

مقاله علمی-پژوهشی

پایش هوشمند کشت گندم دیم و آبی در حوضه مارون-جراحی با استفاده از داده‌های سنتینل و مدل جنگل تصادفی

مسعود سلطانی^۱، بهاره بهمن آبادی^{۲*}، علی مختاران^۳

تاریخ دریافت: ۱۴۰۴/۱۲/۰۳ تاریخ پذیرش: ۱۴۰۵/۰۲/۲۸

چکیده

گندم به عنوان یکی از محصولات راهبردی در تأمین امنیت غذایی، سهم قابل توجهی از اراضی کشاورزی ایران را به خود اختصاص داده است، اما تغییرات اقلیمی، محدودیت منابع آب و ناکارآمدی روش‌های سنتی آماربرداری، پایش دقیق سطح زیرکشت این محصول را با چالش مواجه ساخته است. در این پژوهش، چارچوبی تلفیقی مبتنی بر داده‌های چندمنبعی سنجش از دور و الگوریتم یادگیری ماشین جنگل تصادفی به منظور شناسایی و تفکیک مزارع گندم دیم و آبی و برآورد سطح زیرکشت آن‌ها در حوضه آبخیز مارون ارائه شد. مجموعه‌ای از شاخص‌های طیفی، راداری و حرارتی برای بازنمایی ویژگی‌های بیوفیزیکی، ساختاری و آبی گیاه استخراج و به عنوان ورودی مدل به کار گرفته شد. الگوریتم جنگل تصادفی به دلیل پایداری در برابر بیش‌برازش و توانایی پردازش داده‌های ناهمگون برای طبقه‌بندی انتخاب شد و صحت نتایج با استفاده از ماتریس درهم‌ریختگی، دقت کلی و ضریب کاپا ارزیابی گردید. نتایج نشان داد که مدل پیشنهادی عملکرد بسیار مطلوبی در شناسایی مزارع گندم دارد؛ به طوری که دقت کلی و ضریب کاپا برای داده‌های آزمون به ترتیب حدود ۹۷ و ۹۶ درصد به دست آمد. همچنین، برآورد سطح زیرکشت گندم با بایاس اندک حدود ۳ درصد انجام شد که نشان‌دهنده دقت بالای روش است.

واژه‌های کلیدی: گندم، LGRIP30، Dynamic World، ESA WorldCover

مقدمه

اختصاص داده است که از این میان، ۳۶/۵۳ درصد در اراضی آبی و ۷۴/۵۶ درصد در اراضی دیم مستقر شده‌اند (وزارت جهاد کشاورزی، ۱۴۰۳). با این حال، پایداری تولید این محصول راهبردی در ایران با تهدیدات جدی ناشی از تغییر اقلیم و محدودیت‌های فزاینده منابع آب مواجه است که ضرورت مدیریت هوشمند و بهره‌وری آب را بیش از پیش آشکار می‌سازد (Abbaspour et al., 2009; Heydari & Taran, 2025). در مواجهه با چنین چالش‌های گسترده‌ای، روش‌های سنتی آماربرداری میدانی و نمونه‌برداری‌های زمینی که بر پایه گزارش‌های خوداظهاری بهره‌برداران بنا شده‌اند، به دلیل خطای انسانی، هزینه‌های عملیاتی گزاف و تأخیر در ارائه نتایج، کارایی خود را در مدیریت بحران‌های به‌لحظه از دست داده‌اند (McNairn & Protz, 1993; Rahnema et al., 2019). فقدان دقت مکانی در این روش‌ها مانع از آن می‌شود که تصمیم‌گیران بتوانند پیش از پایان فصل زراعی، تصمیمات مدیریتی لازم را اعمال کنند. این محدودیت‌ها سبب شده است که استفاده از روش‌های نوین و دقیق‌تر برای پایش اراضی کشاورزی مورد توجه پژوهشگران قرار گیرد. تکامل فناوری سنجش از دور ماهواره‌ای در دهه‌های اخیر، با ارائه

گندم به عنوان یکی از ارکان سه‌گانه غلات راهبردی در جهان، در کنار برنج و ذرت، جایگاهی بنیادین در پایداری سیستم‌های اجتماعی و اقتصادی و تأمین امنیت غذایی جوامع بشری ایفا می‌کند (Zhang et al., 2025). اهمیت استراتژیک این محصول در بازارهای بین‌المللی فراتر از یک کالای مصرفی بوده و به عنوان یک محصول استراتژیک و موثر در اقتصاد کشورها شناخته می‌شود. بر اساس گزارش‌های آماری وزارت جهاد کشاورزی ۱۴۰۳-۱۴۰۲، کشت گندم حدود ۵۰ درصد از کل اراضی زیر کشت کشور را به خود

