

بازیابی پارامترهای شاخص سطح برگ، مقدار کلروفیل و سطح پوشیده‌شده با پوشش گیاهی با استفاده از رویکرد آماری - تجربی از تصاویر ابرطیفی ماهواره کریس - پروبا در منطقه باراکس اسپانیا

مهدی نادری*

دانشجوی کارشناسی ارشد سنجش از دور و GIS، دانشگاه تربیت مدرس، تهران، ایران.

دریافت: ۱۳۹۹/۶/۱ پذیرش: ۱۴۰۰/۴/۳۱

چکیده

بازیابی پارامترهای بیوفیزیکی و بیوشیمیایی پوشش گیاهی که در طیف گسترده‌ای از برنامه‌ها مانند مطالعات اقلیمی، نرخ فتوسنتز، وضعیت تغذیه‌ای گیاهان و چرخه‌های ژئوشیمیایی مورد استفاده قرار می‌گیرند، از اهمیت بالایی برخوردار است. آگاهی از این پارامترها دید مناسبی درباره سلامت گیاهان، وضعیت رشد آن‌ها و کیفیت پوشش گیاهی فراهم می‌آورد و امکان بررسی طولانی‌مدت پویایی پوشش گیاهی را میسر می‌سازد. به‌طور کلی، رویکردهای بازیابی پارامتر در دو گروه کلی رویکردهای آماری - تجربی و رویکردهای مبتنی بر مدل‌های فیزیکی قرار می‌گیرند. در این مطالعه بازیابی پارامترهای شاخص سطح برگ (LAI)، مقدار کلروفیل و سطح پوشیده‌شده با پوشش گیاهی (FVC) با استفاده از رویکرد آماری - تجربی از تصاویر ماهواره‌ای کریس - پروبا در منطقه باراکس اسپانیا ارائه شده است. در این رویکرد برای بازیابی پارامترهای پوشش گیاهی، نیازمند داده‌های میدانی جمع‌آوری شده و یک تصویر ماهواره‌ای از منطقه مورد مطالعه بوده تا با فراهم کردن این داده‌ها و برقراری ارتباط بین آن‌ها، مدل کالیبره‌شده و با استفاده از روش‌های رگرسیون خطی و غیرخطی نسبت به بازیابی پارامترهای پوشش گیاهی اقدام کرد. نتایج حاصل از این مطالعه نشان دادند که در بازیابی شاخص سطح برگ، مقدار کلروفیل و سطح پوشیده‌شده با پوشش گیاهی به‌ترتیب مدل‌های Exponential

RMSE= 4.55,) Rational Quadratic GPR, (RMSE= 0.78, $R^2=0.77$, MAE=0.49) GPR
 RMSE=0.11, $R^2=0.71$,) Squared Exponential GPR و ($R^2=0.36$, MAE=3.61
 MAE=0.09) بهترین برآورد و برازش با داده‌های میدانی را ارائه دادند. تجزیه و تحلیل نقشه‌های
 بازایی‌شده شاخص‌های LAI، مقدار کلروفیل و FVC نشان دادند که مدل‌های فرایند گاوسی که از
 روش‌های رگرسیون غیرخطی هستند، در مقایسه با روش‌های رگرسیون خطی و روش‌های ماشین‌بردار
 پشتیبان، در بازایی پارامترهای مذکور عملکرد بهتری داشته‌اند. نقشه‌های بازایی‌شده نشان می‌دهند
 که مدل‌های مختلف گاوسی نه تنها در بازایی شکل اراضی کشاورزی موفق عمل کرده، بلکه تغییرات
 شاخص‌های LAI، مقدار کلروفیل و FVC داخل اراضی کشاورزی را نیز با دقت مطلوبی بازایی کرده‌اند.

واژگان کلیدی: بازایی پارامتر، کریس - پروبا، رویکرد آماری - تجربی، سنجش از دور، باراکس.

۱- مقدمه:

ارزیابی و پایش پوشش گیاهی در طیف گسترده‌ای از برنامه‌ها مانند مدل‌های اکوسیستم،
 چرخه کربن، مدل‌سازی اقلیمی، مدل‌سازی رشد محصول و بهره‌وری از آن‌ها مورد نیاز است
 (Baret et al., 2007; D'Urso et al., 2009). این مدل‌ها به اطلاعاتی درباره خصوصیات
 بیوفیزیکی و بیوشیمیایی نیاز دارند تا بتوانند تبادل تابش، آب، گرما، تبادل گازها با اتمسفر و
 خاک پس‌زمینه را نشان دهند (Monteith and Unsworth, 1990). شاخص سطح برگ
 (LAI)^۱، مقدار کلروفیل^۲ و سطح پوشیده‌شده با پوشش گیاهی (FVC)^۳ از مهم‌ترین
 ویژگی‌های پوشش گیاهی مورد استفاده در این مدل‌ها هستند. شاخص سطح برگ بیانگر
 نسبت سطح اشغال‌شده با برگ‌های سبز در واحد سطح زمین است. مقدار کلروفیل بیانگر
 میزان فتوسنتز گیاهان است و سطح پوشیده‌شده با پوشش گیاهی بیانگر سطحی از زمین است
 که با شاخ و برگ سبز پوشیده شده است. در دسترس بودن اطلاعات مکانی و زمانی از LAI
 برای مطالعات تغییرات اقلیمی از اهمیت زیادی برخوردار است (Sellers et al., 1996) و این
 پارامتر می‌تواند با استفاده از اطلاعات طیفی داده‌های اپتیک با دقت مطلوبی مدل‌سازی شود
 (شمس‌الدینی، ۱۳۹۶). میزان کلروفیل گیاهان با نرخ فتوسنتز، وضعیت نیتروژن و تولید اولیه
 انبوه که معمولاً به‌عنوان شاخص‌های زیستی از وضعیت سلامتی گیاهان شناخته می‌شوند،
 مرتبط است (Gitelson et al, 2006b). FVC نیز عامل کنترل‌کننده تعرق، فتوسنتز، تغییرات

۱. Leaf Area Index

۲. Chlorophyll Content

۳. Fraction of Vegetation Cover

جهانی آب و هوا، چرخه‌های ژئوشیمیایی و مدل‌های اقلیمی است (Arneth, 2015; Jiapaer et al., 2011).

بازیابی خصوصیات بیوفیزیکی و بیوشیمیایی پوشش گیاهی با استفاده از مشاهدات ماهواره‌ای سابقه‌ای طولانی مدت دارد. رویکردهای بازیابی پارامتر معمولاً در دو گروه کلی رویکردهای آماری - تجربی و رویکردهای مبتنی بر مدل‌های فیزیکی قرار می‌گیرند. طبقه‌بندی مناسب‌تری (Baret and Buis, 2008) پیشنهاد شده است که شامل رویکرد مبتنی بر متغیر بیوفیزیکی^۱ و رویکرد مبتنی بر داده رادیومتریک^۲ است. اساس کار رویکردهای آماری - تجربی، ایجاد یک رابطه بین متغیر مورد نظر و میدان تابشی است (Darvishzadeh et al., 2008). انواع مدل‌های رگرسیون و همچنین شاخص‌های پوشش گیاهی در این رویکرد شناخته شده‌اند. هرچند این رویکردها ساده‌اند و از سرعت محاسباتی زیادی برخوردارند، برای یک منطقه خاص توسعه یافته و استفاده از آن‌ها در مواقع دیگر به دلیل ماهیت خاص منطقه، زمان و سنسور بسیار محدود می‌شود (Meroni, Colombo & Panigada, 2004). در مقابل، روش‌های مبتنی بر مدل‌های فیزیکی نیازی به آموزش و کالیبراسیون ندارند و به راحتی در هر منطقه و شرایط جغرافیایی قابلیت استفاده را دارند. در چنین حالتی، یک مدل فیزیکی مانند مدل انتقال تابش که تابش موج الکترومغناطیس را شبیه‌سازی می‌کند، در برابر اندازه‌گیری‌های ماهواره معکوس می‌شود. با توجه به شرایط هادامارد^۳، مدل‌های ریاضی پدیده‌های فیزیکی در صورتی به شکل ریاضی معکوس‌شدنی هستند که علاوه بر اینکه جواب‌های مسئله معکوس وجود داشته باشد و یگانه باشند، به صورت پیوسته به داده‌ها وابسته باشند (Knyazikhin et al., 1998). بنابراین، عدم قطعیت مدل و شکل‌گیری مسئله ناسازگار^۴ ناشی از وارونگی مدل، از جمله مشکلات بالقوه این رویکردها در بازیابی پارامتر است. به منظور برآورد پارامترهای بیوفیزیکی و بیوشیمیایی گیاهان مطالعات متعددی با رویکردهای متنوع انجام شده است. در این باره مطالعه‌ای (Pasolli et al., 2015) به منظور بازیابی LAI در علفزارهای کوهستانی رشته‌کوه‌های آلپ با استفاده از تصاویر مادیس و یک الگوریتم پیشنهادی انجام شد. نتایج ارائه شده نشان دادند که الگوریتم پیشنهادی با $RMSE = 1.68$ برای بازیابی LAI مؤثر است و مقادیر تخمین زده شده توسط الگوریتم پیشنهادی از الگوهای مکانی پیروی می‌کند و با روند زمانی که در زمین اندازه‌گیری شده بود، تطابق دارد. نینواز و همکاران (۲۰۱۶) با استفاده از داده‌های ابرطیفی حرارتی و

