینی غلظت کلروفیل داشتند. علت آنرا می توان در این دانست که در شاخص ها تنها از تعداد محدودی از

اگرچه این مطالعه در سطح آزمایشگاهی انجام شده است، نتایج بدست آمده می توانند قابل استفاده در دنیای واقعی باشند و از این نتایج در طراحی محدوده طیفی سنجنده های هوایی و ماهواره ای استفاده کرد. پیشنهاد می شود در صورت طراحی سنجنده ملی در جهت پایش وضعیت سلامت درختان پسته، محدوده های طیفی انتخاب شده در این مطالعه در طراحی سنسور مورد بررسی قرار گیرند.

سپاسگزاری

در انتها از مسئولان پژوهشگاه علوم و تکنولوژی پیشرفته و علوم محیطی کرمان جهت همکاری های لازم در انجام این مطالعه قدردانی می شود.

اطلاعات طیفی موجود در داده های فرا طیفی استفاده می شود.

در تمام مدل ها هنگام استفاده از داده های تبدیل شده با مشتق اول دقت مدل ها نسبت به استفاده از داده های بازتابی اولیه افزایش یافت.

به طور کلی نتایج این تحقیق نشان داد، داده های انعکاسی فراطیفی پتانسیل بالایی در پایش سلامت و وضعیت رشد محصولات کشاورزی، مخصوصا برگ درختان پسته دارند و همچنین مدل های آماری چند متغیره نسبت به روش های تک متغیره شاخص های گیاهی دقت بالاتری در پایش بینی غلظت کلروفیل برگ درختان پسته داشتند. در پایان نیز اثر تبدیل داده های طیفی در مدل ها نشان داد که داده های انعکاسی تبدیل شده نسبت به داده های طیفی اولیه قابلیت بالاتری در پایش بینی میزان کلروفیل دارند.

مراجع

- [1] Afrousheh, M., Hokmabadi, H. and Mirseyed hosseinil, H. (2010). "Effect of nitrogen, iron, magnesium, manganese and molybdenum deficiencies on biochemical and ecophysiological characteristics of pistachio seedling (*Pistacia vera*) Options Méditerranéennes". XIV GREMPA Meeting on Pistachios and Almonds.
- [2] Atzeberg, C., Jarmer, T., Schlerf, M., Kotz, B., & Werner, W. (2003). "Spectroradiometric determination of wheat bio-physical variables Comparison of different empirical-statistical approaches". Remote Sensing in Transition.
- [3] Blackburn, G. A. (1998). "Quantifying Chlorophylls and Carotenoids at Leaf and Canopy Scales: An Evaluation of Some Hyperspectral Approaches". Remote Sensing of Environment. Vol.66, No.3, PP.273-285.
- [4] Boegh, E., Soegaard, H., Broge, N., Hasager, C. B., Jensen, N. O., Schelde, K., et al. (2002). "Airborne multispectral data for quantifying leaf area index, nitrogen concentration, and photosynthetic efficiency in agriculture". Remote Sensing of Environment. Vol.81, No.2-3, PP.179-193.
- [5] Broge, N. H. and Leblanc, E. (2001). "Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density". Remote Sensing of Environment. Vol.76, No.2, PP.156-172.
- [6] Carter, G. A. (1994). "Ratios of leaf reflectances in narrow wavebands as indicators of plant stress". International Journal of Remote Sensing. Vol.15, No.3, PP.697-703.
- [7] Chatterjee, S. and Price, B. (1991). Regression analysis by example, Wiley.
- [8] Cho, M. A., Skidmore, A., Corsi, F., van Wieren, S. E. and Sobhan, I. (2007). "Estimation of green grass/herb biomass from airborne hyperspectral imagery using spectral indices and partial least squares regression". International Journal of Applied Earth Observation and Geoinformation. Vol.9, No.4, PP.414-424.
- [9] Clevers, J. G. P. W., De Jong, S. M., Epema, G. F., Van Der Meer, F. D., Bakker, W. H., Skidmore, A. K., et al. (2002). "Derivation of the red edge index using the MERIS standard band setting". International Journal of Remote Sensing. Vol.23, No.16, PP.3169-3184.