۱- استادیار گروه علوم و مهندسی آب، دانشکده کشاورزی و منابع طبیعی، دانشگاه بین المللی امام خمینی (ره)، قزوین، ایران

۲- نویسنده مسئول، دکتری آبیاری زهکشی، دانشگاه بین المللی امام خمینی (ره)، قزوین، ایران.

۳- استادیار پژوهشی موسسه تحقیقات فنی و مهندسی کشاورزی، سازمان تحقیقات، آموزش و ترویج کشاورزی، کرج، ایران

(*) نویسنده مسئول: رایانامه: bahmanabadi@alumni.ut.ac.ir

DOI: [10.22034/ijdi.2026.577267.2673](https://doi.org/10.22034/ijdi.2026.577267.2673)

قابلیت‌هایی همچون پوشش وسیع مکانی و سرعت بالا در بازنگری، دریاچه‌های نوینی را به سوی پایش پویای مزارع گشوده است (Bahrami et al., 2022). در مسیر توسعه مدل‌های برآورد عملکرد، چهار رویکرد اساسی شامل مدل‌های آماری تجربی (Evans & Shen, 2021)، مدل‌های نیمه‌تجربی (Guan et al., 2021)، مدل‌های همگون‌سازی داده (Huang et al., 2019)، مدل‌های مبتنی بر یادگیری ماشین (Kamir et al., 2020) شکل گرفته‌اند. اگرچه سه رویکرد نخست گام‌های مهمی در درک روابط فیزیولوژیک گیاه برداشته‌اند، اما وابستگی شدید آن‌ها به ارقام خاص محصول، مراحل رشد مشخص و محدودیت‌های جغرافیایی، مانع از تعمیم‌پذیری آن‌ها در مقیاس‌های کلان شده است. در مقابل، الگوریتم‌های یادگیری ماشین به دلیل توانمندی در پردازش داده‌های کلان و استخراج روابط غیرخطی و پنهان میان متغیرهای محیطی و پاسخ طیفی گیاه، برتری خود را در مدل‌سازی‌های پیچیده به اثبات رسانده‌اند (Obsie et al., 2020). اما، نباید از ماهیت «جعبه سیاه» در این مدل‌ها غافل شد؛ جایی که الگوریتم‌ها بدون در نظر گرفتن محدودیت‌های فیزیولوژیک مانند مراحل رشد و سبزی‌نگی، تنها به الگوهای آماری بسنده می‌کنند. بنابراین، آموزش و تقویت این مدل‌ها با اطلاعات زیستی و پارامترهای فنولوژیکی برای افزایش تفسیرپذیری و انطباق با واقعیت‌های بیوفیزیکی گیاه، اصلی‌ترین بخش در پژوهش‌های نوین است (Zhang et al., 2025). در همین راستا، درک عمیق رفتار فنولوژیکی گیاه و تعامل آن با امواج الکترومغناطیس، پیش‌نیاز هرگونه مدل‌سازی دقیق است. فنولوژی فصلی، که در واقع همان تقویم بیولوژیکی گیاه است، پاسخ طیفی ثبت شده توسط سنجنده‌ها را در مراحل مختلف از پنجه‌زنی و ساقه رفتن تا گلدهی و رسیدگی، تحت تأثیر قرار می‌دهد (Weng et al., 2009). این تغییرات فنولوژیکی منجر به ایجاد امضاهای طیفی منحصر به فردی می‌گردد که امکان تفکیک گندم را از سایر محصولات رقیب فراهم می‌سازد (Ghosh et al., 2014; Pan et al., 2012). استفاده از تصاویر چندزمانه و سری‌های زمانی شاخص‌های پوشش گیاهی^۱ کلیدی‌ترین استراتژی برای کاهش اختلاط کلاس‌ها^۲ و افزایش دقت در نقشه‌برداری‌های منطقه‌ای است (Kayad et al., 2019). با این حال، پایش مبتنی بر شاخص‌های طیفی با یک محدودیت فیزیکی بنیادین مواجه است: حضور ابرها در دوره‌های بحرانی رشد. برای نمونه، در مراحل حساس گلدهی و پر شدن دانه، که تعیین‌کننده نهایی عملکرد هستند، پوشش ابر می‌تواند منجر به گسستگی در داده‌های اپتیک شده و پایش‌بینی‌های مدل را با عدم قطعیت جدی روبرو کند. اینجاست که ضرورت تلفیق داده‌های