۱. Variable-driven Approach

۲. Radiometric data-driven Approach

۳. Hadamard Condition

۴. Ill-Posed Problem

مدل‌های رگرسیون حداقل مربعات جزئی (PLSR)^۱، شبکه عصبی مصنوعی (ANN)^۲ و شاخص‌های پوشش گیاهی به بازیابی LAI چهار گونه مختلف گیاهی اقدام کردند. نتایج نشان دادند که ANN دقت بیشتری در بازیابی LAI نسبت به PLSR و شاخص‌های پوشش گیاهی ارائه می‌دهد و LAI می‌تواند حتی برای مقادیر بالای $5.5(m^2/m^2)$ از داده‌های ابرطیفی حرارتی با موفقیت تخمین زده شود. در مطالعه‌ای (مهدوی‌فرد، خلیل ولی‌زاده و عطا‌زاده، ۱۳۹۹) که با استفاده از داده‌های میدانی و پرازش تصاویر سنتینل-۲ و لندست-۸ به منظور تخمین غلظت کلروفیل-آ انجام گرفت، از الگوریتم‌های بیوپتیک OC2 و OC3 استفاده شد. نتایج تحقیق، عملکرد مناسب الگوریتم OC2 در تصاویر لندست-۸ را نشان داد که می‌تواند با دقت بیشتری مقدار کلروفیل-آ را بازیابی کند. در تحقیقی (Bei et al., 2019) که به منظور بازیابی مقدار کلروفیل برگ گندم با استفاده از شبیه‌سازی داده‌های RapidEye، سنتینل-۲ و EnMAP انجام شد، از روش‌های PLSR و رویکرد مبتنی بر شاخص‌های پوشش گیاهی استفاده شد. نتایج نشان دادند که برای رویکرد PLSR، داده‌های سنتینل-۲ دقت مناسبی در تخمین میزان کلروفیل برگ نشان دادند. برای رویکردهای مبتنی بر شاخص پوشش گیاهی، شاخص کلروفیل زمینی طیف‌سنج تصویربرداری با وضوح متوسط (MERIS)^۳ بیشترین حساسیت را در تخمین مقدار کلروفیل از داده‌های ماهواره‌ای سنتینل-۲ و EnMAP داشت و شاخص‌های پوشش گیاهی اختلافی نرمال شده لبه‌قرمز^۴، شاخص نسبت ساده لبه‌قرمز^۵ و شاخص کلروفیل لبه‌قرمز^۶ بیشترین حساسیت را در تخمین کلروفیل از داده‌های RapidEye نشان دادند. در مطالعه‌ای (Mu et al., 2018) که به منظور تخمین FVC با استفاده از شاخص‌های پوشش گیاهی چند زاویه‌ای انجام شد، از یک رویکرد بهبودیافته به نام Multi VI استفاده شد. نتایج نشان دادند که صحت روش Multi VI در تخمین FVC ($R^2 = 0.866$ و $RMSD = 0.092$) نسبت به محصول نوع ۱ بیوژئوفیزیکی اسپات^۷ بیشتر است ($R^2 = 0.795$ و $RMSD = 0.159$). ژانگ^۸ و همکاران (۲۰۱۸) مدل تخمین جهانی FVC محصولات مختلف را با استفاده از سری

۱. Partial Least Squares Regression

۲. Artificial Neural Network

۳. Moderate Resolution Imaging Spectrometer Terrestrial Chlorophyll Index

۴. Red-edge Normalized Difference Vegetation Index

۵. Red-edge Simple Ratio

۶. Red-edge Chlorophyll Index

۷. SPOT/VEGETATION bio GEO Physical Product Version 1 (GEOV1)

۸. Zhang

بازیابی پارامترهای شاخص سطح برگ مقدار ...
مهدی نادری

زمانی تصاویر دیجیتال و شاخص‌های گیاهی ارائه دادند. نتایج نشان دادند که شاخص رنگی استخراج پوشش گیاهی^۱ بهترین عملکرد را در تخمین FVC محصولات مختلف نشان داد. در این پژوهش با استفاده از تصاویر ابرطیفی ماهواره کریس-پروبا و رویکرد آماری-تجربی به بازیابی پارامترهای شاخص سطح برگ، مقدار کلروفیل و سطح پوشیده‌شده با پوشش گیاهی می‌پردازیم. هدف اصلی این پژوهش، ایجاد یک مدل کالیبره پس از برقراری رابطه بین داده‌های میدانی و مقادیر بازتاب ماهواره با استفاده از مدل‌های رگرسیون خطی و غیرخطی است تا بر اساس آن‌ها بتوان نقشه تغییرات پارامترهای پوشش گیاهی را بازیابی کرد. در نهایت بهترین برازش و نقشه بازیابی‌شده با توجه به تجزیه و تحلیل‌های آماری ارائه می‌گردد.

۲- داده و روش تحقیق:

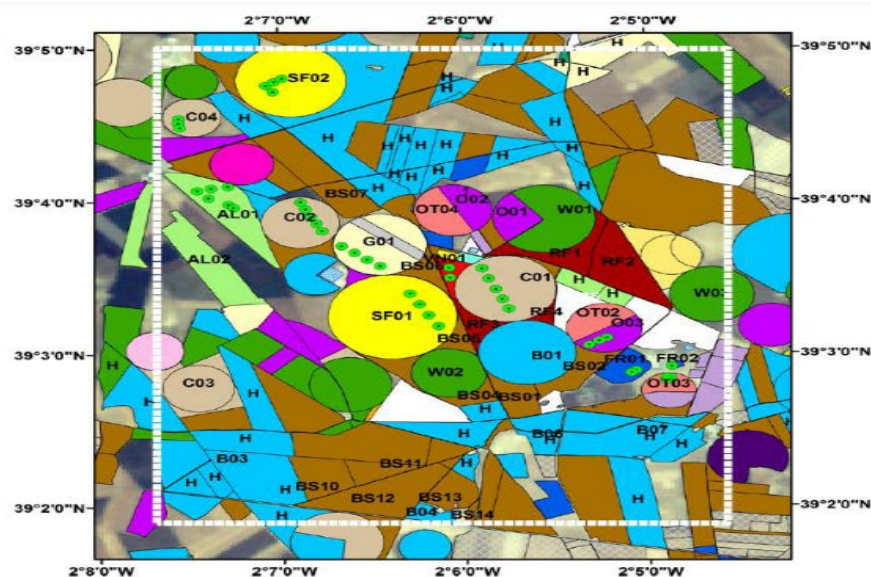
۲-۱- منطقه مورد مطالعه:

منطقه باراکس^۲ (2° 39' N, 60° 3' W) در جنوب شرقی اسپانیا در فلات لمانچا^۳ واقع شده است. این منطقه با مزارع بزرگ کشاورزی یکنواخت شناخته می‌شود که تقریباً ۶۵٪ محصولات دیم و ۳۵٪ محصولات نیز به صورت آبی هستند. توپوگرافی مسطح (با شیب متوسط کمتر از ۲ درصد)، برخی مزارع کشاورزی با قطر یک کیلومتر و تنوع محصولات کشاورزی از جمله مشخصه‌های این منطقه محسوب می‌شوند که اهداف مناسبی برای مطالعات هستند. کشاورزی این منطقه عمدتاً یونجه، ذرت، پیاز، سیر، جو دوسر، سیب‌زمینی، آفتابگردان، جو و گندم را شامل می‌شود. در دو دهه اخیر سازمان فضایی اروپا چندین کمپین گردآوری داده‌های زمینی برای کالیبراسیون-ارزیابی ماهواره‌های جدید در منطقه باراکس اجرا کرده است (موسیوند، ۱۳۹۶). منطقه مورد مطالعه به همراه کاربری اراضی آن در شکل (۱) نشان داده شده است.

۱. Color Index of Vegetation Extraction

۲. Barrax

۳. La Mancha



شکل ۱. نقشه کاربری اراضی منطقه باراکس اسپانیا. خطوط خطچین سفید معرف منطقه مورد مطالعه است. نقاط سبزرنگ، معرف مشاهدات زمینی است و کاربری‌های منطقه در راهنمای نقشه آورده شده است. مزارع جو با نوشته H پیش از فعالیت کمپین گردآوری داده زمینی برداشت شده‌اند (موسیوند، ۱۳۹۶).

۲-۲- داده‌های تحقیق:

۱. به منظور بازیابی پارامترهای شاخص سطح برگ (LAI)، مقدار کلروفیل و سطح پوشیده‌شده با پوشش گیاهی (FVC) از داده‌های میدانی جمع‌آوری‌شده در کمپین SEN3EXP^۱ استفاده شده است. این کمپین را سازمان فضایی اروپا در تاریخ ۲۰ الی ۲۴ ژوئن ۲۰۰۹ با هدف کالیبراسیون و ارزیابی ماهواره سنتینل-۳ انجام داده است. در طول این کمپین، بسیاری از خصوصیات بیوفیزیکی و بیوشیمیایی گیاهان و خاک مانند شاخص سطح برگ، میزان کلروفیل، میزان آب برگ^۲، میزان توده خشک برگ^۳، رطوبت خاک^۴ و سطح پوشیده‌شده با پوشش

۱. Sentinel-3 Experimental Campaign

۲. Leaf Water Content

۳. Dry Matter Content

۴. Soil Moisture

بازیابی پارامترهای شاخص سطح برگ مقدار ...
مهدی نادری

گیاهی همزمان با گذر ماهواره‌ها اندازه‌گیری شده است (موسیوند، ۱۳۹۶). پس از نرمال‌سازی و حذف مقادیر نویز از مجموعه داده‌های میدانی گردآوری‌شده، تعداد ۳۷ نمونه زمینی برای بازیابی پارامترهای LAI و FVC و تعداد ۳۳ نمونه زمینی نیز برای بازیابی پارامتر مقدار کلروفیل استفاده شد.

۲. تصویر ماهواره‌ای سنجنده کریس - پروبا^۱: از جمله ماهواره‌های ابرطیفی است که در ۶۲ باند طیفی در محدوده‌های مرئی و مادون قرمز نزدیک (۴۰۰-۱۰۵۰ نانومتر) تصویربرداری می‌کند که قدرت تفکیک مکانی آن بین ۱۸ تا ۳۴ متر متغیر است. علاوه بر این، این ماهواره قادر به تصویربرداری در چند زاویه نیز است (۵۵±۳۶، ۰). یک تصویر این سنجنده در تاریخ 2009.06.19 در این تحقیق مورد استفاده قرار گرفته است.