- [10] Curran, P. J. (1989). "Remote sensing of foliar chemistry". *Remote Sensing of Environment*. Vol.30, No.3, PP.271-278.
- [11] Darvishzadeh, R., Skidmore, A., Schlerf, M., Atzberger, C., Corsi, F. and Cho, M. (2008). "LAI and chlorophyll estimation for a heterogeneous grassland using hyperspectral measurements". *ISPRS Journal of Photogrammetry and Remote Sensing*. Vol.63, No.4, PP.409-426.
- [12] Dawson, T. P. and Curran, P. J. (1998). "A new technique for interpolating the reflectance red edge position". *International Journal of Remote Sensing*. Vol.19, No.11, PP.2133-2139.
- [13] Deering, D. W., Rouse, J. W., Haas, R. H. and Schell, J. A. (1975). "Measuring Forage Production of Grazing Units from Landsat MSS Data". 10th International Symposium on Remote Sensing of Environment. PP.1169-1178.
- [14] DeJong, S. M., Pebesma, E. and Lacaze, B. (2003). "Above-ground biomass assessment of Mediterranean forests using airborne imaging spectrometry: the DAIS Peyne experiment". *International Journal of Remote Sensing*. PP.1505 – 1520.
- [15] Dunagan, S. C., Gilmore, M. S. and Varekamp, J. C. (2007). "Effects of mercury on visible/near-infrared reflectance spectra of mustard spinach plants (*Brassica rapa* P.)". *Environmental Pollution*. Vol.148, No.1, PP.301-311.
- [16] Fava, F., Colombo, R., Bocchi, S., Meroni, M., Sitzia, M., Fois, N., et al. (2009). "Identification of hyperspectral vegetation indices for Mediterranean pasture characterization". *International Journal of Applied Earth Observation and Geoinformation*. Vol.11, No.4, PP.233-243.
- [17] Ferwerda, J., Skidmore, A. and Mutanga, O. (2005). "Nitrogen detection with hyperspectral NRI across multiple plant species". *International Journal of Remote sensing and environment*. Vol.26, No.18, PP.4083-4095.
- [18] Ferwerda, J. G. and Skidmore, A. K. (2007). "Can nutrient status of four woody plant species be predicted using field spectrometry?". *ISPRS Journal of Photogrammetry and Remote Sensing*. Vol.62, No.6, PP.406-414.
- [19] Gausman, H., Allen, W. and Escobar, D. (1970). "Refractive index of plant cell walls". *Appl.* Vol.9, No.3, PP.522-545.
- [20] Geladi, P. and Kowalski, B. R. (1986). "Partial least-squares regression: a tutorial". *Analytica Chimica Acta*. Vol.185, No.0, PP.1-17.
- [21] Gitelson, A. A. and Merzilyak, M. N. (1998). "Remote sensing of chlorophyll concentration in higher plant leaves". *Advances in Space Research*. Vol.22, PP.689-692.
- [22] Grossman, Y. L., Ustin, S. L., Jacquemoud, S., Sanderson, E. W., Schmuck, G. and Verdebout, J. (1996). "Critique of stepwise multiple linear regression for the extraction of leaf biochemistry information from leaf reflectance data". *Remote Sensing of Environment*. Vol.56, No.3, PP.182-193.
- [23] Haaland, D. M. and Thomas, E. V. (1988). "Partial Least-Squares Methods for Spectral Analyses. Relation to Other Quantitative Calibration Methods and the Extraction of Qualitative Information". Sandia National Laboratories, Albuquerque.
- [24] Haboudane, D., Miller, J. R., Tremblay, N., Zarco-Tejada, P. J. and Dextraze, L. (2002). "Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture". *Remote Sensing of Environment*. Vol.81, No.2–3, PP.416-426.
- [25] Hansen, P. M. and Schjoerring, J. K. (2003). "Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression". *Remote Sensing of Environment*. Vol.86, No.4, PP.542-553.
- [26] Huang, Z., Turner, B. J., Dury, S. J., Wallis, I. R. and Foley, W. J. (2004). "Estimating foliage nitrogen concentration from HYMAP data using continuum removal analysis". *Remote Sensing of Environment*. Vol.93, No.1–2, PP.18-29.
- [27] Kokaly, R. F. and Clark, R. N. (1999). "Spectroscopic Determination of Leaf Biochemistry Using Band-Depth Analysis of Absorption Features and Stepwise Multiple Linear Regression". *Remote Sensing of Environment*. Vol.67, No.3, PP.267-287.