نوری با سنجنده‌های فعال راداری مطرح می‌گردد. تلفیق داده‌های نوری و راداری^۳ پاسخی فنی به محدودیت‌های جوی و نیاز به پایش مستمر است. در حالی که سنجنده‌های نوری اطلاعات ارزشمندی از محتوای کلروفیل و سلامت گیاه ارائه می‌دهند (Zhao et al., 2020)، سنجنده راداری (SAR)، به‌ویژه در باند C ماهواره سنتینل-۱، با توانایی عبور از ابر و غبار، تصویری بدون وقفه از مزارع فراهم می‌کند (Aobpaet, 2022). تحلیل‌های پیشرفته نشان می‌دهند که فیزیک تعامل امواج راداری در پلاریزاسیون‌های VV و VH با ساختار عمودی ساقه و حجم توده زیستی گندم، حساسیت بسیار بالایی به پویایی رشد گیاه دارد (Mandal et al., 2020; Ouaadi et al., 2023). این تلفیق داده‌ای، نه تنها خلأهای ناشی از ابرناکی را پر می‌کند، بلکه با ترکیب اطلاعات طیفی و ساختاری، دقت طبقه‌بندی را در محیط‌های پیچیده زراعی به‌طور چشمگیری ارتقا می‌دهد. مرور نظام‌مند پیشینه پژوهش، گویای صحت بالای این رویکرد تلفیقی است. به عنوان مثال، ژنگ و همکاران با به‌کارگیری شبکه عصبی Pix2Pix برای ادغام تصاویر لندست-۸ و سنتینل-۲ و تمرکز بر شاخص Cired-edge، استانداردهای جدیدی در پایش‌بینی عملکرد خلق کردند (Zhang et al., 2025). در پژوهشی دیگر، لیو و همکاران با استفاده از الگوریتم جنگل تصادفی و داده‌های چندسنجنده، به صحت کلی ۹۲/۹ درصد و ضریب کاپای ۸۵ درصد دست یافتند، در حالی که خطای نسبی برآورد سطح زیرکشت آن‌ها تنها ۳/۹ درصد بود که این نشان‌دهنده دقت بالای مدل‌های مبتنی بر یادگیری ماشین در مدیریت حجم وسیع داده‌ها است (Liu et al., 2018).

در بخش بعدی این تحلیل، مطالعات عراقی و داکاچه پتانسیل پایش‌بینی عملکرد را حتی دو ماه پیش از برداشت، با ضریب همبستگی ۰/۹۳ برای گندم آبی گزارش کرده‌اند که این یافته برای مدیریت استراتژیک غلات حیاتی است. در فضای پژوهشی ایران نیز، مطالعات ارزشمندی صورت گرفته است (Araghi & Daccache, 2025)؛ بوربور و همکاران با تحلیل عملکرد گندم در مناطق خشک، برتری الگوریتم XGBoost را با مقدار NRMSE نرمال شده ۴۲ درصد و خطای مطلق میانگین (MAE) برابر با ۰/۲۵ تن در هکتار نسبت به جنگل تصادفی و ماشین بردار پشتیبان اثبات کردند (Bourbour et al., 2025). همچنین، موسوی و همکاران (۱۳۹۹) در مقایسه روش‌های طبقه‌بندی، نشان دادند که ماشین بردار پشتیبان با دقت ۸۹ درصد عملکردی فراتر از روش حداکثر احتمال^۴ در تفکیک اراضی دیم دارد. سلطانی و همکاران (۱۴۰۴) نیز با ترکیب داده‌های سنتینل-۱ و ۲، موفق شدند به صحت کلی ۹۶ درصد در طبقه‌بندی محصولات حوضه‌ای دست یابند. با این وجود، اکثر این پژوهش‌ها در