۳-۲- روش انجام تحقیق:

روش انجام تحقیق به گونه‌ای است که در وهله اول پس از تحلیل حساسیت^۲ بر روی باندهای سنجنده کریس - پروبا، باندهایی که بیشترین حساسیت را نسبت به تغییرات پارامترهای LAI، مقدار کلروفیل و FVC دارند شناسایی شد و در وهله دوم انواع روش‌های رگرسیون خطی و غیرخطی به منظور ایجاد یک مدل کالیبره بر روی مجموعه داده‌های میدانی اجرا شد. مدل ایجادشده بر اساس متغیرهای وابسته (پارامترهای شاخص سطح برگ، مقدار کلروفیل و سطح پوشیده‌شده با پوشش گیاهی) و متغیر مستقل (باندهای سنجنده کریس - پروبا) است. به منظور آموزش و اعتبارسنجی مدل از روش اعتبارسنجی متقاطع^۳ استفاده شد. عملکرد این روش به گونه‌ای است که مجموعه داده‌ها را به صورت تصادفی به K زیرمجموعه یا لایه^۴ با حجم یکسان تقسیم می‌کند و هر بار یکی از زیرمجموعه‌ها برای اعتبارسنجی و K-1 زیرمجموعه نیز برای آموزش مدل مورد استفاده قرار می‌گیرند. در نهایت، نتیجه این K تکرار برای دستیابی به یک برآورد نهایی، میانگین‌گیری می‌شود (نوایی و کوشا، ۱۳۹۵). باتوجه به وجود همبستگی بالا بین باندهای سنجنده‌های ابرطیفی مانند کریس - پروبا، از تجزیه و تحلیل مؤلفه اصلی^۵ به منظور کاهش همبستگی بین باندها استفاده شد. تجزیه و تحلیل مؤلفه اصلی از جمله روش‌های ریاضی متراکم‌سازی داده‌ها محسوب می‌شود که یک مجموعه با ابعاد پایین را از یک مجموعه با ابعاد

۱. Chris-Proba

۲. Sensitivity Analysis

۳. Cross-validation

۴. Fold

۵. Principal Component Analysis



بالا استخراج می‌کند تا به ثبت اطلاعات بیشتر با تعداد کمتری از متغیرها کمک کند. د نهایت پس از اخذ خروجی‌های مورد نظر و تجزیه و تحلیل آماری، ۳ مدل رگرسیون باتوجه به نتایج تحلیل حساسیت به عنوان مدل‌های نهایی رگرسیون به منظور بازیابی پارامترهای LAI، مقدار کلروفیل و FVC ارائه می‌گردند. گفتنی است که تمامی مراحل تحقیق اعم از آماده‌سازی داده‌ها، کدنویسی و اخذ خروجی‌های مورد نظر در محیط نرم‌افزاری متلب^۱ و ابزار Regression Learner انجام شده است. باتوجه به محدودیت تعداد صفحات مقاله، در ادامه به‌طور مختصر به تفصیل مدل‌های رگرسیون و معیارهای آماری مورد استفاده در این پژوهش می‌پردازیم:

۲-۴-۲- مدل‌های رگرسیون خطی^۲:

رگرسیون خطی یک رویکرد مدل خطی بین متغیر پاسخ^۳ با یک یا چند متغیر توصیفی^۴ است. در این حالت فرض بر این است که یک یا چند متغیر توصیفی که مقدار آن‌ها مستقل از بقیه متغیرها یا تحت کنترل کاربر است، می‌تواند در پیش‌بینی متغیر پاسخ که مقدارش وابسته به متغیرهای توصیفی و تحت کنترل کاربر نیست، مؤثر باشد. مدل‌های رگرسیون خطی استفاده‌شده در این تحقیق در سه گروه رگرسیون خطی، رگرسیون خطی قوی^۵ و رگرسیون خطی گام به گام^۶ قرار می‌گیرند.

۲-۴-۲-۱- رگرسیون خطی:

اساس رگرسیون خطی، ایجاد یک رابطه بین متغیرهای وابسته و مستقل به‌منظور پیش‌بینی متغیر وابسته است. مدل ریاضی همه مدل‌های رگرسیون خطی به‌صورت رابطه (۱) است.

$$Y = a + bX \quad (1)$$

در رابطه فوق، X متغیر توصیفی، Y متغیر وابسته، b شیب خط و a مقدار بایاس^۷ است.

۲-۴-۲-۲- رگرسیون خطی قوی:

۱. Matlab

۲. Linear Regression Models

۳. Response Variable

۴. Explanatory Variable

۵. Robust Linear Regression

۶. Stepwise Linear Regression

۷. Bias

بازیابی پارامترهای شاخص سطح برگ مقدار ...
مهدی نادری

نوعی تجزیه و تحلیل رگرسیون است که برای غلبه بر برخی محدودیت‌های روش‌های سنتی پارامتریک و ناپارامتریک طراحی شده است. در بسیاری از روش‌های رگرسیون، اگر فرضیات اصلی آن‌ها صحیح باشند، بسیار کارآمد و اگر صحیح نباشند، می‌توانند به نتایج گمراه‌کننده‌ای منجر شوند، اما روش رگرسیون قوی به گونه‌ای طراحی شده است که بیش‌ازحد تحت تأثیر نقض مفروضات قرار نمی‌گیرد. رگرسیون خطی قوی جایگزین مناسبی برای رگرسیون حداقل مربعات در صورت عدم تحقق فرضیات اصلی با ماهیت داده‌هاست (Marona et al., 2006).

۲-۴-۳- رگرسیون خطی گام به گام:

هدف رگرسیون گام به گام انتخاب یک مدل با اضافه یا حذف کردن یک متغیر پیش‌بینی‌کننده به‌صورت گام به گام و بر اساس مفهوم آماری است. نتیجه این روند، یک مدل رگرسیون واحد است. تحلیل گام به گام دارای توالی رو به جلو یا عقب^۱ است که توالی رو به جلو متداول‌تر است (Wang and Chen, 2016).

۲-۵- مدل‌های رگرسیون درختی^۲:

یک درخت تصمیم‌باینری است که در هر یک از شاخه‌های آن روابط رگرسیونی خطی وجود دارند که می‌توانند مقادیر عددی را پیش‌بینی کنند. به عبارتی دیگر، این مدل یک فرایند تکراری است که داده‌ها را در شاخه‌های مختلف قرار می‌دهد و این روند تا زمانی که هر شاخه به حداقل اندازه مشخص‌شده توسط کاربر برسد و به یک شاخه نهایی تبدیل شود، ادامه می‌یابد. این مدل‌ها در سه گروه رگرسیون‌های درختی ریز^۳، متوسط^۴ و درشت^۵ قرار می‌گیرند. اختلاف مدل‌های مختلف رگرسیون‌های درختی در حداقل اندازه برگ‌هاست که اندازه برگ‌ها در مدل‌های رگرسیون ریز، متوسط و درشت به ترتیب ۴، ۱۲ و ۳۶ است.

۲-۶- مدل‌های رگرسیون ماشین‌بردار پشتیبان^۶:

۱. Forward or Backward Progression

۲. Regression Tree Models

۳. Fine Tree Regression

۴. Medium Tree Regression

۵. Coarse Tree Regression

۶. Support Vector Machine Regression Models

از جمله روش‌های یادگیری ماشین و تکنیک‌های ناپارامتریک هستند که با بهره‌گیری از توابع کرنل، در دو رویکرد طبقه‌بندی و رگرسیون مورد استفاده قرار می‌گیرند. اساس این روش بر مبنای تئوری یادگیری آماری^۱ استوار است که اصل کمینه‌سازی خطای ساختاری^۲ را به کار برده تا بتواند به یک نتیجه بهینه کلی برسد (Vapnik et al., 1998). رگرسیون بردار پشتیبان (SVR)^۳ که مدل رگرسیونی SVM است، تابع $F(x)$ را برای مجموعه داده‌های آموزشی x به گونه‌ای برازش می‌دهد که بیشینه حاشیه را از مقادیر آموزشی Y داشته باشد. علاوه بر این، این الگوریتم خطاهایی را که در فاصله معینی از مقادیر واقعی قرار دارند نادیده می‌گیرد (Smola and Scholkopf, 2004). مدل‌های رگرسیون SVM می‌توانند حالت خطی یا غیرخطی داشته باشند. در حالت خطی اگر مجموعه داده‌های آموزشی n و D یک مجموعه باشد:

$$D = \{(X_i, Y_i) \mid X_i \in \mathbb{R}^p, Y_i \in \{-1, 1\}\}_{i=1}^n \quad (2)$$

مقدار Y برابر ۱ یا -۱ و هر X_i یک بردار حقیقی با ابعاد p است. در این روش هدف پیدا کردن برازشی است که بیشترین فاصله را با نقاط حاشیه داشته باشد و نقاط $Y_i = 1$ را از نقاط $Y_i = -1$ جدا کند. بنابراین تابع رگرسیون به صورت زیر بیان می‌شود:

$$F(x) = \langle w, x \rangle + b \quad (3)$$

که $\langle \cdot \rangle$ ضرب داخلی است. به منظور مطالعه نحوه محاسبه پارامترهای w و b و همچنین بهینه‌سازی تابع فوق، علاقه‌مندان می‌توانند به مقاله «شکری، صادقی و احمدی مروس، ۱۳۹۲» رجوع کنند. در حالت غیرخطی، اگر یک مجموعه همانند رابطه (۲) داشته باشیم، تابع رگرسیون به شرح زیر خواهد بود:

$$F(x, w) = w^T x + b \quad (4)$$

به منظور برآورد تابع F نیاز داریم تا پارامترهای w و b را محاسبه کنیم. با توجه به محدودیت تعداد صفحات مقاله، علاقه‌مندان می‌توانند به منظور نحوه بهینه‌سازی و محاسبه‌های پارامترهای مذکور به مقاله «احمدی و امیری بختیار، ۱۳۹۷» رجوع کنند. یکی از متداول‌ترین روش‌ها

۱. Statistical Learning theory

۲. Structural Risk Minimization

۳. Support Vector Regression

بازیابی پارامترهای شاخص سطح برگ مقدار ...
 مهدی نادری

به منظور حل مسائل غیرخطی، استفاده از توابع کرنل است که بر اساس ضرب داخلی داده‌های مفروض تعریف می‌شود. بر اساس توابع کرنل اعمال شده بر روی مدل‌های غیرخطی SVM، می‌توان انواع مدل‌های رگرسیون مانند رگرسیون‌های خطی^۱، درجه دوم^۲، درجه سوم^۳، گاوسی ریز^۴، گاوسی متوسط^۵ و گاوسی درشت^۶ را تعریف کرد که همه آن‌ها در این تحقیق مورد استفاده قرار گرفته‌اند. گفتنی است که اختلاف مدل‌های رگرسیون گاوسی تنها در ایجاد ساختارهای ریز تا درشت در داده‌ها است.

