



سنجش ازدور و سامانه اطلاعات جغرافیایی در منابع طبیعی (سال یازدهم / شماره دوم) تابستان ۱۳۹۹

نمایه شده در سایت: پایگاه استنادی علوم جهان اسلام، جهاد دانشگاهی، مگ ایران، نورمگز، سیولیکا، گوگل اسکولار
آدرس وب سایت: <http://girs.iaubushehr.ac.ir>



بهره‌گیری از رویکرد بوت‌استرپ برای مقایسه روش‌های مدل‌سازی آماری در برآورد سنجش ازدوری زی‌توده روی زمینی جنگل‌های زاگرس

امیر صفری^۱، هرمز سهرابی

دریافت: ۸ اردیبهشت ۱۳۹۸ / پذیرش: ۳۰ اردیبهشت ۱۳۹۹

دسترسی اینترنتی: ۱۵ تیر ۱۳۹۹

چکیده

۳) ارزیابی عملکرد و مقایسه نتایج آن‌ها با استفاده از آزمون نکویی برازش و روش‌های ارزیابی متقابل انجام گرفته است. با توجه به نقش انتخاب روش مدل‌سازی مناسب، مطالعه حاضر باهدف بررسی نه روش مدل‌سازی شامل رگرسیون خطی، مدل تجمعی تعمیم‌یافته، جنگل تصادفی، ماشین‌بردار پشتیبان، درختان رگرسیون تقویت‌شده، k-AMIN نزدیک‌ترین همسایه، مدل پردازش گوسی و رگرسیون اسپلاین تطبیقی چندگانه با استفاده از فرآیند بوت‌استرپ برای برآورد زی‌توده روی زمینی جنگل‌های زاگرس با استفاده از تصویرهای لندست ۸ انجام شد.

مواد و روش‌ها مطالعه حاضر در جنگل‌های کرمانشاه که در غرب ایران در کوه‌های زاگرس واقع شده است با گونه درختی غالب بلوط انجام شد. جنگل‌های زاگرس اغلب به صورت تنک و باز بوده و حدود ۲۰ درصد از سطح ایران و ۴۰ درصد از مناطق جنگلی این کشور را شامل می‌شود. برای اجرای این مطالعه دو منطقه جنگلی با شدت دخالت‌های انسانی مختلف انتخاب شد: منطقه سرفیروزآباد با شدت تخریب بالا و منطقه جنگلی گهواره با حداقل تخریب. مختصات جغرافیایی این مناطق به ترتیب برای سرفیروزآباد و گهواره عبارتند از ۳۳ درجه و ۵۷ دقیقه تا ۳۴ درجه و ۴ دقیقه شمالی و ۴۷ درجه و ۳ دقیقه تا ۴۷ درجه و ۱۷ دقیقه شرقی، و ۳۴ درجه و ۲۱ دقیقه تا ۳۴ درجه و ۲۴ دقیقه شمالی و ۴۶ درجه و ۱۶ دقیقه تا ۴۶ درجه و ۲۳ دقیقه شرقی است.

پیشینه و هدف با توجه به اهمیت اکوسیستم‌های جنگل‌ها در پروژه‌های تعدیل تغییر اقلیم، روش‌های دقیق و کم‌هزینه برای برآورد زی‌توده روی زمینی موردنیاز است. روش‌های معمول مورد استفاده برای برآورد زی‌توده روی زمینی شامل اندازه‌گیری در عرصه، محاسبه زی‌توده با استفاده از معادلات آلومتری و استفاده از فنون سنجش از دوری است. سنجش‌ازدور به‌طور گسترده برای برآورد زی‌توده جنگل‌ها در دهه‌های اخیر مورد استفاده قرار گرفته است. یکی از عوامل بسیار مهم در استفاده از داده‌های سنجش‌ازدوری برای برآورد زی‌توده روی زمینی روش مدل‌سازی آماری مورد استفاده است. تعداد زیادی تحقیق در زمینه استفاده از روش‌های مدل‌سازی انجام شده است. به‌رحال، این مطالعات با چالش‌های مختلفی روبرو هستند که برای مثال می‌توان به موارد زیر اشاره کرد: ۱) هیچ روش مدل‌سازی به‌عنوان بهترین روش معرفی نشده است، ۲) عملکرد این روش‌ها متأثر از نوع جنگل، ساختار جنگل و میزان آشفستگی‌های موجود است.

امیر صفری^۱، هرمز سهرابی^۲ (✉)

۱. دکتری جنگلداری، دانشکده منابع طبیعی و علوم دریایی نور، دانشگاه تربیت مدرس، تهران، ایران

۲. دانشیار گروه جنگلداری، دانشکده منابع طبیعی و علوم دریایی نور، دانشگاه تربیت مدرس، تهران، ایران

پست الکترونیکی مسئول مکاتبات: hsohrabi@modares.ac.ir

نتایج و بحث مشخصه‌های آماری اندازه‌گیری شده از قطعه نمونه‌های زمینی نشان داد که میانگین زی‌توده روی زمینی برای منطقه‌های سرفیروزآباد و گهواره به ترتیب ۱۲/۶ و ۲۰/۵ تن در هکتار است. تجزیه واریانس نشان داد که اختلاف معنی داری در سطح ۰/۰۱ برای مشخصه‌های ضریب تبیین و جذر میانگین مربعات خطای محاسبه شده برای ۱۰۰۰ مرتبه تکرار توسط ۱۰-fold وجود دارد. روش مدل‌سازی کویست در منطقه سرفیروزآباد با میانگین ضریب تبیین محاسبه شده ۰/۶۱ نتایج بهتری نسبت به سایر روش‌ها ارائه کرد. این نتایج برای منطقه گهواره بیانگر کارایی بهتر روش‌های رگرسیون خطی، مدل جمعی تعمیم‌یافته و k-امین نزدیکترین همسایه با میانگین ضریب تبیین ۰/۸۷ بود. مقایسه‌های روش‌های مختلف مدل‌سازی در مورد جذر میانگین مربعات خطای برآوردها با استفاده از آزمون تکی Tukey نشان داد که در منطقه سرفیروزآباد روش کویست با مقدار میانگین ۳/۳ تن در هکتار و روش‌های k-امین نزدیکترین همسایه و جنگل تصادفی با میانگین ۵/۸ تن در هکتار اختلاف معنی داری با سایر روش‌ها داشت. به‌طور کلی، نتایج این تحقیق بیانگر کارایی مناسب تصویر لندست ۸ در برآورد زی‌توده روی زمینی جنگل‌های زاگرس است. این موضوع می‌تواند به دلیل مقدار کم زی‌توده در منطقه‌های مورد بررسی و نرسیدن آنها به نقطه اشباع به‌عنوان یکی از چالش‌های استفاده از تصویرهای نوری مانند لندست باشد. نتایج دیگر این تحقیق ارزیابی اثر روش مدل‌سازی برای افزایش صحت برآورد سنجش‌زدوری زی‌توده روی زمینی است. برخلاف نتایج پژوهش‌های قبلی، رگرسیون خطی نتایج بهتری نسبت به روش‌های ناپارامتری ارائه کرد که دلیل آن می‌تواند وجود رابطه خطی زی‌توده روی زمینی و متغیرهای طیفی استخراج شده از تصویرهای لندست باشد. در بین متغیرهای طیفی مختلفی که استفاده شد مقدارهای باند قرمز، مادون قرمز نزدیک و مادون قرمز با طول موج کوتاه ۱ و ۲ به عنوان متغیر نهایی در اغلب روش‌های مدل‌سازی استفاده شدند.

نتیجه‌گیری در این تحقیق، کارایی روش‌های مختلف مدل‌سازی را در برآورد زی‌توده روی زمینی با استفاده از تصویرهای لندست مورد بررسی قرار گرفت. برآوردهای زی‌توده با استفاده از ۹ روش پارامتری، نیمه‌پارامتری و ناپارامتری و با استفاده از ارزیابی متقابل ۱۰-fold و با ۱۰۰۰ مرتبه تکرار مورد مقایسه قرار گرفتند. نتایج بیانگر قابلیت خوب تصویرهای لندست برای برآورد زی‌توده روی زمینی جنگل‌های بلوط زاگرس با هزینه کم است. برآورد زی‌توده در منطقه جنگلی گهواره با دست‌خوردگی کمتر صحت بالاتری نسبت به منطقه جنگلی سرفیروزآباد با تخریب بالاتر داشت.

واژه‌های کلیدی: زی‌توده روی زمینی، بوت‌استرپ، مدل‌سازی

آمار، جنگل‌های زاگرس، لندست

برای جمع‌آوری نمونه‌های زمینی در دو منطقه مورد مطالعه، از نقشه شاخص سطح برگ تهیه شده از تصویرهای لندست براساس مدل جهانی استفاده شد. هر دو منطقه به سه طبقه کم، متوسط و زیاد شاخص سطح برگ طبقه بندی شده و در هر طبقه موقعیت مکانی نمونه‌ها با استفاده از آماربرداری منظم در قالب یک شبکه ۲۰۰×۲۰۰ متر مشخص گردید. نمونه‌های زمینی به تعداد ۱۲۴ نمونه مربعی شکل (۶۳ نمونه در منطقه گهواره و ۶۱ نمونه در منطقه سرفیروزآباد) با ابعاد ۳۰×۳۰ متر به اندازه پیکسل‌های تصویر لندست مورد استفاده برداشت شد. مقدار زی‌توده روی زمینی هر درخت یا جست گروه توسط معادله آلومتری پیشنهاد شده برای گونه‌های بلوط جنگل‌های زاگرس محاسبه شد. معادله آلومتری مورد استفاده برای برآورد اندوخته زی‌توده هر درخت یا جست گروه از دو قطر عمود بر هم تاج درختان استفاده می‌کند. از جمع اندوخته زی‌توده تک تک درختان موجود در قطعه نمونه، مقدار قطعه زی‌توده در سطح قطعه نمونه بر حسب تن در هکتار محاسبه شد. منطقه‌های مورد مطالعه در یک فریم از تصویرهای لندست ۸ به شماره گذر ۱۶۷ و شماره ردیف ۳۶ قرار داشتند. یک تصویر لندست بدون پوشش ابرناکی مربوط به تاریخ ۱۹ مرداد سال ۱۳۹۴ (۱۰ آگوست ۲۰۱۵) مربوط به زمانی که تاج درختان کاملاً بسته بوده و نزدیک به زمان آماربرداری زمینی از سایت USGS دانلود شد. پیش‌پردازش تصویر مورد استفاده شامل تصحیحات رادیومتری و توپوگرافی به روش C اعمال شد که در مطالعات قبلی پیشنهاد شده است. برای برآورد زی‌توده روی زمینی در منطقه‌های مورد مطالعه با استفاده از سنجش‌زدوری، تعداد ۳۸ متغیر طیفی شامل مقدارهای بانندی، نسبت‌های ساده بانندی، شاخص‌های گیاهی و تبدیل‌های خطی رایج مانند تسلدکپ و آنالیز مولفه‌های اصلی از تصویر لندست ۸ مورد استفاده استخراج شدند.