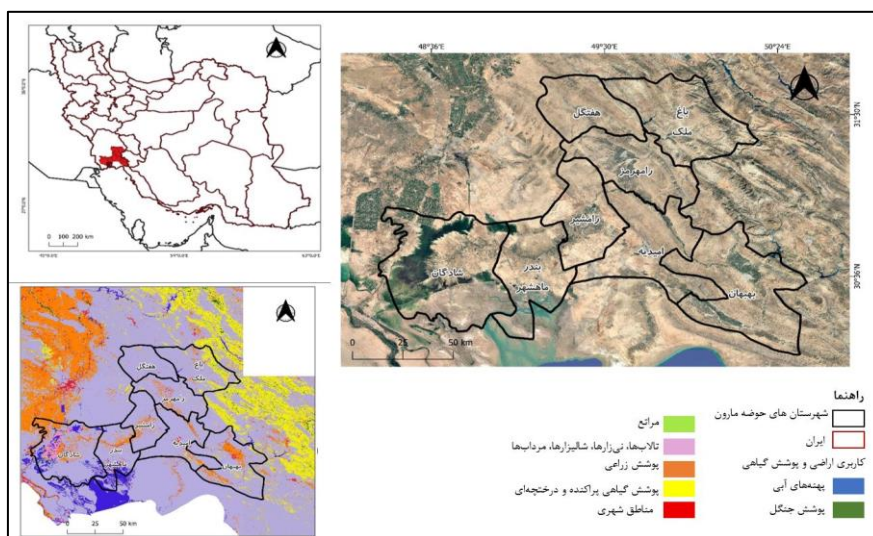
و تضمین امنیت غذایی در دوران تغییرات اقلیمی ایران مورد استفاده قرار گیرد. مطالعه حاضر با هدف رفع خلأ موجود در پایش دقیق و تفکیک اراضی دیم و آبی گندم در مناطق خشک و نیمه‌خشک، یک چارچوب نوین مبتنی بر پردازش ابری ارائه می‌دهد که بر ادغام هم‌زمان داده‌های چندمنبعی و چندزمانه استوار است.

مواد و روش‌ها

منطقه مورد مطالعه

محدوده مورد مطالعه شامل زیرحوضه آبخیز رودخانه مارون (بخشی از حوضه درجه ۱ خلیج فارس)، واقع در جنوب‌غرب ایران و استان خوزستان با محدوده جغرافیایی ۵۰°۰۵' الی ۵۱°۱۱' طول شرقی و ۳۹°۳۰' الی ۳۱°۲۱' عرض شمالی می‌باشد. این زیرحوضه مساحتی در حدود ۲۷۵۰ کیلومتر مربع را در بر می‌گیرد و از نظر توپوگرافی دارای اختلاف ارتفاع قابل توجهی است؛ به‌طوری‌که بیشینه ارتفاع آن به حدود ۳۴۸۳ متر در ارتفاعات زاگرس می‌رسد، در حالی که کمینه ارتفاع در محدوده پایین‌دست و در نزدیکی سد مارون حدود ۳۶۴ متر گزارش شده است (Sharifi paichoon et al., 2019). رودخانه مارون از به‌هم‌پیوستن سرشاخه‌های اصلی سقاوه، لوداب، شور و چاروساق در ارتفاعات رشته‌کوه زاگرس شکل می‌گیرد و پس از طی مسیری طولانی و عبور از تنگ تکاب، وارد دشت بهبهان می‌شود. این رودخانه به‌عنوان یکی از سرشاخه‌های اصلی رود جراحی، نقش حیاتی در تأمین آب مورد نیاز شهرها و روستاهای متعدد، هزاران هکتار اراضی کشاورزی، باغ‌ها، نخلستان‌ها و واحدهای صنعتی منطقه ایفا کرده و در نهایت به تالاب بین‌المللی شادگان منتهی می‌گردد که اهمیت زیست‌محیطی بالایی دارد (سلطانی و همکاران، ۱۴۰۴). بر اساس طبقه‌بندی اقلیمی کوپن (Köppen, 1884)، منطقه مورد مطالعه در پهنه اقلیمی خشک قرار می‌گیرد. همچنین، مطابق گزارش‌های سازمان هواشناسی کشور و طبقه‌بندی دومارتن، اقلیم منطقه خشک تا نیمه‌خشک بوده و با محدودیت‌های جدی منابع آب مواجه است. حداقل، میانگین و حداکثر دمای سالانه منطقه به‌ترتیب برابر با ۶-، ۲۲/۸ و ۵۳/۳ درجه سانتی‌گراد گزارش شده و میانگین بارش سالانه نیز حدود ۶۴۶/۵۱ میلی‌متر برآورد شده است (Adib et al., 2023). در شکل ۱، موقعیت جغرافیایی منطقه مورد مطالعه و وضعیت کاربری اراضی نمایش داده شده است.