۲-۱-۶- تابع کرنل رگرسیون خطی SVM :

$$K(x_j, x_k) = x_j^T x_k \quad (5)$$

۲-۲-۶- تابع کرنل رگرسیون چندجمله‌ای SVM :

$$K(x_j, x_k) = 1 + (x_j^T x_k)^q \quad (6)$$

۲-۳-۶- تابع کرنل رگرسیون گاوسی SVM :

$$K(x_j, x_k) = \exp(-\|x_j - x_k\|^2) \quad (7)$$

x_j و x_k بردارهای ورودی هستند که تنها مرز حاشیه را لمس می‌کنند (Madhu, Paul and Roy, 2019).

۲-۷- مدل‌های رگرسیون ترکیبی درختی^۷:

به روش‌هایی اطلاق می‌شود که با ترکیب چند مدل رگرسیون ضعیف، منجر به تولید یک مدل رگرسیون دقیق می‌شوند و مدل نهایی باتوجه به مشخصه‌هایی نظیر دقت بالا می‌تواند به منظور پیش‌بینی، طبقه‌بندی یا رگرسیون مورد استفاده قرار گیرد. این مدل‌ها در سه گروه کلی روش‌های دسته‌بندی^۸، روش‌های تقویت کردن^۹ و درخت تصادفی^۱ قرار می‌گیرند که در این

-
۱. Linear Support Vector Machine(SVM) Regression
 ۲. Quadratic Support Vector Machine(SVM) Regression
 ۳. Cubic Support Vector Machine(SVM) Regression
 ۴. Fine Gaussian Support Vector Machine(SVM) Regression
 ۵. Medium Gaussian Support Vector Machine(SVM) Regression
 ۶. Coarse Support Vector Machine(SVM) Regression
 ۷. Ensembles of Trees Regression Models
 ۸. Bagging Methods
 ۹. Boosting methods

مطالعه از دو روش دسته‌بندی و تقویت کردن استفاده شده است. در روش دسته‌بندی، یک زیرمجموعه از مجموعه داده‌های اصلی به هر کدام از طبقه‌بندها داده می‌شود تا مدل خود را بر اساس همان قسمت از داده‌هایی که در اختیارش قرار گرفته است، بسازد. روش تقویت کردن به جای انتخاب نمونه‌های آموزشی به صورت تصادفی و با استفاده از توزیع یکنواخت، نمونه‌هایی را که طبقه‌بند اول نتوانسته است به درستی طبقه‌بندی کند، با احتمال بیشتری توسط طبقه‌بند دوم انتخاب و آموزش می‌دهد. در هر دو روش، بعد از چند مرحله تکرار و با در نظر گرفتن میانگین وزنی هر طبقه‌بند و مدل رگرسیون ساخته شده توسط آن، مدل نهایی رگرسیون تولید می‌شود.

۲-۸- مدل‌های رگرسیون فرایند گاوسی:

از جمله مدل‌های احتمالی مبتنی بر کرنل ناپارامتریک با مجموعه متناهی از متغیرهای تصادفی با توزیع چندمتغیره هستند. در این روش فرض می‌شود که مقدار خطا نرمال و دارای توزیع یکسان با میانگین صفر و واریانس (σ^2) است. یک فرایند گاوسی با استفاده از توابع میانگین و کوواریانس تعریف می‌شود. اگر یک مجموعه داده با تعداد n مشاهده وجود داشته باشد که x بردار ورودی با ابعاد D و y مقادیر تابع هدف باشد:

$$M(x) = E(F(x)) \quad , \quad K(x, x') = \text{COV}(F(x), F(x')) = E\{(F(x) - m(x))(F(x') - m(x'))\}$$

با توجه به توابع فوق می‌توان فرایند گاوسی را به صورت معادله زیر تعریف کرد:

$$y = F(x) \approx \text{GP}(m(x), K(x, x')) \quad (8)$$

مدل‌های رگرسیون فرایند گاوسی که در این مطالعه مورد استفاده قرار گرفتند، در چهار گروه کلی مدل‌های رگرسیون گاوسی سهمی منطقی^۲، مربع نمایی^۳، مترن $5/2$ ^۴ و نمایی^۵ قرار می‌گیرند. در بسیاری از توابع کرنل استاندارد، پارامترهای کرنل بر اساس انحراف معیار سیگنال (σ_f) و پارامتر مقیاس طول (σ_l) ^۱ تعریف می‌شوند. هر دو پارامتر نیاز دارند که بزرگ‌تر از صفر باشند، اگر:

-
۱. Random Forest
 ۲. Rational Quadratic Gaussian Process Regression(GPR) Models
 ۳. Squared Exponential Gaussian Process Regression(GPR) Models
 ۴. Matern 5/2 Gaussian Process Regression(GPR) Models
 ۵. Exponential Gaussian Process Regression(GPR) Models
 ۶. Signal Standard Deviation

بازیابی پارامترهای شاخص سطح برگ مقدار ...
 مهدی نادری

$$\theta_1 = \log \sigma_1, \quad \theta_2 = \log \sigma_f$$

۲-۸-۱- تابع کرنل رگرسیون گاوسی سهمی منطقی:

$$K(x_i, x_j | \theta) = \sigma_f^2 \left(1 + \frac{r^2}{2\alpha\sigma_1^2} \right)^{-\alpha}, \quad r = \sqrt{(x_i - x_j)^T (x_i - x_j)} \quad (9)$$

α پارامتر مثبت ترکیبی مقیاس^۲ و r فاصله اقلیدسی بین x_i و x_j است.

۲-۸-۲- تابع کرنل رگرسیون گاوسی مربع نمایی:

این مدل شبیه مدل گاوسی نمایی است، با این تفاوت که فاصله اقلیدسی مربع است. مزیت این روش در استفاده از مجموعه داده‌های بزرگ با حداقل مقدار خطاست. علاوه بر این، عدم پیوستگی را به خوبی بهبود می‌بخشد.

$$K(x_i, x_j | \theta) = \sigma_f^2 \exp \left[-\frac{1}{2} \frac{(x_i - x_j)^T (x_i - x_j)}{\sigma_1^2} \right] \quad (10)$$

۲-۸-۳- تابع کرنل رگرسیون گاوسی مترن^{۵/۲}:

یک کرنل ثابت و تعمیم‌یافته کرنل تابع شعاعی محور (RBF)^۳ است. این روش تراکم طیفی از یک کرنل ثابت را گرفته و تبدیل فوری و سریع از کرنل RBF ایجاد می‌کند.

$$K(x_i, x_j | \theta) = \sigma_f^2 \left(1 + \frac{\sqrt{5}r}{\sigma_1} + \frac{5r^2}{3\sigma_1^2} \exp \left(\frac{\sqrt{5}r}{\sigma_1} \right) \right), \quad r = \sqrt{(x_i - x_j)^T (x_i - x_j)} \quad (11)$$

۲-۸-۴- تابع کرنل رگرسیون گاوسی نمایی:

این روش محصولات داخلی توابع پایه را با کرنل‌ها جایگزین می‌کند، اما سرعت عملکرد کرنل نسبت به روش مربع نمایی کمتر است. اگرچه این روش توابع صاف را با حداقل خطا به کار می‌برد، در بهبود ناپیوستگی‌ها عملکرد مناسبی ندارد.

$$K(x_i, x_j | \theta) = \sigma_f^2 \exp \left(-\frac{r}{\sigma_1} \right) \quad (12)$$

۲-۹- معیارهای ارزیابی نتایج:

۱. Characteristics Length Scale

۲. Positive-valued Scale-mixture Parameter

۳. Radial-basis function kernel

به‌منظور تجزیه و تحلیل و همچنین ارزیابی مدل‌های رگرسیون استفاده شده در روند بازایی پارامترهای LAI، مقدار کلروفیل و FVC از معیارهای ضریب همبستگی^۱، جذر میانگین مربعات خطا^۲ و میانگین خطای مطلق^۳ استفاده شد.

$$R^2 = 1 - \frac{SSR}{SST} = 1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (۱۳)$$

SSR مجموع مربعات خطای رگرسیون و SST مجموع مربعات خطای کل است.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\text{Predicted} - \text{Actual})^2}{N}} \quad (۱۴)$$

Predicted مقادیر پیش‌بینی‌شده، Actual مقادیر واقعی و N تعداد کل داده‌های آموزشی است.

$$MAE = \frac{1}{n} \sum_{i=1}^n |X_i - \hat{X}_i| \quad (۱۵)$$

X_i مقادیر واقعی و $\hat{X}_i$ مقادیر برآورد شده است.