به‌طور کلی کارایی ۹ روش مدل‌سازی آماری مختلف شامل روش‌های پارامتری (رگرسیون خطی)، نیمه‌پارامتری (مدل جمعی تعمیم‌یافته) و ناپارامتری (جنگل تصادفی، ماشین‌بردار پشتیبان، k-امین نزدیکترین همسایه، درختان رگرسیون تقویت‌شده، رگرسیون اسپلاین تطبیقی چندمتغیره، کویست و مدل پردازش گوسی) برای برآورد زی‌توده روی زمینی مورد مقایسه قرار گرفت. ارزیابی مدل‌ها با استفاده از روش ۱۰-fold و با ۱۰۰۰ مرتبه تکرار و با محاسبه دو آماره ضریب تبیین و جذر میانگین مربعات خطا انجام گرفت. این تعداد اطمینان از پایداری نتایج را تأمین می‌کند.

مقدمه

برآورد دقیق انتشار گازهای گلخانه‌ای و پایش جزئیات چرخه کربن برای متعادل کردن تغییرات اقلیمی ضروری است. (۱۶ و ۳۶). تغییرات اقلیمی و اثرات زیان‌بار آن بر حیات زیست‌کره زمین سبب شده تا کاهش اثرات تغییرات اقلیمی به‌عنوان یک موضوع مهم موردتوجه دانشمندان و سیاست‌گذاران علوم مختلف قرار گرفته است (۴ و ۱۷). یکی از مهم‌ترین راهکارهای معرفی شده برای مقابله با تغییرات اقلیمی کاهش انتشار و افزایش ترسیب کربن است (۱ و ۴). جنگل‌ها با ذخیره ۸۰ درصد کربن روی زمینی و ۴۰ درصد کربن زیرزمینی به‌عنوان یک منبع حیاتی جذب‌کننده کربن معرفی شده است که می‌تواند سالانه تقریباً ۳۳ درصد از انتشار کربن ناشی از سوخت‌های فسیلی و تغییرات کاربری اراضی را ترسیب کنند (۲۳، ۲۵ و ۲۷). با توجه به اینکه ۵۰ درصد زی‌توده درختان کربن است (۱۱)، بنابراین، برآورد زی‌توده درختان برای کمی کردن اندوخته کربن و مدل‌سازی تغییرات اقلیم به‌عنوان یک ابزار کلیدی معرفی شده است (۳۲).

اندازه‌گیری زی‌توده با استفاده از نمونه‌برداری زمینی با مشکلات مختلفی مانند صرف زمان و هزینه زیاد، محدودیت در مقیاس‌های وسیع و مشکلات اجرایی متعدد به‌ویژه در مناطق صعب‌العبور و دور از دسترس مواجه است (۳۲ و ۳۵). در مقابل، استفاده از مدل‌های شبیه‌سازی نیازمند اطلاعات کامل و طولانی‌مدت است که استفاده از آن‌ها را در کشورهای درحال‌توسعه با چالش اساسی روبرو می‌کند. ازاین‌رو توسعه روش‌های ساده، سریع، کم‌هزینه و دقیق برای برآورد زی‌توده جنگل در مقیاس‌های مختلف یک اقدام حیاتی و ضروری است (۱۰ و ۲۰). خوشبختانه با توسعه فنون و داده‌های سنجش‌ازدوری و ترکیب آن‌ها با اطلاعات جمع‌آوری شده از نمونه‌برداری زمینی امکان برآورد زی‌توده جنگل‌ها را با حداقل عملیات میدانی و هزینه فراهم آورده است (۵ و ۳۲). مطالعات مختلفی برای برآورد زی‌توده جنگل از مقیاس‌های محلی تا جهانی با استفاده از داده‌های مختلف سنجش‌ازدوری مانند داده فعال (لیدار و رادار) و داده‌های غیرفعال (انواع

تصویرهای نوری) انجام شده است. اغلب این مطالعات در جنگل‌های متراکم با میانگین بالای زی‌توده روی زمینی مناطق گرمسیری و مرطوب انجام شده است (۱۸). بااین‌حال، تحقیقات بسیار اندکی در این رابطه در جنگل‌های مناطق خشک و نیمه‌خشک با میانگین اندوخته کم زی‌توده انجام گرفته است. در چنین مناطقی به دلیل تراکم پایین پوشش گیاهی، انعکاس طیفی ثبت شده توسط سنجنده‌های نوری به‌شدت تحت تأثیر بازتاب آمیخته خاک، پوشش گیاهی و سایه قرار می‌گیرند (۸ و ۱۰). یکی از بارزترین چنین منطقی جنگل‌های زاگرس است که با وسعت حدود ۶ میلیون هکتار می‌تواند نقش مهمی در تعدیل تغییرات اقلیمی و فراهم‌کننده آب در مقیاس محلی تا جهانی باشد (۳۱). از آنجاکه جنگل‌های زاگرس تولیدکننده چوب صنعتی به‌عنوان اقتصادی‌ترین محصول جنگل‌ها نیستند، بودجه کمی برای آماربرداری و پایش آن‌ها اختصاص داده می‌شود. بنابراین، توسعه روش‌های سنجش‌ازدوری با حداقل هزینه برای برآورد زی‌توده این جنگل‌ها لازم و ضروری است. تصویرهای لندست به دلیل داشتن آرشیو رایگان و قابل‌دسترس برای عموم و پوشش زمانی و مکانی مناسب می‌تواند به‌عنوان ابزاری مفید برای برآورد و پایش زی‌توده موردتوجه قرار گیرد. دقت برآورد زی‌توده با استفاده از داده‌های سنجش‌ازدوری وابسته به نکات مختلفی است که از مهم‌ترین آن‌ها می‌توان به انتخاب روش مدل‌سازی مناسب اشاره کرد. تحقیق فسنناچت و همکاران (۹) اظهار کردند که روش مدل‌سازی اهمیتی برابر با نوع داده سنجش‌ازدوری دارد. روش‌های مدل‌سازی مختلفی برای برآورد سنجش‌ازدوری مشخصه‌های کمی جنگل از جمله زی‌توده مورد استفاده قرار گرفته که در دودسته کلی مدل‌های پارامتری (رگرسیون‌های خطی) و مدل‌های ناپارامتری (ماشین‌بردارپشتیبان، جنگل تصادفی و...) تقسیم‌بندی می‌شوند که در دهه‌های اخیر توسعه یافته‌اند (۲۴ و ۳۰). مرور منابع بیانگر این موضوع است که روش‌های مدل‌سازی در مطالعات مختلف نتایج متفاوتی ارائه کرده‌اند و نمی‌توان یک روش را به‌عنوان بهترین روش مدل‌سازی معرفی کرد. مطالعات گورگن و همکاران (۱۵)، کاستیلو و همکاران (۳) و دومینگو و همکاران (۵) از کارایی مناسب روش‌های رگرسیون خطی در

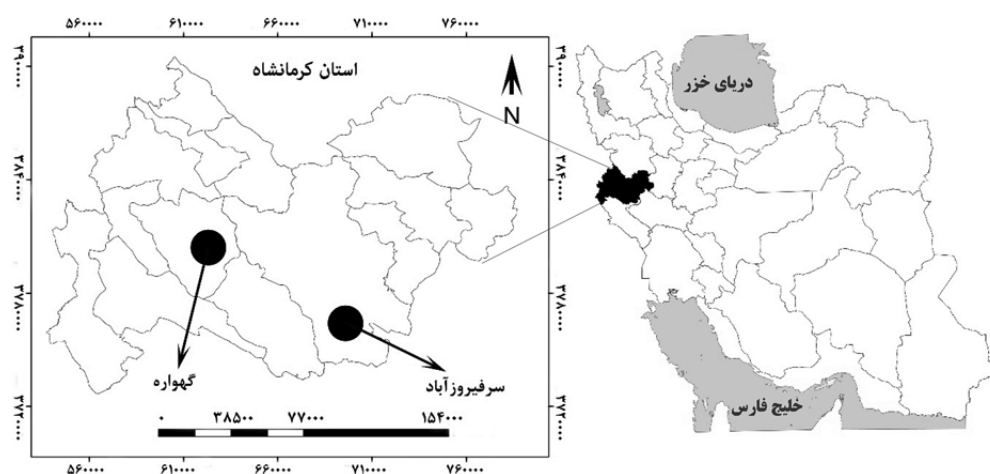
جنگل‌های زاگرس است.

روش تحقیق

منطقه مورد مطالعه

تحقیق حاضر در بخشی از جنگل‌های زاگرس، در استان کرمانشاه و دو منطقه جنگلی گهواره در شهرستان دالاهو با طول ۶۱۶۷۹۵ تا ۶۲۸۰۷۷ و عرض ۳۷۹۸۴۳۳ تا ۳۸۰۸۷۰۲ جنگل‌های سرفیروزآباد در شهرستان کرمانشاه با طول ۶۸۶۳۴۸ تا ۷۱۳۶۹۵ و عرض ۳۷۵۴۳۶۳ تا ۳۷۷۰۹۵۱ در سیستم مختصات UTM در زون ۳۸ انجام گرفته است (شکل ۱). جنگل‌های گهواره و سرفیروزآباد به ترتیب دارای اقلیم‌های نیمه مرطوب و نیمه خشک هستند. گهواره در دامنه ارتفاعی ۱۴۶۲ تا ۲۰۹۷ متر از سطح آزاد دریا و سرفیروزآباد در دامنه ۱۲۳۰ تا ۲۱۲۷ متر از سطح دریا قرار دارند. فاصله دو منطقه از هم حدود ۷۰ کیلومتر بوده و ساختار هر دو منطقه جنگلی به صورت شاخه‌زاد است.