- [28] le Maire, G., François, C. and Dufrière, E. (2004). "Towards universal broad leaf chlorophyll indices using PROSPECT simulated database and hyperspectral reflectance measurements". *Remote Sensing of Environment*. Vol.89, No.1, PP.1-28.
- [29] Lee, K.-S., Cohen, W. B., Kennedy, R. E., Maieringer, T. K. and Gower, S. T. (2004). "Hyperspectral versus multispectral data for estimating leaf area index in four different biomes". *Remote Sensing of Environment*. Vol.91, No.3-4, PP.508-520.
- [30] Li, B., Morris, J. and Martin, E. B. (2002). "Model selection for partial least squares regression". *Chemometrics and Intelligent Laboratory Systems*. Vol.64, No.1, PP.79-89.
- [31] Li, L. (2008). "Quantifying lunar soil composition with partial least squares modeling of reflectance". *Advances in Space Research*. Vol.42, No.2, PP.267-274.
- [32] Liang, S. (2004). "Quantitative Remote Sensing of Land Surfaces". Hoboken, NJ, USA: Wiley & Sons.
- [33] Lichtenthaler, H. K. (1987). "Chlorophylls and carotenoids: Pigments of photosynthetic biomembranes". *Methods Enzymol*. PP.350-382.
- [34] Nguyen, H. T. and Lee, B.-W. (2006). "Assessment of rice leaf growth and nitrogen status by hyperspectral canopy reflectance and partial least square regression". *European Journal of Agronomy*. Vol.24, No.4, PP.349-356.
- [35] Pearson, R. L. and Miller, L. D. (1972). "Remote mapping of standing crop biomass for estimation of the productivity of the shortgrass prairie, Pawnee National Grasslands, Colorado". *Proceedings of the eighth international symposium on remote sensing of environment*. PP.1357-1381.
- [36] Rouse, J. W. and Haas, R. H. (1973). "Monitoring Vegetation Systems in the Great Plains with ERTS". *Third ERTS Symposium, NASA*. PP. 309-317.
- [37] Savitzky, A. and Golay, M. J. E. (1964). "Smoothing and differentiation of data by simplified least square procedure". *Analytical Chemistry*. Vol. 36, No. 8, PP.1627-1638.
- [38] Sergent, M., Mathieu, D., Phan-Tan-Luu, R. and Drava, G. (1995). "Correct and incorrect use of multilinear regression". *Chemometrics and Intelligent Laboratory Systems*. Vol.27, No.2, PP.153-162.
- [39] Stimson, H. C., Breshears, D. D., Ustin, S. L. and Kefauver, S. C. (2005). "Spectral sensing of foliar water conditions in two co-occurring conifer species: *Pinus edulis* and *Juniperus monosperma*". *Remote Sensing of Environment*. Vol.96, No. 1, PP.108-118.
- [40] Uno, Y., Prasher, S. O., Lacroix, R., Goel, P. K., Karimi, Y., Viau, A., et al. (2005). "Artificial neural networks to predict corn yield from Compact Airborne Spectrographic Imager data". *Computers and Electronics in Agriculture*. Vol.47, No.2, PP.149-161.
- [41] Wold, S., Sjöström, M. and Eriksson, L. (2001). "PLS-regression: a basic tool of chemometrics". *Chemometrics and Intelligent Laboratory Systems*. Vol.58, No.2, PP.109-130.
- [42] Wu, C., Niu, Z., Tang, Q. and Huang, W. (2008). "Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation". *Agricultural and Forest Meteorology*. Vol.148, No.8-9, PP.1230-1241.
- [43] Yao, X., Zhu, Y., Tian, Y., Feng, W. and Cao, W. (2010). "Exploring hyperspectral bands and estimation indices for leaf nitrogen accumulation in wheat". *International Journal of Applied Earth Observation and Geoinformation*. Vol.12, No.2, PP.89-100.