۳- بحث و نتایج:

۳-۱- بازایی شاخص سطح برگ (LAI):

بر اساس شکل (۲-الف) که نمودار پراکندگی LAI اندازه‌گیری‌شده در مقابل LAI پیش‌بینی‌شده توسط مدل‌های رگرسیون را نشان می‌دهد، مدل Exponential GPR با $RMSE=0.78$ ، $R^2=0.77$ و $MAE=0.49$ بهترین برآورد از مقادیر زمینی LAI را نشان داد. مدل Rational Quadratic GPR با $RMSE=0.79$ ، $R^2=0.76$ و $MAE=0.54$ و مدل Matern 5/2 با $RMSE=0.80$ ، $R^2=0.75$ و $MAE=0.56$ به‌ترتیب بهترین برازش را بعد از مدل Exponential GPR ارائه دادند. بر اساس شکل (۲-الف)، هر ۳ مدل در تخمین مقادیر پایین LAI عملکرد قابل قبولی داشته‌اند، ولی با افزایش مقادیر LAI شاهد پراکندگی بیشتر داده‌ها نسبت به خط نرمال هستیم. علاوه بر این، در مقادیر بالای LAI کم‌تخمینی^۴ مشهود است که این کم‌تخمینی برای هر سه مدل دارای رفتار مشابهی است. نقشه محدود تغییرات LAI بازایی‌شده (بر حسب m^2/m^2) که در شکل (۳-الف) نمایش داده شده است با میانگین گرفتن از مقادیر LAI برای هر پیکسل در ۵ فولد اعتبارسنجی متقاطع محاسبه شده است. با توجه به نقشه خروجی، مدل Exponential GPR عملکرد قابل قبولی در بازایی LAI فراهم آورده است؛

۱. Correlation Coefficient

۲. Root Mean Square Error (RMSE)

۳. Mean Absolute Error (MAE)

۴. Under-fitting

به گونه‌ای که این مدل در بازیابی ذرت با پوشش تقریباً متراکم و $LAI (m^2/m^2)$ 2.9 عملکرد نسبتاً قابل قبولی داشته است، اگرچه در بعضی نقاط بیش‌ازاندازه تخمین زده شده است. همچنین این مدل در بازیابی پیاز با پوشش تنک و LAI کمتر از $0.2 (m^2/m^2)$ ، آفتابگردان با پوشش تنک و LAI کمتر از $2 (m^2/m^2)$ و گیاه یونجه با پوشش متراکم عملکرد مناسبی داشته است. نتایج دیگری نیز حاصل شده است، اما باتوجه به محدودیت تعداد صفحات مقاله آورده نشده‌اند. بالاین‌حال، این مدل در بازیابی مقادیر پایین LAI (مانند کاربری‌های خاک لخت و مزارع جو) عملکرد مناسبی را نداشته است و در بازیابی مقادیر بالای LAI بهتر عمل کرده است. عملکرد مدل Rational Quadratic GPR نیز همانند مدل Exponential GPR است و کاربری‌های موجود در منطقه با دقت مطلوبی بازیابی شده‌اند، اما این مدل نیز در برآورد مقادیر پایین LAI عملکرد خوبی نداشته است، به گونه‌ای که میزان LAI خاک لخت توسط مدل مذکور تقریباً $0.5 (m^2/m^2)$ برآورد شده است که بیان‌کننده ضعف مدل است. مدل Matern 5/2 نیز در بازیابی مقادیر بالای LAI عملکردی مشابه سایر مدل‌ها داشته است، به گونه‌ای که برای کاربری یونجه در بخش‌هایی عملکرد مطلوب‌تری نسبت به مدل‌های قبلی داشته است، اما باوجود این، در تخمین مقادیر پایین LAI دچار مشکل بوده و عملاً مقادیر کمتر از $1 (m^2/m^2)$ را بازیابی نکرده است. نتایج نشان می‌دهند که روش رگرسیون فرایند گاوسی که ازجمله روش‌های ناپارامتریک محسوب می‌شود، دقت قابل‌قبولی در بازیابی شاخص سطح برگ نسبت به سایر روش‌ها ارائه داده است. اگرچه در مطالعات متعددی به‌منظور بازیابی LAI از محدوده‌های لبه‌قرمز و مادون قرمز نزدیک استفاده شده است (zhu et al., 2011; Huang et al., 2011a)، نتایج حاصل از تحلیل حساسیت نشان داد که باندهایی که در محدوده مرئی، لبه‌قرمز^۱ و مادون قرمز نزدیک هستند، منجر به افزایش دقت بازیابی LAI می‌شوند. عملکرد تجزیه و تحلیل مؤلفه اصلی در کاهش همبستگی بین باندها در بازیابی LAI نیز مورد بررسی قرار گرفت که تفاوت معناداری در نتایج حاصل‌شده (با اعمال/بدون اعمال تجزیه و تحلیل مؤلفه اصلی) در مجموعه باندی مد نظر مشاهده نشد.

۲-۳- بازیابی مقدار کلروفیل:

باتوجه به نمودار پراکندگی مقدار کلروفیل که مقادیر اندازه‌گیری‌شده در مقابل مقادیر پیش‌بینی‌شده با استفاده از مدل‌های رگرسیون را نشان می‌دهد (شکل (۲-ب))، مدل Rational Quadratic GPR با $RMSE=4.55$ ، $R^2=0.36$ و $MAE=3.61$ به‌عنوان بهترین مدل رگرسیون

۱. Red Edge

در برآورد مقدار کلروفیل منطقه شناخته شد. مدل Matern 5/2 GPR با $RMSE=4.62$ ، $R^2=0.34$ و $MAE=3.70$ و مدل Fine Gaussian SVM با $RMSE=4.70$ ، $R^2=0.32$ و $MAE=3.31$ به ترتیب بهترین برازش را بعد از مدل Rational Quadratic GPR ارائه دادند. باتوجه به شکل ۲-ب، رفتار هر ۳ مدل در تخمین مقادیر کلروفیل تقریباً مشابه یکدیگر است. هر سه مدل در تخمین مقادیر متوسط کلروفیل دقت مناسبی داشته‌اند، اما در تخمین مقادیر بالا و پایین کلروفیل عملکرد مناسبی نداشته‌اند و پراکندگی داده‌ها از خط نرمال در مقادیر حدود بالا و پایین به چشم می‌خورد. باتوجه به نقشه محدوده تغییرات کلروفیل (بر حسب mg/m^3) که در شکل (۳) نشان داده شده است و با میانگین گرفتن از مقادیر کلروفیل برای هر پیکسل در ۸ فولد اعتبارسنجی متقاطع محاسبه شده است، مدل Rational Quadratic GPR عملکرد قابل قبولی در بازیابی مقدار کلروفیل فراهم آورده است؛ به گونه‌ای که در بازیابی مقدار کلروفیل در مزارع آفتابگردان با مقدار کلروفیل $28(mg/m^3)$ بخش‌هایی از کاربری ذرت با مقدار کلروفیل $40(mg/m^3)$ و یونجه با مقدار کلروفیل $35(mg/m^3)$ دقت مطلوبی داشته است. مقدار کلروفیل در مزارع پیاز با دقت نسبتاً کمتری بازیابی شده است و در بخش‌هایی مانند کاربری خاک لخت و مزارع جو، بازیابی به خوبی صورت نگرفته و در نتیجه مدل در تفکیک این کاربری‌ها و در نهایت بازیابی مقدار کلروفیل دقت لازم را نداشته است. مدل Matern 5/2 GPR نیز عملکردی مشابه مدل Rational Quadratic GPR داشته است، با این تفاوت که تنها حد بالا و پایین تغییر کرده است، اما مقدار کلروفیل همانند مدل Rational Quadratic GPR بازیابی شده است. مدل Fine Gaussian SVM نسبت به دو مدل دیگر محدوده تغییرات بیشتری را نشان می‌دهد، اما بازیابی محتوای کلروفیل در اکثر کاربری‌ها یکسان بوده است. برای نمونه می‌توان به کاربری یونجه با تراکم بالا و مقدار کلروفیل $35(mg/m^3)$ و آفتابگردان با تراکم پایین و مقدار کلروفیل $28(mg/m^3)$ اشاره کرد که دقت بازیابی در هر دو کاربری توسط این مدل تقریباً یکسان بوده است که بیان‌کننده ضعف این مدل است، اما با این حال در بازیابی کلروفیل از کاربری ذرت، گندم و جو دوسر دقت مطلوبی را ارائه می‌دهد. نتایج به دست آمده، عملکرد قابل قبول روش رگرسیون گاوسی در بازیابی مقدار کلروفیل را تصدیق می‌کنند. نتایج حاصل از تحلیل حساسیت نشان دادند که مقدار کلروفیل نسبت به محدوده مرئی (۴۰۰-۷۰۰ نانومتر) حساسیت بیشتری نشان می‌دهد. عملکرد تجزیه و تحلیل مؤلفه اصلی نیز مورد بررسی قرار گرفت و در حالتی که مؤلفه اصلی اول ۹۹٪ واریانس موجود در داده‌ها را دربرمی‌گیرد، نتایج بهتری حاصل می‌شود.