برآورد سنجش از دوری زی توده یادکرده‌اند، درحالی‌که، تعدادی از پژوهش‌های دیگر نشان داده‌اند که روش‌های ناپارامتری کاربرد بهتری نسبت به روش‌های پارامتری دارند (۲۱، ۲۶ و ۳۳). جمع‌بندی مرور منابع نشان می‌دهد استفاده از روش‌های مدل‌سازی با سه چالش عمده روبرو است؛ هیچ روشی تاکنون به‌عنوان بهترین معرفی نشده است، عملکرد این روش‌ها تحت تأثیر نوع و ساختار جنگل دارد، اغلب مطالعات صحت روش‌های مدل‌سازی فقط با استفاده از روش ارزیابی متقابل سنجیده شده است. به‌رحال چالش آخر سبب می‌شود تا استفاده از این مدل‌ها در تعداد داده‌های زمینی کم که اغلب پژوهش‌ها با آن روبرو است، نتواند به‌طور کامل توان و پایداری نتایج روش‌های مدل‌سازی را منعکس کند. یکی از رویکردهای مفید و قابل قبول برای مقایسه روش‌های مدل‌سازی استفاده از نمونه‌برداری مجدد مانند فرآیند بوت‌استرپ است (۱۲ و ۲۲). هدف اصلی مطالعه حاضر، مقایسه روش‌های مختلف مدل‌سازی برای برآورد کم‌هزینه زی توده روی زمینی



شکل ۱. منطقه‌های مورد مطالعه

Fig 1. Study area

به صورت شاخه‌زاد و به صورت غالب از گونه بلوط ایرانی (*Quercus brantii* Lindl) تشکیل شده‌اند؛ بنابراین، برای نمونه‌بندی جنگل‌های مورد بررسی از نقشه شاخص سطح برگ (LAI: Leaf Area Index) استفاده شد، زیرا استفاده از این شاخص کمک می‌کند تا قطعات نمونه در کل ساختار تاج

نمونه‌گیری داده زمینی

مونه‌بندی جنگل سبب پراکنش متناسب قطعات نمونه در شرایط مختلف توده‌های جنگلی برای جمع‌آوری اطلاعات زمینی شده و در نهایت دقت برآورد ویژگی‌های جنگل را بهبود بخشد (۲۱). با توجه به اینکه، ساختار جنگل در هر دو منطقه

سال ۲۰۱۵ برابر با ۱۹ مرداد سال ۱۳۹۴ مربوط به زمانی که تاج درختان کاملاً بسته است و نزدیک به زمان آماربرداری از وب‌سایت USGS (earthexplorer.usgs.gov) دریافت شد. برای پیش‌پردازش تصویر لندست ۸ (OLI) طبق مطالعه صفری و همکاران (۳۰) تصحیحات رادیومتری و توپوگرافی استفاده شد. تصحیح رادیومتری با استفاده از مشخصات حس‌گر اخذ کننده تصویر انجام شده که سبب تبدیل تصویر به انعکاس بالای اتمسفر شد. تصحیح توپوگرافی با استفاده از روش C انجام شد که دارای کارایی مناسب برای تصاویر لندست است (۲۸)، جهت اجرای تصحیح توپوگرافی از مدل رقومی زمین استخراج شده از تصاویر ASTER استفاده شد. تصحیح‌های رادیومتری و توپوگرافی در نرم‌افزار منبع باز R و با استفاده از بسته‌های landsat و RStoolbox انجام شد.

متغیرهای طیفی استخراج شده

تعداد زیادی از انواع متغیرهای طیفی از تصاویر لندست ۸ استخراج شد که شامل باندهای اصلی، نسبت‌گیری‌های باندهای ساده، شاخص‌های گیاهی، تبدیل‌های ساده مانند تسلدکپ و تجزیه مؤلفه‌های اصلی بود (جدول ۱). در مجموع ۳۸ متغیر طیفی از تصویر لندست ۸ مورد استفاده تهیه (شکل ۲) و مقدارهای طیفی متناظر با قطعات نمونه زمینی برداشت شده استخراج شد (۱۳، ۱۸، ۲۶، ۳۰ و ۳۷).

جنگل موردبررسی پراکنده شوند. نقشه LAI با استفاده از تصاویر لندست ۸ و یک مدل جهانی ارائه شده برای این کار تهیه شد (۲۵). پس از تهیه نقشه LAI، هر دو جنگل موردبررسی به سه طبقه (با تراکم کم، متوسط و بالا) با مساحت‌های تقریباً برابر تقسیم و در هر طبقه چند پلی‌گون انتخاب شد. پس از انتخاب پلی‌گون‌ها، با استفاده از جنگل‌گردی در نهایت در هر طبقه یک چند پلی‌گون برای نمونه‌برداری تعیین شد. نمونه‌برداری در هر طبقه با استفاده از طرح منظم با شروع نقطه تصادفی و در غالب یک شبکه آماربرداری ۲۰۰ × ۲۰۰ متر انجام گرفت. تعداد ۱۲۴ قطعه‌نمونه مربعی شکل به ابعاد ۳۰ × ۳۰ متر مشابه با ابعاد پیکسل لندست در مردادماه سال ۱۳۹۵ برداشت شد (۶۳ نمونه در جنگل گهواره و ۶۱ نمونه در جنگل سرفیروزآباد). پس از تعیین موقعیت مکانی نمونه‌ها با استفاده از دستگاه موقعیت‌یاب جهانی (GPS) دستی، اندازه‌گیری اندوخته زی‌توده تک‌تک درختان با اندازه‌گیری تاج پوشش آن‌ها و استفاده از روابط آلومتری توسعه داده شده برای جنگل‌های زاگرس (۱۷) محاسبه شد، که در نهایت با جمع مقدار زی‌توده تمامی درختان موجود در قطعه‌نمونه میزان زی‌توده قطعه‌نمونه برحسب تن در هکتار محاسبه شد.

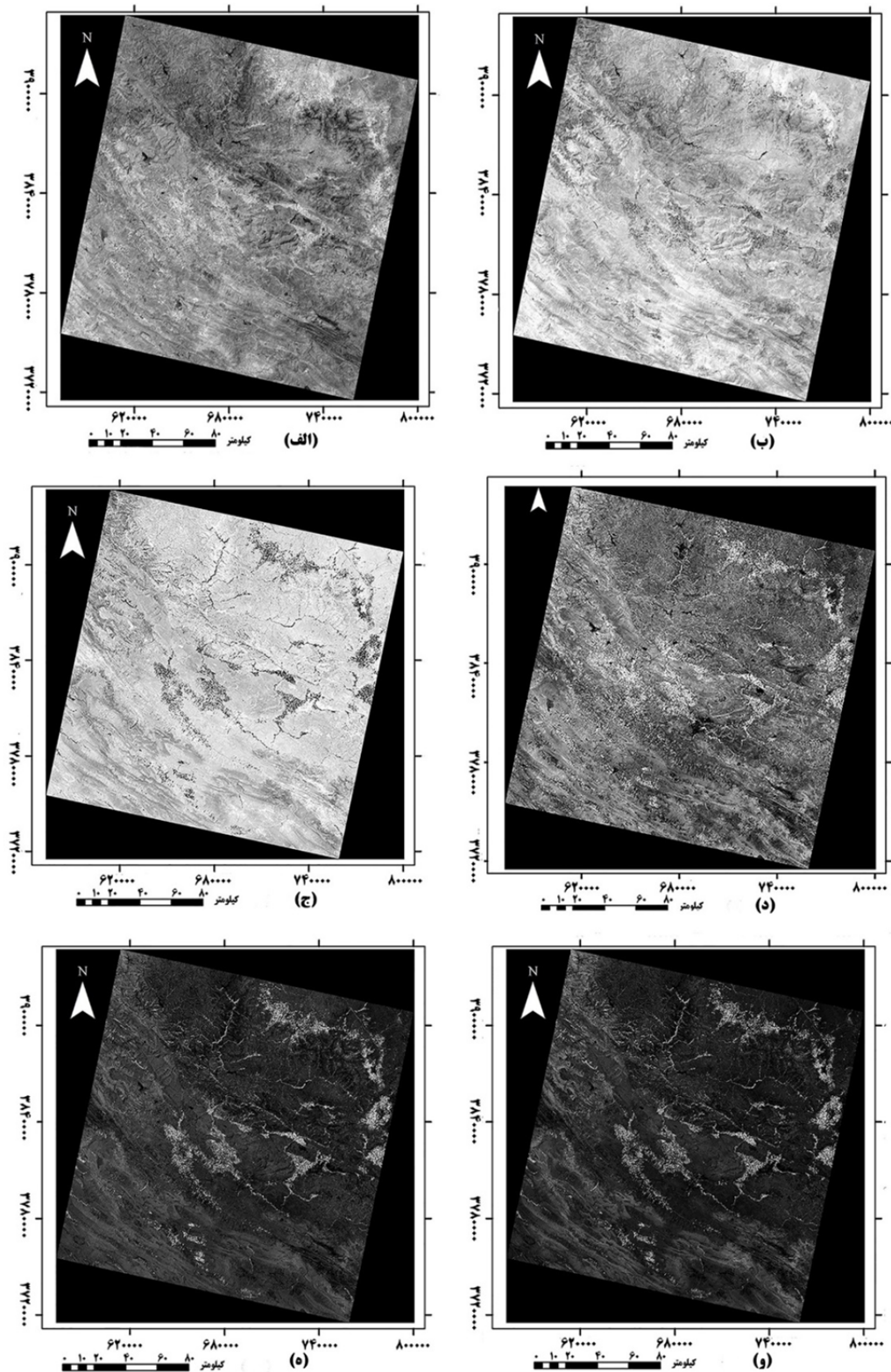
تصاویر ماهواره‌ای

در این مطالعه یک فریم از تصاویر لندست ۸ (OLI)، به شماره گذر ۱۶۷ و ردیف ۳۶ مربوط به تاریخ دهم آگوست

جدول ۱. متغیرهای طیفی استخراج شده از تصویر لندست ۸

Table 1. Landsat 8 derived spectral variables

ردیف	متغیر طیفی	شرح
۱	باندهای طیفی	آبی (B)، سبز (G)، قرمز (R)، مادون قرمز (NIR)، مادون قرمز با طول موج کوتاه اول (SWIR1) و مادون قرمز با طول موج کوتاه دوم (SWIR2)
۲	نسبت‌های باندهای ساده	B/G, B/R, B/N, B/S1, B/S2, G/R, G/N, G/S1, G/S2, N/R, N/S1, N/S2, S1/S2, R/S1, R/S2
۳	شاخص‌های گیاهی	NDVI, DVI, MSAVI2, SAVI, NDVIC, SATVI
۴	تسلدکپ و تجزیه مؤلفه‌های اصلی	PC1, PC2, PC3, PC4, PC5, PC6, BRIGHTNESS, GREENESS, WETNESS, TCA TCD



شکل ۲. متغیرهای به دست آمده از لندست ۸ شامل؛ (الف) مادون قرمز نزدیک، (ب) مادون قرمز نزدیک با طول موج کوتاه ۲، (ج) نسبت باند قرمز به مادون قرمز نزدیک، (د) نسبت مادون قرمز نزدیک با طول موج کوتاه ۱ به ۲، (ه) شاخص گیاهی DVI و (و) شاخص گیاهی NDVI
 Fig 2. Landsat 8 derived variables such as; (A) near infrared, (B) shortwave infrared 2, (C) ratio red to near infrared, (D) ratio shortwave infrared 1 to shortwave infrared 2, (E) difference vegetataion index (DVI) and (F) normalized difference vegetation index (NDVI)

روش‌های مدل‌سازی

به‌طور کلی ۹ روش مدل‌سازی شامل روش‌های پارامتری (رگرسیون خطی)، نیمه‌پارامتری (مدل جمعی تعمیم‌یافته) و ناپارامتری (جنگل تصادفی، ماشین‌بردار پشتیبان، k-AMIN نزدیک‌ترین همسایه، درختان رگرسیون تقویت‌شده، رگرسیون اسپلاین تطبیقی چندمتغیره، کوپیست و مدل پردازش گوسی) برای برآورد زی‌توده روی زمینی مورد مقایسه قرار گرفت (۱۲ و ۲۹).

ارزیابی عملکرد روش‌های مدل‌سازی

برای ارزیابی صحت روش‌های مدل‌سازی مورد استفاده از فرآیند بوت‌استرپ و ارزیابی متقابل fold-۱۰ با ۱۰۰۰ مرتبه تکرار انجام شد. این تعداد تکرار کمک می‌کند تا یک ارزیابی قابل قبول از اجرای مدل‌ها ارائه گردد. ضریب تبیین (R^2) و ریشه میانگین مربعات خطای برآوردها (RMSPE) محاسبه شده از هر تکرار مدل‌سازی، به‌عنوان دو شاخص آماری عمومی برای مقایسه مدل‌ها محاسبه شد (رابطه‌های ۱ و ۲).

$$A = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{\sum_{i=1}^n (y_i - \bar{y}_i)} \quad [1]$$

$$RMSPE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad [2]$$

در این رابطه؛ R^2 ضریب تبیین، y_i مقدار مشاهده‌شده، \hat{y}_i مقدار برآورد شده، \bar{y} مقدار میانگین مشاهدات، RMSPE ریشه میانگین مربعات خطای برآوردها و n تعداد مشاهدات. در نهایت، وجود اختلاف معنی‌دار بین روش‌های مورد بررسی از نظر آماری R^2 و RMSPE توسط تجزیه واریانس یک‌طرفه و مقایسه میانگین‌ها توسط آزمون آماری Tukey مورد بررسی قرار گرفت. تحلیل‌های آماری این تحقیق با استفاده از نرم‌افزار R انجام شد.

نتایج

خلاصه مشخصه‌های آماری محاسبه‌شده برای زی‌توده درختان منطقه مورد بررسی با استفاده از اطلاعات جمع‌آوری‌شده از ۱۲۴ قطعه نمونه برداشت‌شده در جدول (۲) ارائه شده است. منطقه جنگلی گهواره دارای میانگین زی‌توده بالاتری (۲۰/۵ تن در هکتار) در مقایسه با منطقه جنگلی سرفیروزآباد (۱۲/۶ تن در هکتار) دارد. میانگین زی‌توده در منطقه جنگلی سرفیروزآباد ۳۹ درصد کمتر از میانگین زی‌توده در منطقه گهواره بود.

جدول ۲. خلاصه مشخصه‌های آماری زی‌توده روی زمینی در منطقه‌های مورد بررسی بر حسب تن در هکتار

Table 2. Summary of descriptive statistics of AGB in study areas (unit: ton/ha)

انحراف	کمینه	بیشینه	میانگین	تعداد	منطقه
۱۱/۹	۰	۵۳/۰	۲۰/۵	۶۳	گهواره
۵/۰	۴/۷	۲۰/۵	۱۲/۶	۶۱	سرفیروزآباد
۹/۹	۰	۵۳/۰	۱۶/۶	۱۲۴	مجموع دو

بار تکرار وجود دارد ($p\text{-value} < ۰/۰۰۱$). در سرفیروزآباد از نظر ضریب تبیین، روش‌های مدل‌سازی اختلاف معنی‌داری در سطح ۰/۰۵ داشته و روش مدل‌سازی کوپیست با کسب ضریب تبیین ۰/۶۱ نسبت به سایر روش‌ها عملکرد بهتری ارائه کرده است. برای منطقه گهواره، مقدارهای ضریب تبیین

مقایسه روش‌های مدل‌سازی با استفاده از روش fold-۱۰ با ۱۰۰۰ مرتبه تکرار

نتایج تجزیه واریانس نشان داد اختلاف معنی‌داری بین روش‌های مدل‌سازی برای هر دو شاخص آماری ضریب تبیین و ریشه میانگین مربعات خطای برآوردها محاسبه‌شده از ۱۰۰۰

(۶).

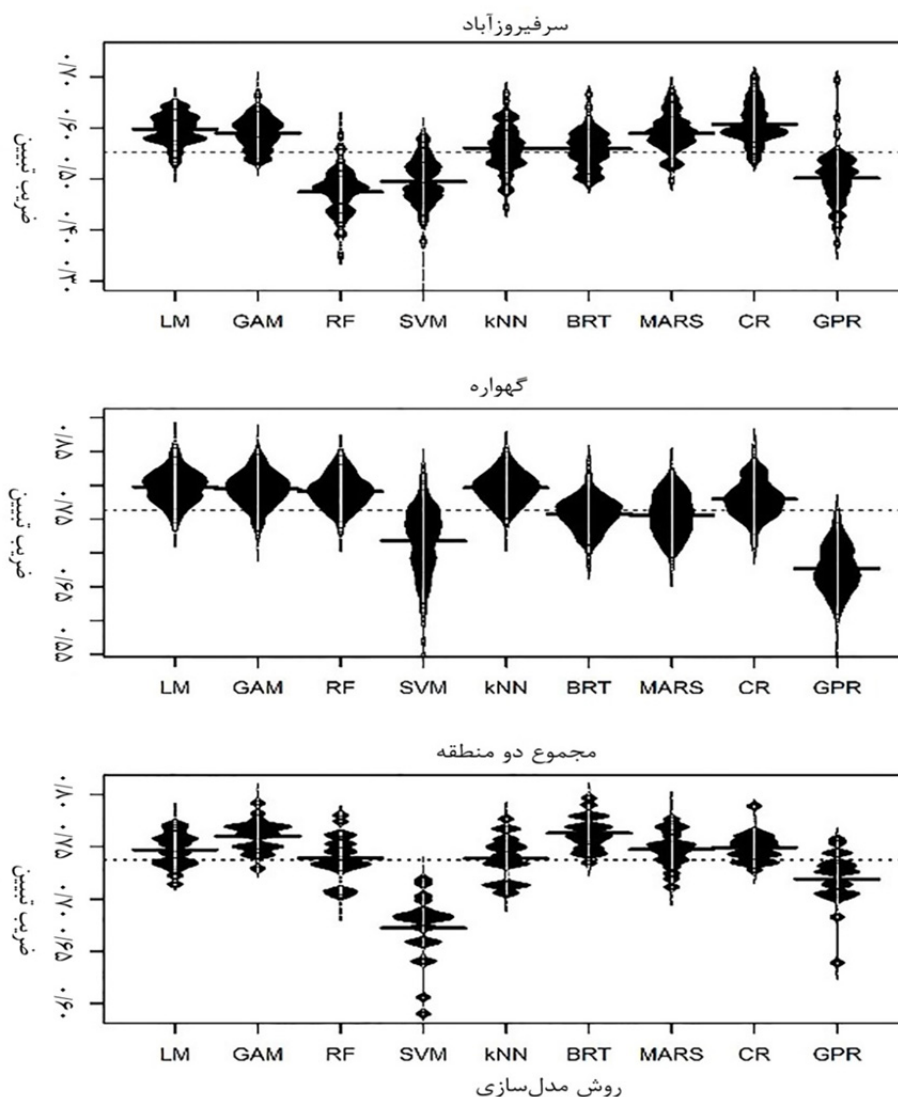
در سرفیروزآباد، میانگین مقدار ضریب تبیین بین $0/48$ (محاسبه شده با مدل جنگل تصادفی) و $0/61$ (محاسبه شده توسط مدل کویست) بود. این مقادیر برای گهواره از $0/68$ (محاسبه شده با مدل پردازش گوسی) و $0/87$ برای روش های k-امین نزدیک ترین همسایه و رگرسیون خطی بود. بیشینه و کمینه ریشه میانگین مربعات خطای برآوردها برای سرفیروزآباد به ترتیب $4/0$ تن در هکتار (محاسبه شده با روش جنگل تصادفی) و $3/3$ تن در هکتار (محاسبه شده با روش کویست) به دست آمده آمد. کمینه و بیشینه ریشه میانگین مربعات خطای برآوردها برای منطقه گهواره به ترتیب $5/8$ تن در هکتار (محاسبه شده با روش جنگل تصادفی) و $7/4$ تن در هکتار (محاسبه شده با مدل پردازش گوسی) به دست آمد.

به طور کلی، حداکثر و حداقل میانگین ریشه میانگین مربعات خطای برآوردها برای مجموع دو منطقه به ترتیب $6/0$ و $5/1$ تن در هکتار به ترتیب توسط روش های ماشین-بردار پشتیبان و درختان رگرسیون تقویت شده محاسبه شد.

محاسبه شده توسط نمونه برداری مجدد در فرآیند بوت استرپ توسط مدل رگرسیون خطی، k-امین نزدیک ترین همسایه و مدل جمعی تعمیم یافته با محاسبه بالاترین ضریب تبیین با مقدار $0/87$ نسبت به سایر روش ها بهتر عمل کردند، با این حال، سه روش مدل سازی اختلاف معنی داری با هم نداشتند.

مدل جمعی تعمیم یافته و درختان رگرسیون تقویت شده برای محاسبه ضریب تبیین با مقدار $0/76$ در مورد مجموع دو منطقه نتایج بهتری از دیگر روش ها کسب کرد (شکل های ۳ و ۴).

مدل کویست حداقل میانگین مربعات خطای برآوردها ($3/3$ تن در هکتار) را در سرفیروزآباد محاسبه کرد. روش های k-امین نزدیک ترین همسایه و جنگل تصادفی در گهواره با ریشه میانگین مربعات خطای برآوردها ($5/8$ تن در هکتار) به طور معنی داری با روش های دیگر اختلاف داشتند. در حالت مجموع کل نمونه ها بدون تفکیک به دو منطقه، روش مدل سازی درختان رگرسیون تقویت شده با میانگین $5/3$ تن در هکتار عملکرد بهتری از دیگر روش ها ارائه نمودند (شکل های ۵ و