۳-۳- بازیابی سطح پوشیده شده با پوشش گیاهی (FVC):

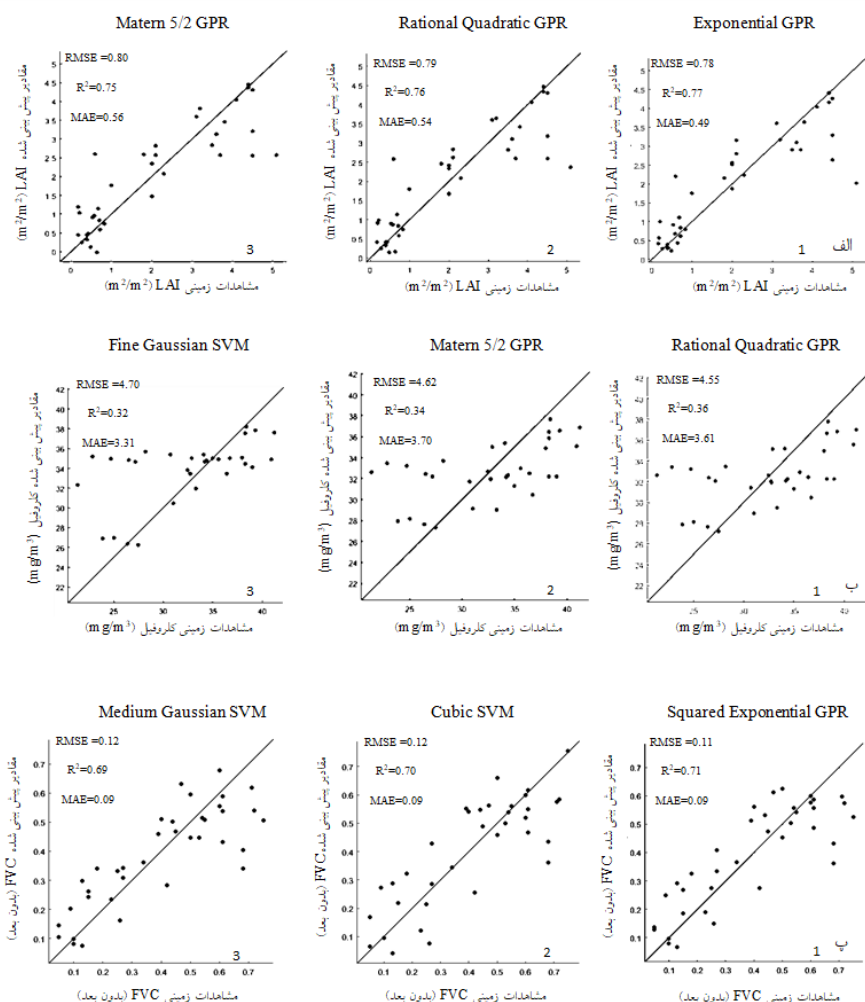
شکل (۲-ج) نمودار پراکندگی FVC را برای مقادیر اندازه گیری شده در مقابل مقادیر پیش بینی شده توسط مدل های رگرسیون را نشان می دهد. همان طور که مشخص است مدل Squared Exponential GPR با $RMSE=0.11$ ، $R^2=0.71$ و $MAE=0.09$ بهترین برآورد از FVC را ارائه می دهد. مدل Cubic SVM با $RMSE=0.12$ ، $R^2=0.70$ و $MAE=0.09$ و مدل Medium Gaussian SVM با $RMSE=0.12$ ، $R^2=0.69$ و $MAE=0.09$ به ترتیب بعد از مدل Squared Exponential GPR بهترین برازش را از شاخص FVC داشتند. باتوجه به شکل ۲-ج، در تخمین مقادیر بالا و پایین، هر سه مدل عملکرد مطلوبی داشته اند و ملاحظه می کنیم که خط نرمال در فاصله مناسبی از نقاط قرار دارد، هر چند در مقادیر بالای FVC شاهد کم تخمینی هستیم. باتوجه به نقشه محدوده تغییرات شاخص FVC هر کدام از مدل های Squared Exponential GPR، Cubic SVM، Medium Gaussian SVM و Exponential GPR که در شکل (۳) نشان داده است و از میانگین گیری مقادیر FVC برای هر پیکسل در ۵ فولد اعتبارسنجی متقاطع محاسبه شده است، عملکرد خوب مدل Squared Exponential GPR کاملاً به چشم می خورد. مقدار FVC کاربری ذرت، بخش هایی از کاربری آفتابگردان، گندم، جو دوسر، یونجه، پیاز و سیر به خوبی با استفاده از این مدل با دقت بالایی بازیابی شده است. مقادیر حد پایین FVC که مرتبط با کاربری هایی مانند خاک لخت و مزارع جو است نیز با دقت مطلوبی بازیابی شده است، اگرچه مقادیر عددی تخمین زده شده برای این دو کاربری بالاتر از حد انتظار است. مدل Cubic SVM اگرچه از لحاظ آماری عملکرد مطلوبی در بازیابی شاخص FVC داشته است، زمانی که نقشه محدوده تغییرات این شاخص را بررسی می کنیم، نه تنها مقادیر حد بالا و پایین FVC به خوبی بازیابی نشده اند، بلکه این مدل در بازیابی FVC از کاربری های مختلف نیز دقت مطلوبی نداشته است. مزارع گندم، جو دوسر، ذرت و بخش هایی از کاربری آفتابگردان با دقت نسبتاً قابل قبولی بازیابی شدند، اما این مدل در بازیابی مقادیر پایین FVC (مانند کاربری خاک لخت و مزارع جو) عملکرد ضعیفی داشته است. مدل Medium Gaussian SVM که از لحاظ آماری سومین مدل مناسب در برآورد FVC بود، تخمین بسیار مناسبی از FVC منطقه داشته و تقریباً شبیه مدل Squared Exponential GPR عمل کرده است، به گونه ای که حتی مقادیر حد پایین را بهتر از مدل Squared Exponential GPR بازیابی کرده است. باتوجه به نقشه مذکور متوجه می شویم که FVC تمامی کاربری ها اعم از ذرت، آفتابگردان، گندم، یونجه، جو دوسر و پیاز به خوبی و با دقت بالا بازیابی شده است. این مدل برخلاف مدل Cubic SVM که در بازیابی FVC کاربری هایی مانند خاک لخت و مزارع جو کارایی مناسبی نداشت، نتایج



قابل قبولی را ارائه می دهد و مقادیر پیش بینی شده به واقعیت زمینی نزدیک تر است. نتایج نشان می دهند که اگرچه روش ماشین بردار پشتیبان دقت مطلوبی در بازیابی این پارامتر داشته است، روش گاوسی عملکرد قابل قبول تری ارائه داده است.

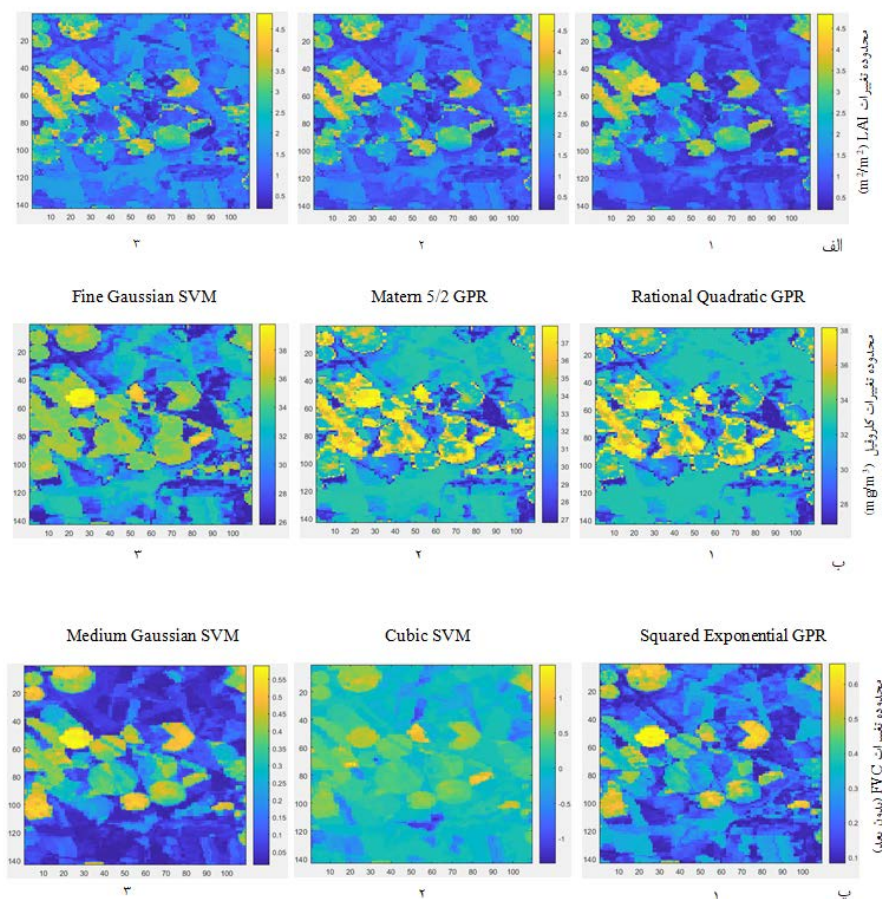
نتایج تحلیل حساسیت نشان داد که این پارامتر بیشترین حساسیت را به محدوده مرئی، لبه قرمز و مادون قرمز نزدیک دارد و همچنین نتایج تجزیه و تحلیل مؤلفه اصلی نشان داد که زمانی که مؤلفه اصلی اول ۹۵٪ واریانس موجود در داده ها را دربرمی گیرد، بهترین نتایج را ارائه می دهد.

بازیابی پارامترهای شاخص سطح برگ مقدار ...
 مهدی نادری



شکل ۲. نمودار پراکندگی پارامترهای بیوفیزیکی و بیوشیمیایی اندازه‌گیری شده در مقابل مقادیر پیش‌بینی شده توسط مدل‌های رگرسیون. الف-نمودار پیش‌بینی شده LAI با مدل (۱) Exponential GPR، (۲) Rational Quadratic GPR و (۳) Matern 5/2 GPR. ب- نمودار پیش‌بینی شده مقدار کلروفیل با مدل (۱) Rational Quadratic GPR، (۲) Matern 5/2 GPR و (۳) Fine Gaussian SVM. پ- نمودار پیش‌بینی شده FVC با مدل (۱) Squared Exponential GPR، (۲) Cubic SVM و (۳) Medium Gaussian SVM.