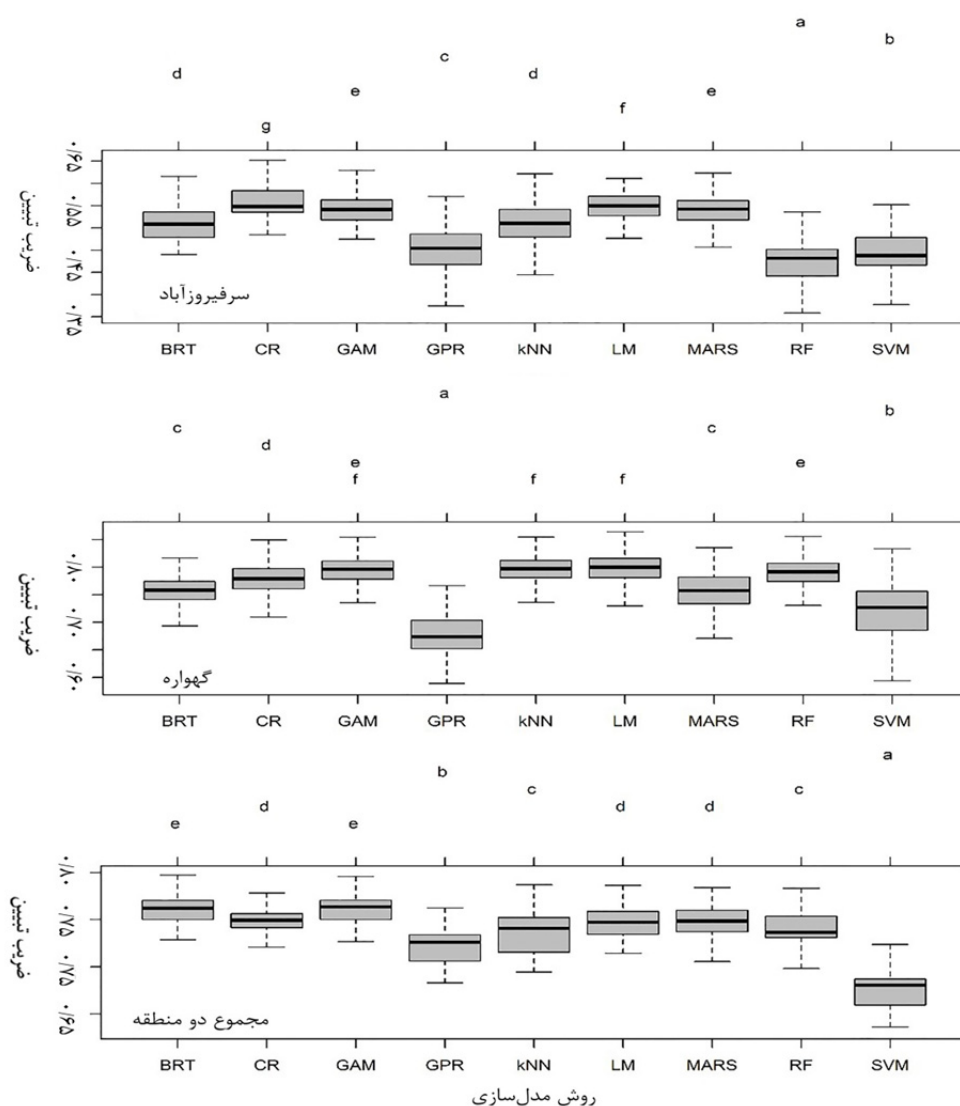
شکل ۳. ضریب تبیین محاسبه‌شده برای منطقه سرفیروزآباد (نمودار بالا)، منطقه گهواره (نمودار وسط) و مجموع دو منطقه (نمودار پایین) با استفاده از روش ارزیابی متقابل ۱۰-fold و ۱۰۰۰ مرتبه تکرار برای ۹ روش مدل‌سازی مورد استفاده، LM: رگرسیون خطی، GAM: مدل جمعی تعمیم‌یافته، RF: جنگل تصادفی، SVM: ماشین بردار پشتیبان، kNN: k-امین نزدیک‌ترین همسایه، BRT: درختان رگرسیون تقویت‌شده، MARS: رگرسیون اسپلاین تطبیقی چندمتغیره، CR: کوئیست و GPR: مدل پردازش گوسی

Fig 3. Calculated R^2 for SarfiruzAbad (top), Gahvareh (middle) and combined two sites (bottom) with 1000-repeated 10-fold cross-validation for nine modeling methods: including linear regression (LR), generalized additive model (GAM), random forest (RF), boosted regression tree (BRT), support vector machine (SVM), k-nearest neighbor (kNN), cubist regression (CR), Gaussian process model (GPR), multivariate adaptive regression spline (MARS)

های بانندی ساده، تبدیل خطی (مانند TCA و PCA)، و شاخص‌های گیاهی به ترتیب ۴۸، ۲۱، ۲۰ و ۱۱ درصد کل متغیرهای استفاده‌شده در تمامی مدل‌ها را به خود اختصاص داده‌اند.

اهمیت متغیرهای پیشگو

نتایج مطالعه نشان داد که متغیرهای طیفی استخراج‌شده از تصویرهای لندست توانایی مختلفی را در مدل‌های مختلف از خود نشان دادند (جدول ۳). به‌طور کلی، فراوانی نسبی نسبت-



شکل ۴. نمودار جعبه‌ای محاسبه شده برای ضریب تبیین برای منطقه سرفیروزآباد (نمودار بالا)، منطقه گهواره (نمودار وسط) و مجموع دو منطقه (نمودار پایین) با استفاده از روش ارزیابی متقابل ۱۰-fold و ۱۰۰۰ مرتبه تکرار برای ۹ روش مدل‌سازی مورد استفاده، LM: رگرسیون خطی، GAM: مدل جمعی تعمیم یافته، RF: جنگل تصادفی، SVM: ماشین بردار پشتیبان، kNN: k-امین نزدیک‌ترین همسایه، BRT: درختان رگرسیون تقویت شده، MARS: رگرسیون اسپلاین تطبیقی چندمتغیره، CR: کویست و GPR: مدل پردازش گوسی

Fig 4. Calculated box-plot for R^2 in SarfiruzAbad (top chart), Gahvareh (middle chart) and combined two sites (bottom chart) with 1000-repeated 10-fold cross-validation for nine modeling methods: including linear regression (LR), generalized additive model (GAM), random forest (RF), boosted regression tree (BRT), support vector machine (SVM), k-nearest neighbor (kNN), cubist regression (CR), Gaussian process model (GPR), multivariate adaptive regression spline (MARS)

جدول ۳. متغیرهای طیفی استفاده‌شده در روش‌های مدل‌سازی

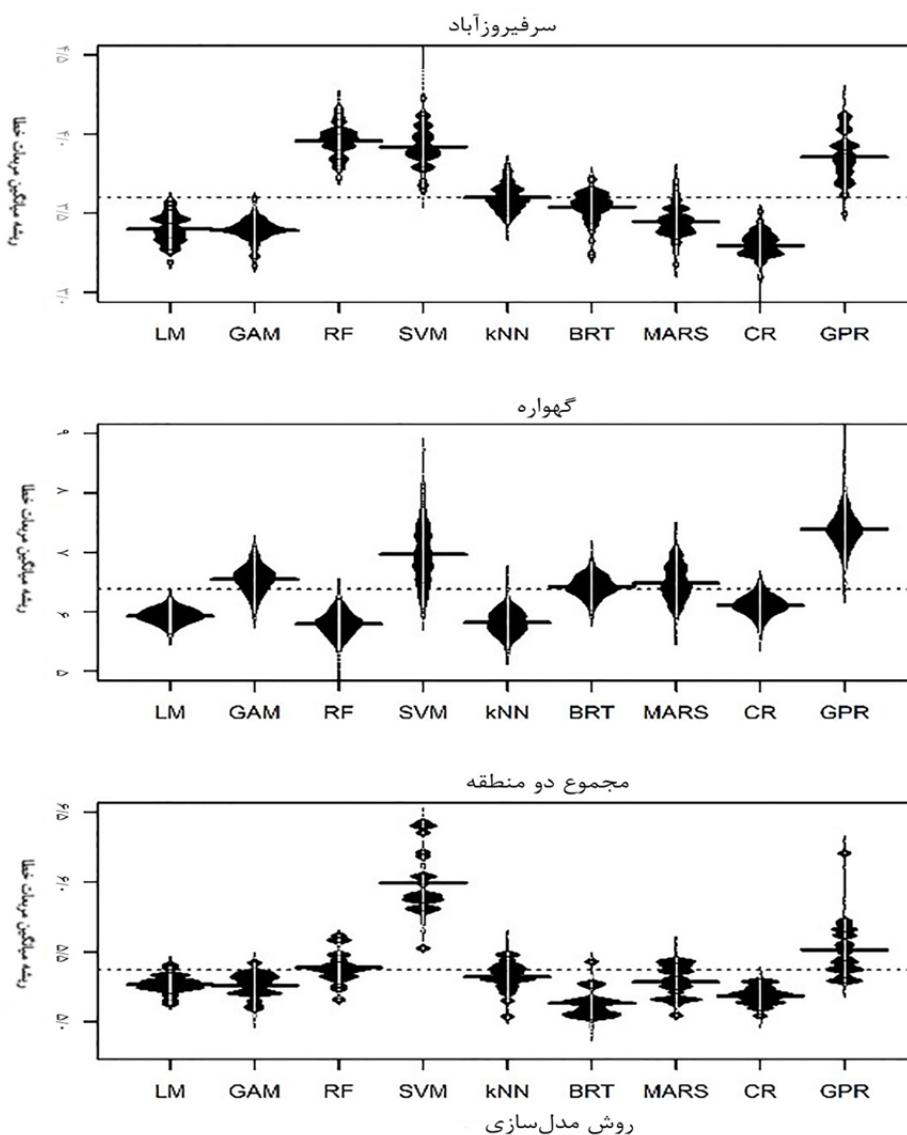
Table 3. Applied spectral variables in modeling methods

روش مدل‌سازی	منطقه مورد مطالعه		
	سرفیروزآباد	گهواره	مجموع دو منطقه
LM	NR, R	S1S2, MSAVI2	NR, Greeness
GAM	R, N	GS1, NS2	NR, S1S2
RF	NR, R	S1S2, MSAVI2	NR, Greeness
SVM	NR, R	S1S2, MSAVI2	NR, GREENESS
kNN	R, NDVIC	TCA, GREENESS	NR, GN, S1S2
BRT	NR, NDVI	TCA, S1S2	NR, GN, S1S2
MARS	NR, BR	TCA, S1S2	NR, S1S2
CR	G, DVI, GREENESS	S1, S2, GREENESS	N, R, B
GPR	NR, R, NDVIC	TCA, NR	PC5, GREENESS, S1S2

جدول ۴. متغیرهای پیشگو نهایی انتخاب‌شده در هر مدل

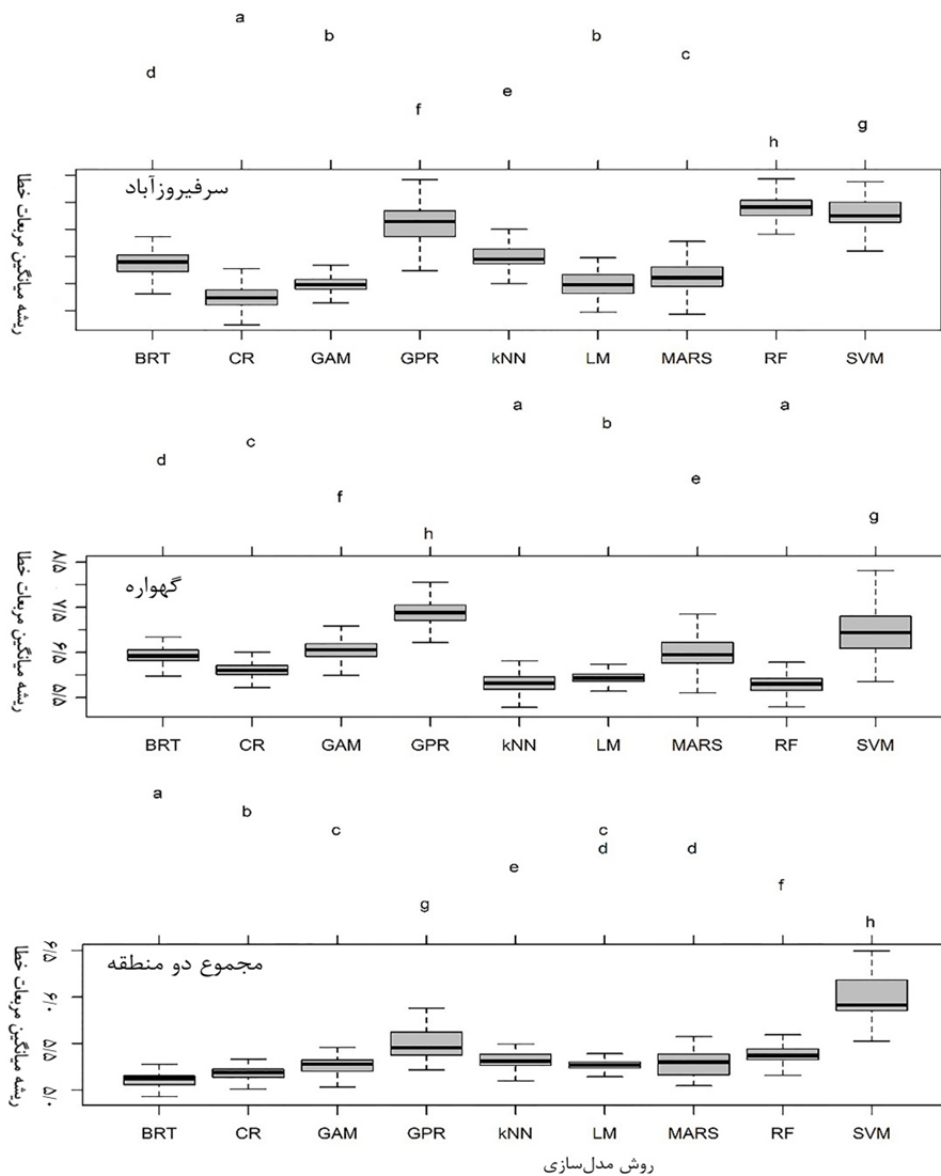
Table 4. selected explanatory variables in each model

روش مدل‌سازی	منطقه مورد مطالعه		
	سرفیروزآباد	گهواره	مجموع دو منطقه
LM	NR, R	S1S2, MSAVI2	NR, GREENESS
GAM	R, N	GS1, NS2	NR, S1S2
RF	NR, R	S1S2, MSAVI2	NR, GREENESS
SVM	NR, R	S1S2, MSAVI2	NR, GREENESS
kNN	R, NDVIC	TCA, GREENESS	NR, GN, S1S2
BRT	NR, NDVI	TCA, S1S2	NR, GN, S1S2
MARS	NR, BR	TCA, S1S2	NR, S1S2
CR	G, DVI, GREENESS	S1, S2, GREENESS	N, R, B
GPR	NR, R, NDVIC	TCA, NR	PC5, GREENESS, S1S2



شکل ۵. ریشه میانگین مربعات خطای محاسبه شده برای منطقه سرفیروزآباد (نمودار بالا)، منطقه گهواره (نمودار وسط) و مجموع دو منطقه (نمودار پایین) برحسب تن در هکتار با استفاده از روش ارزیابی متقابل ۱۰-fold و ۱۰۰۰ مرتبه تکرار برای ۹ روش مدل سازی مورد استفاده، LM: رگرسیون خطی، GAM: مدل جمعی تعمیم یافته، RF: جنگل تصادفی، SVM: ماشین بردار پشتیبان، kNN: k-امین نزدیک ترین همسایه، BRT: درختان رگرسیون تقویت شده، MARS: رگرسیون اسپلاین تطبیقی چندمتغیره، CR: کویست و GPR: مدل پردازش گوسی

Fig 5. Calculated RMSPE for SarfiruzAbad (top chart), Gahvareh (middle chart) and combined two sites (bottom chart) with 1000-repeated 10-fold cross-validation for nine modeling methods: including linear regression (LR), generalized additive model (GAM), random forest (RF), boosted regression tree (BRT), support vector machine (SVM), k-nearest neighbor (kNN), cubist regression (CR), Gaussian process model (GPR), multivariate adaptive regression spline (MARS)



شکل ۶. نمودار جعبه‌ای برای ریشه میانگین مربعات خطای محاسبه‌شده برای منطقه سرفیروزآباد (نمودار بالا)، منطقه گهواره (نمودار وسط) و مجموع دو منطقه (نمودار پایین) با استفاده از روش ارزیابی متقابل ۱۰-fold و ۱۰۰۰ مرتبه تکرار برای ۹ روش مدل‌سازی مورد استفاده، LM: رگرسیون خطی، GAM: مدل جمعی تعمیم‌یافته، RF: جنگل تصادفی، SVM: ماشین بردار پشتیبان، k-NN: k-امین نزدیک‌ترین همسایه، BRT: درختان رگرسیون تقویت‌شده، MARS: رگرسیون اسپلاین تطبیقی چندمتغیره، CR: کویست و GPR: مدل پردازش گوسی

Fig 4. Calculated box-plot for RMSPE in SarfiruzAbad (top chart), Gahvareh (middle chart) and combined two sites (bottom chart) with 1000-repeated 10-fold cross-validation for nine modeling methods: including linear regression (LR), generalized additive model (GAM), random forest (RF), boosted regression tree (BRT), support vector machine (SVM), k-nearest neighbor (kNN), cubist regression (CR), Gaussian process model (GPR), multivariate adaptive regression spline (MARS)

بحث و نتیجه‌گیری

تصمیم‌گیری و اجرای برنامه‌های تعدیل تغییر اقلیم برای اکوسیستم‌های جنگلی، نیازمند وجود اطلاعات منطقی و ابزارهای لازم برای برآورد دقیق اندوخته کربن در مقیاس‌های منطقه‌ای تا جهانی است (۲۶). تحقیق حاضر باهدف مقایسه روش‌های مختلف مدل‌سازی بر صحت برآورد زی‌توده با استفاده از تصویرهای لندست ۸ انجام شد. تعداد زیادی از متغیرهای استخراج‌شده از تصویر ماهواره‌ای لندست مانند شاخص‌های گیاهی، مقدارهای بانندی، تبدیل‌های ساده می‌تواند برای برآورد زی‌توده استفاده شود. به‌هرحال، تعداد محدودی از متغیرهای سنجش‌ازدوری دارای ضریب همبستگی بالایی با مشخصه‌های کمی جنگل مانند زی‌توده هستند. مثلاً مادون‌قرمز نزدیک، شاخص گیاهی تفاضلی نرمال شده و سبزینه از تسلدکپ اطلاعات مشابهی دارند و ترکیبی از چنین متغیرهایی نمی‌تواند سبب ارتقاء صحت مدل‌سازی زی‌توده شوند.

نقش متغیرهای طیفی در برآورد سنجش‌ازدوری زی‌توده بستگی به وضعیت ساختار جنگل دارد. در تحقیق حاضر، متغیرهای مورد استفاده در مدل‌های نهایی در دو منطقه باهم متفاوت بودند. این مطالعه نشان داد که نسبت‌های بانندی ساده حاصل مادون‌قرمز نزدیک با طول‌موج کوتاه ۱ به مادون‌قرمز نزدیک با طول‌موج کوتاه ۲ و مادون‌قرمز نزدیک به قرمز یک نقش مهم‌تری از سایر متغیرهای طیفی داشته‌اند. کارایی بهتر باندهای مادون‌قرمز نزدیک با طول‌موج کوتاه به سبب این موضوع است که این باندها حساسیت بیشتری به مؤلفه‌های ساختار توده‌های جنگلی داشته و شرایط اتمسفر اثر کمتری بر روی ضبط داده‌ها در این طول‌موج‌ها نسبت به سایر باندها دارد، این نتیجه مشابه با نتایج کارلسون و همکاران (۱۸) است. دیگر آنکه، برای مدل‌سازی زی‌توده متغیرهای مختلف از داده‌های سنجش‌ازدوری در هرکدام از مدل‌های مورد استفاده به کار گرفته شد و یک یا چند متغیر واحد که بتواند به‌طور مؤثر پیچیدگی ساختار توده‌های جنگلی را مدل کرده و بهبود قابل‌توجهی در برآوردهای زی‌توده ارائه دهد قابل‌گزارش نیست. زی‌توده برآورد شده توسط روش‌های مدل‌سازی مختلف برای دو منطقه مورد بررسی بیانگر صحت بالاتر برآوردها در منطقه گهواره نسبت به سرفیروزآباد است.

فعالیت‌های انسانی که منجر به قطع و تغییر شکل درختان در سرفیروزآباد شده سبب دو عامل مهم شده است، اول اینکه باعث برآورد نادرست زی‌توده در سطح تک‌درخت می‌شود. از طرف دیگر، کاهش سطح تاج درختان سبب افزایش انعکاس خاک بدون پوشش و اختلاط بیشتر بازتاب‌های منعکس شده در سطح پیکسل می‌گردد.

پیکسل‌های آمیخته به‌عنوان یکی از عوامل تأثیرگذار بر صحت برآورد زی‌توده در مناطق خشک و نیمه‌خشک با پوشش گیاهی تنک مانند زاگرس می‌گذارد (۲ و ۸). مقایسه نتایج نشان‌دهنده برآورد بهتر زی‌توده روی زمینی جنگل‌های زاگرس در مقایسه با سایر تحقیقات انجام‌شده در خصوص برآورد زی‌توده با تصویرهای لندست است (۶، ۱۴ و ۲۶). این موضوع می‌تواند به دلیل تراکم نسبتاً کم زی‌توده در منطقه مورد بررسی باشد که از وقوع پدیده اشباع جلوگیری می‌کند. اشباع یکی از مشکلات موجود در استفاده از تصاویر ماهواره‌ای در تراکم بالای زی‌توده است که سبب انحراف برآوردهای سنجش‌ازدوری می‌گردد (۸ و ۱۴). زیرا مقدار نسبتاً کم زی‌توده در تاج پوشش کم تراکم سبب انعکاس کم در دامنه باند مادون‌قرمز شده که می‌تواند بر پدیده اشباع غلبه کند (۱۳ و ۳۷).

کارایی روش‌های مختلف مدل‌سازی مورد استفاده در دو منطقه مورد بررسی و مجموع دو منطقه باهم متفاوت است و نمی‌توان یک روش را به‌عنوان بهترین روش برای مناطق مورد بررسی معرفی کرد. روش‌های کویست، k-امین نزدیک‌ترین همسایه و درختان رگرسیون تقویت‌شده به ترتیب در منطقه‌های سرفیروزآباد، گهواره و مجموع دو منطقه برآورد بهتری ارائه کردند. مطالعات گذشته نشان‌دهنده کارایی بالای روش جنگل تصادفی در برآورد زی‌توده است، که در این مطالعه چنین نتیجه‌ای حاصل نشد. جنگل تصادفی به‌طور معمول نیازمند تعداد زیاد نمونه برای بهبود برآوردها است (۷). علاوه بر این کراونیکلر و روسمن (۱۹) اظهار کردند که روش جنگل تصادفی از ترکیب مرزهای خطی در تصمیم‌گیری استفاده کرده و این کار در مورد تعداد نمونه‌های

برای جنگل‌های زاگرس و دیگر مناطق مشابه باشد. اهمیت این نتیجه برای جنگل‌های غیراقتصادی و در کشورهای درحال توسعه که بودجه محدودی برای مدیریت جنگل‌ها اختصاص می‌دهند بسیار قابل توجه‌تر خواهد بود.