Rational Quadratic Exponential GPR GPR
 Matern 5/2 GPR



شکل ۳. نقشه محدوده تغییرات پارامترهای بیوفیزیکی و بیوشیمیایی. الف - نقشه محدوده تغییرات LAI با مدل (۱) Exponential GPR، (۲) Rational Quadratic GPR و (۳) Matern 5/2 GPR. ب - نقشه محدوده تغییرات مقدار کلروفیل با مدل (۱) Rational Quadratic GPR، (۲) Matern 5/2 و (۳) Fine Gaussian SVM. پ - نقشه محدوده تغییرات FVC با مدل (۱) Squared Exponential GPR، (۲) Cubic SVM و (۳) Medium Gaussian SVM.

۴- نتیجه‌گیری:

با پیشرفت‌های اخیر سنجش از دور و وجود ماهواره‌های ابرطیفی، داده‌ها و اطلاعات بسیار زیادی از کمیت‌های مکانی و زمانی پوشش‌های گیاهی در دسترس قرار گرفته است. این امر

باعث شده تکنیک‌های پردازش قدرتمند و کارآمدی برای بازیابی متغیرهای بیوفیزیکی و بیوشیمیایی گیاهان به وجود آید. در این زمینه از روش‌های خطی و خصوصاً غیرخطی رویکرد آماری - تجربی به دلیل انعطاف‌پذیری بالا، عدم نیاز به داده‌های اضافی، عدم نیاز به دانش تخصصی زیاد، سرعت و دقتی که دارند، برای بازیابی پارامتر در سنجش از دور استفاده می‌شود (Rivera-Caicedo et al., 2017). در این پژوهش از رویکرد آماری - تجربی به منظور بازیابی پارامترهای LAI، مقدار کلروفیل و FVC در منطقه باراکس اسپانیا استفاده شد. به منظور تهیه نقشه‌های شاخص سطح برگ و سطح پوشیده‌شده با پوشش گیاهی از اعتبارسنجی متقاطع ۵ فولدی و برای مقدار کلروفیل از اعتبارسنجی متقاطع ۸ فولدی برای جلوگیری از بیش برآزش مدل‌ها استفاده شد. پس از آنالیزهای آماری و اخذ خروجی‌های مورد نظر، بهترین مدل به منظور بازیابی هر کدام از پارامترها ارائه شد. در بازیابی LAI، مدل Exponential GPR (RMSE = 0.78, R²=0.77, MAE=0.49)، در بازیابی مقدار کلروفیل، مدل Rational Squared (RMSE = 4.55, R²=0.36, MAE=3.61) و در بازیابی FVC، مدل Exponential GPR (RMSE = 0.11, R²=0.71, MAE=0.09) بهترین برآورد و تخمین را ارائه دادند. مدل Exponential GPR در بازیابی مقادیر پایین LAI (مانند کاربری‌های خاک لخت و مزارع جو) عملکرد مناسبی را نداشته است و در بازیابی مقادیر بالای LAI بهتر عمل کرده است. مدل Rational Quadratic GPR عملکرد قابل‌قبولی در بازیابی مقدار کلروفیل فراهم آورده است، به گونه‌ای که در بازیابی مقدار کلروفیل در مزارع آفتابگردان با مقدار کلروفیل (28 mg/m³)، بخش‌هایی از کاربری ذرت با مقدار کلروفیل (40 mg/m³) و یونجه با مقدار کلروفیل (35 mg/m³) دقت مطلوبی داشته است. مقدار FVC کاربری ذرت، بخش‌هایی از کاربری آفتابگردان، گندم، جو دوسر، یونجه، پیاز و سیر به‌خوبی با استفاده از مدل Squared Exponential GPR با دقت بالایی بازیابی شده است. مقادیر حد پایین FVC که مرتبط با کاربری‌هایی مانند خاک لخت و مزارع جو است نیز با دقت مطلوبی بازیابی شده است، اگرچه مقادیر عددی تخمین‌زده‌شده برای این دو کاربری بالاتر از حد انتظار است. تجزیه و تحلیل نقشه‌های بازیابی‌شده شاخص‌های LAI، مقدار کلروفیل و FVC نشان دادند که مدل‌های فرایند گاوسی که از روش‌های رگرسیون غیرخطی هستند د مقایسه با روش‌های رگرسیون خطی و همچنین روش‌های ماشین بردار پشتیبان در بازیابی پارامترهای مذکور عملکرد بهتری داشته‌اند. بر اساس نتایج (Verrelst et al., 2012a) روش‌های فرایند گاوسی علاوه بر اینکه عملکرد خوبی در داده‌های آموزشی دارند، بلکه منجر به کاهش بیش تخمینی نیز می‌شوند. نقشه‌های بازیابی‌شده نشان می‌دهند که مدل‌های مختلف فرایند گاوسی نه تنها در بازیابی شکل

اراضی کشاورزی موفق عمل کرده، بلکه تغییرات شاخص‌های LAI، مقدار کلروفیل و FVC داخل اراضی کشاورزی را نیز با دقت مطلوبی بازایی کرده‌اند.

نتایج تحلیل حساسیت نشان داد که LAI نسبت به محدوده مرئی، لبه‌قرمز و مادون قرمز نزدیک حساسیت بیشتری نشان می‌دهد. باتوجه به اینکه تغییرات LAI در محدوده مادون قرمز موج کوتاه^۱ نیز اتفاق می‌افتد و سنجنده کریس - پروبا فقط محدوده ۱۰۵۰-۴۰۰ نانومتر را پوشش می‌دهد، لذا پیشنهاد می‌گردد از سنجنده‌هایی استفاده شود که باندهایی در محدوده مادون قرمز موج کوتاه نیز داشته باشند تا دقت تخمین LAI با اضافه شدن این محدوده مورد بررسی قرار گیرد. همچنین مقدار کلروفیل بیشترین حساسیت را به محدوده مرئی (۴۰۰-۷۰۰ نانومتر) و FVC نیز بیشترین حساسیت را به محدوده مرئی، لبه‌قرمز و مادون قرمز نزدیک نشان داد. عملکرد تجزیه و تحلیل مؤلفه اصلی در روند بازایی شاخص‌های مقدار کلروفیل و FVC نیز مورد بررسی قرار گرفت و به ترتیب زمانی که مؤلفه اصلی اول، ۹۹٪ و ۹۵٪ واریانس داده‌ها را دربرمی‌گیرد، نتایج بهتری حاصل می‌شود. گفتنی است که عملکرد تجزیه و تحلیل مؤلفه اصلی در بازایی LAI معنادار نبود. درخصوص استفاده از داده‌های ابرطیفی کریس - پروبا باید عنوان شود اگرچه داده‌های ابرطیفی تکنیک‌های تجزیه و تحلیل طول موج را تسهیل کرده و به‌طور معمول عملکرد بازایی پارامتر را بهبود می‌بخشند، کاهش نسبت سیگنال به نویز از مشکلات بالقوه استفاده از این داده‌هاست. باین‌حال، این داده‌ها پتانسیل خوبی در کاهش عدم قطعیت داده‌ها نشان می‌دهند و در نتیجه باعث بهبود دقت و ثبات در برآورد غیرمستقیم پارامترهای گیاهی می‌شوند. علی‌رغم نتایج قابل قبولی که از تعداد کم داده‌های زمینی در این پژوهش حاصل شد و نقشه‌های بازایی‌شده پارامترهای LAI، مقدار کلروفیل و FVC این موضوع را تصدیق می‌کنند، پیشنهاد می‌شود که در مطالعات بعدی تعداد بیشتری از داده‌های میدانی در فرایند بازایی مورد استفاده قرار بگیرد تا عملکرد مدل‌ها افزایش یابد. علاوه بر این، استفاده از سایر روش‌های رگرسیون نظیر روش‌های توسعه‌یافته شبکه عصبی نیز در مطالعات بعدی پیشنهاد می‌شود تا کارایی و عملکرد آن‌ها در بازایی پارامترهای مختلف مورد مقایسه و بررسی قرار گیرد.

قدردانی و تشکر:

از جناب دکتر علی موسیوند، عضو محترم هیئت علمی گروه سنجش از دور و GIS دانشگاه تربیت مدرس، به دلیل در اختیار گذاشتن داده‌های این پژوهش نهایت سپاس و تشکر را دارم.

۱. Short Wave Infrared

۵- منابع:

- احمدی، رضا؛ امیری بختیار، محمد صادق، «به کارگیری مدل رگرسیون ماشین بردار پشتیبان به منظور تخمین میزان اشباع‌شدگی آب سازند یکی از میدان‌های نفتی بزرگ جنوب غرب ایران»، نشریه پژوهش‌های ژئوفیزیک کاربردی، د ۴. ش ۲. صص ۱۹۹-۲۱۰. ۱۳۹۷. شناسه دیجیتال: 10.22044/JRAG.2017.5552.1106.
- شکری، سعید؛ صادقی، محمدتقی؛ احمدی مروست، مهدی، «ارائه روش ترکیبی پیش‌پردازش داده‌ها در ماشین بردار رگرسیون پیش‌بینی کیفیت گازوییل پالایش شده»، پژوهش نفت، س ۲۳. ش ۷۵. صص ۱۰۲-۱۱۶. ۱۳۹۲. شناسه دیجیتال: 10.22078/PR.2013.317.
- شمس‌الدینی، علی، «قابلیت‌سنجی کارایی داده‌های لیدار و اپتیک به منظور استخراج پارامترهای ساختاری جنگل»، مجله برنامه‌ریزی و آمایش فضا، د ۲۱. ش ۲. صص ۱۱۹-۱۴۵. ۱۳۹۶.
- موسیوند، علی جعفر، «شبیه‌سازی تصاویر ماهواره‌ای ابرطیفی بالای اتمسفر از طریق مدل‌های انتقال تابش»، مجله برنامه‌ریزی و آمایش فضا، د ۲۱. ش ۴. صص ۱۶۱-۱۸۲. ۱۳۹۶.
- مهدوی‌فرد، مصطفی؛ خلیل ولی‌زاده، کامران؛ عطا‌زاده، احسان، «تخمین غلظت کلروفیل-آ با استفاده از داده‌های میدانی و پردازش تصاویر ماهواره‌ای سنتینل-۲ و لندست-۸» (مطالعه موردی: خور تياب)، فصلنامه سنجش از دور و سامانه اطلاعات جغرافیایی در منابع طبیعی، س ۱۱. ش ۱. صص ۷۲-۸۳. ۱۳۹۹.
- رضایی نوایی، سمیرا؛ کوشا، حمیدرضا، «به کارگیری و ارزیابی تکنیک‌های داده‌کاوی جهت پیش‌بینی رویگردانی مشتری در صنعت بیمه»، نشریه بین‌المللی مهندسی صنایع و مدیریت تولید، ش ۴. صص ۶۵۳-۶۳۶. ۱۳۹۵.
- Ahmadi, R., Amiri Bakhtiar, M.S. (2018), « Application of support vector regression to estimate the formation water saturation in one of the largest oil fields located in the southwest of Iran», *JOURNAL OF RESEARCH ON APPLIED GEOPHYSICS*, 4(2), 199-210. doi: 10.22044/JRAG.2017.5552.1106.
- Shokri, S., Sadeghi, M.T., Marvast, M.A. (2013), « An Integrated Method of Data Pre-processing in Support Vector Regression for the Quality Prediction of Treated