References

1. Aghababae M, Ebrahimi A, Tahmasebi P. 2018. Comparison vegetation indices and tasseled cap transformation for estimates of soil organic carbon using Landsat-8 OLI images in a semi-steppe rangelands. *Journal of RS and GIS for Natural Resources*, 9(3): 58-59. (In Persian)
2. Calvao T, Palmeirim J. 2004. Mapping Mediterranean scrub with satellite imagery: biomass estimation and spectral behaviour. *International Journal of Remote Sensing*, 25(16): 3113-3126. doi:https://doi.org/10.1080/01431160310001654978.
3. Castillo JAA, Apan AA, Maraseni TN, Salmo SG. 2017. Estimation and mapping of above-ground biomass of mangrove forests and their replacement land uses in the Philippines using Sentinel imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 134: 70-85. doi:https://doi.org/10.1016/j.isprsjprs.2017.10.016.
4. Dai L, Jia J, Yu D, Lewis BJ, Zhou L, Zhou W, Zhao W, Jiang L. 2013. Effects of climate change on biomass carbon sequestration in old-growth forest ecosystems on Changbai Mountain in Northeast China. *Forest Ecology and Management*, 300: 106-116. doi:https://doi.org/10.1016/j.foreco.2012.06.046.
5. Domingo D, Lamelas MT, Montealegre AL, García-Martín A, De la Riva J. 2018. Estimation of total biomass in aleppo pine forest stands applying parametric and nonparametric methods to low-density airborne laser scanning data. *Forests*, 9(4): 158. doi:https://doi.org/10.3390/f9040158.
6. Du H, Cui R, Zhou G, Shi Y, Xu X, Fan W, Lü Y. 2010. The responses of Moso bamboo (*Phyllostachys heterocycla* var. *pubescens*) forest aboveground biomass to Landsat TM spectral reflectance and NDVI. *Acta Ecologica Sinica*, 30(5): 257-263. doi:https://doi.org/10.1016/j.chnaes.2010.08.005.
7. Dube T, Mutanga O. 2015. Evaluating the utility of the medium-spatial resolution Landsat 8 multispectral sensor in quantifying aboveground biomass in uMgeni catchment, South Africa. *ISPRS Journal of Photogrammetry and Remote Sensing*, 101: 36-46. doi:https://doi.org/10.1016/j.isprsjprs.2014.11.001.
8. Eisfelder C, Kuenzer C, Dech S. 2012. Derivation

کم امکان‌پذیر نیست. این نتیجه مشابه با تحقیق گائو و همکاران (۱۲) است که بیان کردند وقتی که مقادیرهای زی‌توده کم است روش‌های ناپارامتری می‌توانند سبب بهبود برآورد زی‌توده شوند. این موضوع نشان‌دهنده اهمیت موضوع انتخاب روش مدل‌سازی برای بهبود صحت برآورد، مخصوصاً در مورد مناطق با تراکم کم زی‌توده مانند زاگرس است. دیگر نتایج این مطالعه بیان کرد که نمی‌توان یک روش مدل‌سازی واحد را برای هر دو منطقه و مجموع دو منطقه معرفی کرد، و بهتر است که در مطالعات آتی روش‌های مختلف مدل‌سازی برای بهبود صحت برآوردها مورد آزمون قرار گیرند. به‌رحال باوجود نتایج قابل‌قبول این مطالعه، هنوز بخشی از واریانس زی‌توده تبیین نشده است. قطعات نمونه‌برداری یک از بنیان‌های برآورد سنجش‌ازدوری زی‌توده است. عدم قطعیت داده‌های قطعه‌نمونه می‌تواند ناشی از رویکرد نمونه‌برداری، اندازه نمونه، مدل آلومتری و اندازه‌گیری ویژگی‌های درختان طی آماربرداری باشد. تعداد کافی قطعه‌نمونه سبب بهبود صحت برآورد و استفاده از مدل‌های آلومتری تا حدود ۲۰ درصد خطای برآورد را شامل می‌شوند (۱۲ و ۳۴). علاوه بر این نکات، مقدار کم زی‌توده سبب پیچیدگی ترکیب نوع پوشش زمین از قبیل زمین بدون پوشش و پوشش گیاهی و درختی و تشکیل پیکسل‌های آمیخته می‌گردد، که بازتاب ثبت‌شده توسط سنجنده دقیقاً جداکننده عوارض موجود در پیکسل نیست (۸ و ۱۰). ازاین‌رو سبب کاهش همبستگی بین مقدار زی‌توده اندازه‌گیری شده در قطعه‌نمونه و اطلاعات استخراج‌شده از داده‌های سنجش‌ازدوری متناظر با موقعیت مکانی قطعه‌نمونه می‌شود.

نتایج این تحقیق نشان داد که عملکرد روش‌های مدل‌سازی بسته به شرایط توده موردبررسی متفاوت بوده و ازاین‌رو نمی‌توان یک روش را به‌عنوان بهترین روش معرفی کرد. علاوه براین، صحت نسبتاً بالای برآوردهای زی‌توده با استفاده از تصویرهای رایگان لندست این موضوع را نمایان می‌سازد که استفاده از روش‌های مدل‌سازی ناپارامتری و تصویر لندست ۸ می‌تواند یک‌راه ساده، کم‌هزینه و اجرایی

9. of biomass information for semi-arid areas using remote-sensing data. *International Journal of Remote Sensing*, 33(9): 2937-2984. doi:<https://doi.org/10.1080/01431161.2011.62003>.
10. Fassnacht FE, Hartig F, Latifi H, Berger C, Hernández J, Corvalán P, Koch B. 2014. Importance of sample size, data type and prediction method for remote sensing-based estimations of aboveground forest biomass. *Remote Sensing of Environment*, 154: 102-114. doi:<https://doi.org/10.1016/j.rse.2014.07.028>.
11. Fernández-Manso O, Fernández-Manso A, Quintano C. 2014. Estimation of aboveground biomass in Mediterranean forests by statistical modelling of ASTER fraction images. *International Journal of Applied Earth Observation and Geoinformation*, 31: 45-56. doi:<https://doi.org/10.1016/j.jag.2014.03.005>.
12. Fu L, Zhao Y, Xu Z, Wu B. 2015. Spatial and temporal dynamics of forest aboveground carbon stocks in response to climate and environmental changes. *Journal of Soils and Sediments*, 15(2): 249-259. doi:10.1007/s11368-014-1050-x.
13. Gao Y, Lu D, Li G, Wang G, Chen Q, Liu L, Li D. 2018. Comparative analysis of modeling algorithms for forest aboveground biomass estimation in a subtropical region. *Remote Sensing*, 10(4): 627. doi:<https://doi.org/10.3390/rs10040627>.
14. Gasparri NI, Parmuchi MG, Bono J, Karszenbaum H, Montenegro CL. 2010. Assessing multi-temporal Landsat 7 ETM+ images for estimating above-ground biomass in subtropical dry forests of Argentina. *Journal of Arid Environments*, 74(10): 1262-1270. doi:<https://doi.org/10.1016/j.jaridenv.2010.04.007>.
15. Gizachew B, Solberg S, Næsset E, Gobakken T, Bollandsås OM, Breidenbach J, Zahabu E, Mauya EW. 2016. Mapping and estimating the total living biomass and carbon in low-biomass woodlands using Landsat 8 CDR data. *Carbon balance and management*, 11(1): 13. doi:<https://doi.org/10.1186/s13021-016-0055-8>.
16. Görgens EB, Montaghi A, Rodriguez LCE. 2015. A performance comparison of machine learning methods to estimate the fast-growing forest plantation yield based on laser scanning metrics. *Computers and Electronics in Agriculture*, 116: 221-227. doi:<https://doi.org/10.1016/j.compag.2015.07.004>.
17. Huffman T, Liu J, McGovern M, McConkey B, Martin T. 2015. Carbon stock and change from woody biomass on Canada's cropland between 1990 and 2000. *Agriculture, Ecosystems & Environment*, 205: 102-111. doi:<https://doi.org/10.1016/j.agee.2014.10.009>.
18. Iranmanesh Y. 2013. Assessment on biomass estimation methods and carbon sequestration of *Quercus brantii* Lindl. in Chaharmahal & Bakhtiari Forests. PhD Thesis, Tarbiat Modares University, 106 pp, (In Persian)
19. Karlson M, Ostwald M, Reese H, Sanou J, Tankoano B, Mattsson E. 2015. Mapping tree canopy cover and aboveground biomass in Sudano-Sahelian woodlands using Landsat 8 and random forest. *Remote Sensing*, 7(8): 10017-10041. doi:<https://doi.org/10.3390/rs70810017>.
20. Krahwinkler P, Rossman J. 2011. Using decision tree based multiclass support vector machines for forest mapping. In: Lena Halounová EE (ed) IEEE Int. Geoscience and Remote Sensing Symp., Vancouver, Canada. pp 307-318.
21. Kwak D-A, Lee W-K, Cho H-K, Lee S-H, Son Y, Kafatos M, Kim S-R. 2010. Estimating stem volume and biomass of *Pinus koraiensis* using LiDAR data. *Journal of Plant Research*, 123(4): 421-432. doi:10.1007/s10265-010-0310-0.
22. Latifi H, Fassnacht F, Koch B. 2012. Forest structure modeling with combined airborne hyperspectral and LiDAR data. *Remote Sensing of Environment*, 121: 10-25. doi:<https://doi.org/10.1016/j.rse.2012.01.015>.
23. McRoberts RE, Magnussen S, Tomppo EO, Chirici G. 2011. Parametric, bootstrap, and jackknife variance estimators for the k-Nearest Neighbors technique with illustrations using forest inventory and satellite image data. *Remote Sensing of Environment*, 115(12): 3165-3174. doi:<https://doi.org/10.1016/j.rse.2011.07.002>.
24. Nolè A, Law B, Magnani F, Matteucci G, Ferrara A, Ripullone F, Borghetti M. 2009. Application of the 3-PGS model to assess carbon accumulation in forest ecosystems at a regional level. *Canadian Journal of Forest Research*, 39(9): 1647-1661. doi:<https://doi.org/10.1139/X09-077>.
25. Noorian N, Shataee S, Mohamadi J. 2019. Evaluation of RapidEye satellite data for estimation some quantitative structure variables in the Caspian forests of Gorgan region. *Journal of RS and GIS for Natural Resources*, 9(4): 1-16. (In Persian)
26. Pan Y, Birdsey RA, Fang J, Houghton R, Kauppi PE, Kurz WA, Phillips OL, Shvidenko A, Lewis SL, Canadell JG. 2011. A large and persistent carbon sink in the world's forests. *Science*, 333(6045): 988-993. doi:<https://doi.org/10.1126/science.1201609>.
27. Powell SL, Cohen WB, Healey SP, Kennedy RE, Moisen GG, Pierce KB, Ohmann JL. 2010. Quantification of live aboveground forest biomass dynamics with Landsat time-series and field inventory data: A comparison of empirical modeling approaches. *Remote Sensing of Environment*, 114(5): 1053-1068. doi:<https://doi.org/10.1016/j.rse.2009.12.018>.
28. Rajashekar G, Fararoda R, Reddy RS, Jha CS,