Gas-oil», *Journal of Petroleum Research*, 23(75), 102-116.doi:10.22078/PR.2013.317.

– Shamseddini, A. (2017) « LiDAR and optical data capability assessment for plantation structural parameter estimation Assessment of LiDAR and optical data capability in the estimation of structural parameters of plantations », *Journal of Spatial Planning*, 21(2), 119-145.

– Mousivand, A. (2018) «Top of Atmosphere Hyperspectral Image Simulation through Radiative Transfer Models», *Journal of Spatial Planning*, 21(4), 161-182.

– MahdaviFard, M., Khalil ValiZadeh, K., Atazadeh, E. (2020), « Estimation of chlorophyll-a concentration using ground data and Sentinel-2 and Landsat-8 Satellite images processing (Case study: Tiab Estuary)», *Journal of RS and GIS for Natural Resources*, 11(1), 72-83.

– Rezaei Navaei, S., Koosha, H. (2016), « Applying Data Mining Techniques for customer churn prediction in insurance industry», *International journal on industrial Engineering & production Management*, 27(4), 635-653.

– Arneth, A. (2015), «Climate Science: Uncertain Future for Vegetation Cover», *Nature*, 524(7563), 44-45.doi: 10.1038/524044a.

– Baret, F., Buis, S. (2008), «Estimating Canopy Characteristics from Remote Sensing Observation: Review of Methods and Associated Problems», *Advanced in Land Remote Sensing*, 173-201.

– Baret, F., Hagolee, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., Berthelot, B., Nino, Weiss, M., Samain, O., Roujean, J., Leroy, M. (2007), «LAI, FAPAR and fCover Cyclope Global Products Derived from Vegetation Part1: Principals of the algorithm», *Remote Sensing and Environment*, 110, 275-286.

بازیابی پارامترهای شاخص سطح برگ مقدار...
مهدی نادری

- Bei, C., Qian-jun, ZH., Wen-jiang, H., Xiao-yu, S., Hui-chun, Y., Xian-feng, ZH. (2019), «Leaf Chlorophyll Content Retrieval of Wheat by Simulated RapidEye, Sentinel-2 and EnMAP data», *Journal of Integrative Agriculture*, 18(6):1230-1245. doi: 10.1016/S20953119(18)62093-3.
- Darvishzadeh, R., Skidmore, A., Schlerf, M., & Atzberger, C. (2008), « Inversion of a Radiative Transfer Model for Estimating Vegetation LAI and Chlorophyll in a Heterogeneous Grassland», *Remote Sensing of Environment*, 112(5), 2592-2604. doi: 10.1016/j.rse.2007.12.003.
- D'Ursom G., Gomez, S., Vuolo, F., Dini, L. (2009), « Estimation of Land Surface Parameter through Modeling Inversion of Earth Observation Optical Data», *Advances in Modeling Agricultural Systems*, 317-338. doi: 10.1007/978-0-387-75181-8_15.
- Gitelson, A.A., Vina, A., Verma, S.B., Rundquist, D.C., Arkebauer, T.J., Keydan, G., Leavitt, B., Ciganda, V., Burba, G.G., Suyker, A.E. (2006), «Relationship between Gross Primary Production and Chlorophyll Content in Crops: Implication for the Synoptic Monitoring of Vegetation Productivity», *Journal of Geophysical Research*, 111(D8). doi: 10.1029/2005JD006017.
- Huang, J., Zeng, Y., Kuusk, A., Wu, B., Dong, L., Mao, K., Chen, J. (2011a), «Inverting a forest canopy reflectance model to Retrieve the Overstorey and Understory Leaf Area Index for Forest Stands, *International Journal of Remote Sensing*, 32(22), 7591-7611. doi: 10.1080/01431161.2010.525259.
- Jiapear, G., Chen, X., Bao, A. (2011), «A Comparison of Methods for Estimating Fractional Vegetation Cover in Arid Regions», *Agriculture and Forest Meteorology*, 151(12), 1698-1710. doi: 10.1016/j.agrformet.2011.07.004.
- Knyazikhin, Y., Kranigk, J., Myneni, R.B., Panforyov, O., & Gravenhorst, G. (1998), «Influence of Small-scale Structure on Radiative Transfer and



- Photosynthesis in Vegetation Canopies», *Journal of Geophysical Research*, 103(D6), 6133-6144.doi: 10.1029/97JD03380.
- Madhu, B., Paul, A.K., Roy, R. (2019), «Performance Comparison of Various Kernels of Support Vector Machine Regression for Predicting Option Price», *International Journal of Discrete Mathematics*, 4(1): 21-31.doi: 10.11648/j.dmath.20190401.14.
- Marona, R., Martin, R., Yohai, V.J. (2006), « Robust Statistics and Methods », *John Wiley & Sons Ltd*, England.
- Meroni, M., Colombo, R., Panigada, C. (2004), « Inversion of a Radiative Transfer Model With Hyperspectral Observations for LAI Mapping in Poplar Plantations», *Remote Sensing of Environment*, 92(2), 195–206.doi: 10.1016/j.rse.2004.06.005.
- Mu, X., Song, W., Gao, Zh., Mcvicar, T.R., Donohue, R.J., Yan, G. (2018), «Fractional Vegetation Cover Estimation by Using Multi-angle Vegetation Index», *Remote Sensing of Environment*, 216:44-56.doi: 10.1016/j.rse.2018.06.022.
- Monteith, J., Unsworth, M. (1990), «Principals of Environmental Physics», London, UK: Edward Arnold.
- Neinavaz, E., Skidmore, A.K., Darvishzadeh, R., Groen, T.A. (2016), «Retrieval of Leaf Area Index in Different Plant Species Using Thermal Hyperspectral Data, *ISPRS Journal of Photogrammetry and Remote Sensing*, 119:390-401.doi: 10.1016/j.isprsjprs.2016.07.001.
- Pasolli, L., Asam, S., Castelli, M., Bruzzone, L., Wohlfahrt, G., Zebisch, M., Notarnicola, C. (2015), «Retrieval of Leaf Area Index in Mountain Grasslands in the Alps from MODIS Satellite Imagery», *Remote Sensing and Environment*, 165,159-174. doi: 10.1016/j.rse.2015.04.027.

بازیابی پارامترهای شاخص سطح برگ مقدار...
مهدی نادری

- _ Ranwana, R., Palade, V. (2006), «Multi-classifier Systems: Review and a Roadmap for Developers», *International Journal of Hybrid Intelligent Systems*, 3(1), 35–61.doi: 10.3233/HIS-2006-3104.
- _ Rivera-Caicedo , J.P., Verrelst, J., Munoz-Mari, J., Camps-Valls, G., & Moreno, J. (2017). Hyperspectral dimensionality reduction for biophysical variable statistical retrieval, *ISPRS journal of photogrammetry and remote sensing*, 132, 88-101.
- _ Sellers, P.J., Los, S.O., Tucker, C.J., Justice, C.O., Dazlich, D.A., Collatz, G.J., Randall, D.A. (1996), «A Revised Land Surface Parametrization (SiB2) for Atmospheric GCMs: Part II. The Generation of Global Field of Terrestrial Biophysical Parameters from Satellite Data», *Journal of Climate*, 9(4), 706-737.
- _ Smola. A.J., Scholkopf, B. (2004), «A Tutorial on Support Vector Regression», *Statistics and Computing*, 14, 199-222.
- _ Vapnik, Vladimir. (1995), «The nature of statistical learning theory», *Springer*, New York.
- _ Verrelst, J., Alonso, L., Caicedo, J.P.R., Moreno, J., & Camps-Valls, G. (2012a), « Gaussian process retrieval of chlorophyll content from imaging spectroscopy data», *IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing*, 6, 867-874.
- _ Wang, K., Chen, Zh. (2016), «Stepwise Regression and All Possible Subsets Regression in Education», *Electronic International Journal of Education, Arts and Science*, 2(1), 60-81.
- _ Zhang, D., Mansaray, L.R., Jin, H., Sun, H., Kuang, Zh., Huang, J. (2018), «A Universal Estimation Model of Fractional Vegetation Cover for Different Crops based on Time Series Digital Photographs», *Computers and Electronics in Agriculture*, 151, 93-103.doi: 10.1016/j.compag.2018.05.030.



– Zhu, X., Feng, X., Zhao, Y. (2011), «Multi-scale MSDT inversion based on LAI spatial knowledge», *Science China Earth Sciences*, 55(8), 246–255.doi: 10.1007/s11430-011- 4312-0.