- Ganeshaiyah KN, Singh JS, Dadhwal VK. 2018. Spatial distribution of forest biomass carbon (Above and below ground) in Indian forests. *Ecological Indicators*, 85: 742-752. doi:<https://doi.org/10.1016/j.ecolind.2017.11.024>.
29. Riaño D, Chuvieco E, Salas J, Aguado I. 2003. Assessment of different topographic corrections in Landsat-TM data for mapping vegetation types (2003). *IEEE Transactions on Geoscience and Remote Sensing*, 41(5): 1056-1061. doi:<https://doi.org/10.1109/TGRS.2003.811693>.
30. Safari A, Sohrabi H. 2019. The effect of digital preprocessing and modeling method on an estimation of aboveground carbon stock of Zagros forests using Landsat 8 imagery. *Journal of RS and GIS for Natural Resources*, 9(4): 73-89. (In Persian)
31. Safari A, Sohrabi H, Powell SL. 2018. Comparison of satellite-based estimates of aboveground biomass in coppice oak forests using parametric, semiparametric, and nonparametric modeling methods. *Journal of Applied Remote Sensing*, 12(4): 046026. doi:<https://doi.org/10.1117/1.JRS.12.046026>.
32. Sagheb-Talebi K, Pourhashemi M, Sajedi T. 2014. *Forests of Iran*. The Netherlands: Springer Netherlands, 152 pp. <https://doi.org/10.1007/978-94-007-7371-4>.
33. Sarker LR, Nichol JE. 2011. Improved forest biomass estimates using ALOS AVNIR-2 texture indices. *Remote Sensing of Environment*, 115(4): 968-977. doi:<https://doi.org/10.1016/j.rse.2010.11.010>.
34. Shao Z, Zhang L. 2016. Estimating forest aboveground biomass by combining optical and SAR data: A case study in Genhe, Inner Mongolia, China. *Sensors*, 16(6): 834. doi:<https://doi.org/10.3390/s16060834>.
35. Wang X, Shao G, Chen H, Lewis BJ, Qi G, Yu D, Zhou L, Dai L. 2013. An Application of Remote Sensing Data in Mapping Landscape-Level Forest Biomass for Monitoring the Effectiveness of Forest Policies in Northeastern China. *Environmental Management*, 52(3): 612-620. doi:[10.1007/s00267-013-0089-6](https://doi.org/10.1007/s00267-013-0089-6).
36. Zhao K, Suarez JC, Garcia M, Hu T, Wang C, Londo A. 2018. Utility of multitemporal lidar for forest and carbon monitoring: Tree growth, biomass dynamics, and carbon flux. *Remote Sensing of Environment*, 204: 883-897. doi:<https://doi.org/10.1016/j.rse.2017.09.007>.
37. Zhao P, Lu D, Wang G, Wu C, Huang Y, Yu S. 2016. Examining spectral reflectance saturation in Landsat imagery and corresponding solutions to improve forest aboveground biomass estimation. *Remote Sensing*, 8(6): 469. doi:<https://doi.org/10.3390/rs8060469>.
38. Zhu X, Liu D. 2015. Improving forest aboveground biomass estimation using seasonal Landsat NDVI time-series. *ISPRS Journal of Photogrammetry and Remote Sensing*, 102: 222-231. doi:<https://doi.org/10.1016/j.isprsjprs.2014.08.014>.



Using the bootstrap approach for comparing statistical modeling methods to estimate remotely-sensed aboveground biomass in Zagros forests

Amir Safari, Hormoz Sohrabi

Received: 28 April 2019/ Accepted: 19 May 2020
Available online 5 July 2020

Abstract

Background and Objective Considering the increasing importance of forest ecosystems in climate change mitigation projects, reliable and cost-effective methods are required to estimate the aboveground biomass (AGB). Common methods used to estimate the aboveground biomass (AGB) include in-situ measurement, the biomass calculation using allometric equations and using remote sensing techniques. Remote sensing has been widely used to estimate the biomass of forests in recent decades. The used statistical modeling method is one of the most important factors to use remotely-sensed data for estimation of the aboveground biomass. A large number of researches have been carried out about using the modeling methods. However, these studies face the following different challenges: 1) no modeling method has been recommended as the best method 2) the performance of these modeling methods is affected by forest type, the forest structure, and the present disturbance intensity 3) the performance evaluation and the comparison of the results of these methods were done by using goodness-of-fit test and cross-validation methods. The purpose of this study is to considering the role of choosing statistical modeling methods to estimate remotely-sensed aboveground biomass, the current study was conducted to investigate nine statistical

modeling methods including linear regression (LR), generalized additive model (GAM), random forest (RF), support vector machine (SVM), boosted regression tree (BRT), k-nearest neighbor (kNN), cubist regression (CR), Gaussian process model (GPR), multivariate adaptive regression spline (MARS) using bootstrap process and 1000-repeated 10-fold cross-validation approach to estimate the aboveground biomass of Zagros forests using Landsat 8 images.

Materials and Methods The current study was conducted in Kermanshah forests which is mostly dominated by oak species trees (*Quercus* spp.) and is located in western Iran on the Zagros Mountains. Zagros forests are generally sparse and open and comprise approximately 20% of Iran's area and 40% forest regions of Iran. In order to conduct this study, two forest regions with different levels of human disturbances were chosen; SarfiruzAbad region with highly degraded (HD) forests, and Gahvareh forest region with minor degradation (MD).

Geographical coordinates of SarfiruzAbad and Gahvareh regions are 33°57'-34°04'N / 47°03'-47°17'E & 34°21'- 34°24'N / 46°16'-46°23' E respectively. The Leaf area index (LAI) map derived from the Landsat images based on a global model was used to collect field-based sample plots in both regions of the study. Both regions were divided into three low, moderate and high Leaf area index (LAI) strata, and the locations of the sample plots were located by using a systematic inventory at the intersections of a 200m×200 m grid in each stratum.

A. Safari¹, H. Sohrabi(✉)²

1. PhD of Forestry, Faculty of Natural Resources and Marine Sciences, University of Tarbiat Modares, Tehran, Iran
2. Associate Professor, Department of Forestry, Faculty of Natural Resources and Marine Sciences, University of Tarbiat Modares, Iran
e-mail: hsohrabi@modares.ac.ir

124 georeferenced square plots of field-based sample plots (63 plots in Gahvareh region and 61 plots in SarfiruzAbad region) with 30m×30m dimensions the same size as a Landsat 8 image's pixel were collected. Allometric equation developed for oak tree in Zagros forests was used to calculate the amount of the aboveground biomass of each individual tree or sprout-clump. The allometric equation used in this study uses two vertical tree crown diameters to estimate the amount of the biomass of each individual tree or sprout-clump. The sum of the amount of the biomass of each individual tree in sample plot was used to calculate the amount of the biomass plot in sample plot level at a ton per hectare. Our study regions were located in a frame of Landsat 8 images (path/row:167/36). A cloud-free Landsat image relating to 19th Mordad 1394 (10th August 2015) relating to the time when the tree canopies are completely closed and near to the date of land inventory was downloaded from earthexplorer.usgs.gov site. Based on the previous studies, the pre-processing of the used image comprising the radiometric and topographic corrections was done using C method.

To estimate the aboveground biomass in the study areas by using remote sensing, 38 spectral variables including band values, simple band ratios, vegetation indices and common linear transformations like tasseled cap and principle component analysis were extracted from the used Landsat 8 image. Generally, the efficiency of nine different statistical modeling methods including parametric methods (Linear Regression, LR), semiparametric (Generalized Additive Model, GAM), and nonparametric Random Forest (RF), Support Vector Machine (SVM), K-nearest neighbor (KNN), Boosted regression trees (BRT), multivariate additive regression splines, cubist regression (CR), and Gaussian processes regression/model) were compared in order to estimate aboveground biomass. To assess the models, two common quality statistics: (i) determination coefficient and (2) root mean square error via 10 fold cross validation repeated 1000 times approach were calculated. This number of repeats helps to ensure an acceptable assessment of robustness of the results.

Results and Discussion The measured statistical characteristics of the field sample plots showed that the mean aboveground biomass of SarfiruzAbad and Gahvareh regions were 12.6 ton/ha and 20.5 ton/ha respectively. ANOVA indicated significant differences between modelling methods (treatment effect: $p < 0.001$) for both R^2 and RMSPE calculated in 1000-time repeats using 10-fold cross-validation.

The Cubist modeling method with the mean determination coefficient of 0.61 outperformed other methods in SarfiruzAbad region. These results for Gahvareh region showed better efficiency of linear regression (LR), generalized additive model (GAM), and k-nearest neighbor (KNN) with the mean determination coefficient of 0.87.

The multiple comparisons of different models by using Tukey test concerning RMSE showed that in SarfiruzAbad region, cubist method with the mean of RMSE 3.3 ton/ha and KNN and RF methods with the mean of RMSE 5.8 ton/ha had a significant difference in comparison to the other methods. Totally, the results of the research revealed the suitable efficiency of Landsat 8 image for AGB estimation in Zagros forests. The acceptable results are due to the low AGB in our study regions that did not reach the saturation point as one of challenges of using optical images like Landsat.

The other results of this research is the assessment of the efficiency of modeling method in order to increase the accuracy of the estimation of remotely-sensed aboveground biomass. Unlike the results of the previous studies, linear regression yielded better results compared to nonparametric methods that can be due to the presence of the linear relationship between aboveground biomass and spectral variables derived from Landsat images. Among the used various spectral variables, red, near infrared, and shortwave infrared 1 and 2 band ratios were selected as the final variable in most modeling methods.

Conclusion In this study, we evaluated the efficiency of different statistical modeling methods to estimate AGB in Zagros forests by using Landsat images. The biomass estimations were compared by using nine parametric, semi-parametric, and non-parametric methods and using 1000-repeated 10-fold cross-validation. The results illustrated the acceptable potentiality of Landsat images for cost-efficient AGB estimating in Zagros oak forests. The accuracy of AGB estimation in Gahvareh region with low-degraded forest stands was higher than SarfiruzAbad region with highly degraded stands.

Keywords Aboveground biomass, Bootstrap, Statistical modeling, Zagros forest, Landsat 8