مواجهه با چالش تفکیک اراضی آبی و دیم در اقلیم‌های خشک و نیمه‌خشک، که در آن‌ها تفاوت پاسخ طیفی در سال‌های کم‌باران به حداقل می‌رسد، همچنان با عدم قطعیت‌های فضایی روبرو هستند. مسئله تفکیک اراضی آبی و دیم در ایران، به دلیل کوچک و پراکنده بودن قطعات مزارع و تداخل پاسخ‌های طیفی در فصول خشک، پیچیدگی مضاعفی دارد. روش‌های مرسوم مانند موجودی منابع FAO به دلیل تکیه بر آمارهای شهرستان‌محور، فاقد وضوح مکانی برای مدیریت محلی هستند (FAO, 2010). نقشه‌های جهانی ارائه شده توسط USGS و IWMI مانند GLCC و GIMA نیز اگرچه در مقیاس قاره‌ای کاربردی‌اند، اما در مقیاس منطقه‌ای به دلیل دخالت عوامل ذهنی در کلاس‌بندی و عدم کالیبراسیون با شرایط محلی ایران، دچار خطاهای سیستماتیک می‌شوند (Loveland et al., 2009; Thenkabail et al., 2009). رویکردهای نوین مانند آنچه پویا و همکاران با متدولوژی Fuzzy-AHP در استان کردستان پیاده کردند، نشان داد که نسبت عملکرد واقعی به بالقوه در کلاس‌های مختلف تناسب زراعی به چه میزانی تحت تأثیر عوامل مکانی است (Pooya et al., 2025). همچنین محمدی‌منور و زیبازاده (۱۴۰۱) با برجسته کردن نقش شاخص NDWI در شناسایی افت رطوبت پس از اوج رشد، راهکار مؤثری برای تشخیص اراضی دیم ارائه دادند. با این حال، مشکل «پیکسل‌های مختلط» در قطعات کوچک مزارع ایران همچنان به عنوان یک مانع فنی برای آستانه‌گذاری‌های دستی باقی مانده است که تنها با استفاده از داده‌های با وضوح بالا و یادگیری ماشین پیشرفته قابل حل است. در پاسخ به این شکاف‌های متدولوژیک، پژوهش حاضر چارچوبی نوآورانه و چندمنبعی را تدوین کرده است که فراتر از رویکردهای تک‌سجده‌ای متداول عمل می‌کند. در این تحقیق، از تلفیق داده‌های اپتیک سنتینل ۲، داده‌های راداری سنتینل ۱-، دمای سطح زمین^۱ و محصولات جهانی پیشرفته‌ای همچون ESA WorldCover، LGRIP30 و Dynamic World استفاده شده است. این ساختار تلفیقی اجازه می‌دهد تا تفاوت‌های دینامیکی در رفتار فنولوژیکی گندم آبی و دیم، به‌ویژه در مراحل تنش آبی، با دقت بسیار بالایی استخراج شود. برخلاف روش‌های سنتی که به تنظیمات دستی آستانه‌ها وابسته‌اند، این سیستم با بهره‌گیری از الگوریتم جنگل تصادفی، الگوهای تمایز را به صورت خودکار شناسایی کرده و بر محدودیت‌های ناشی از پراکندگی مزارع غلبه می‌کند. توانایی این روش در ارائه نقشه‌های مجزای کشت دیم و آبی در سطح شهرستان‌ها امکان برنامه‌ریزی دقیق منابع آب، سیاست‌گذاری‌های کلان کشاورزی و تصمیم‌سازی‌های اقتصادی را فراهم می‌آورد. این مدل‌سازی نوین می‌تواند به‌عنوان روشی کارآمد برای کشاورزی دقیق



شکل ۱- کاربری اراضی براساس پروداکت Dynamic world و موقعیت جغرافیایی منطقه مطالعاتی (earthengine.google.com)

تصاویر ماهواره‌ای

در این تحقیق به منظور بهره‌برداری مؤثر از مشاهدات ماهواره‌ای، از داده‌های دو مأموریت مهم برنامه کوپرنیکوس^۱ متعلق به سازمان فضایی اروپا استفاده شد. سنیتل-۲ یک سامانه سنجش از دور با حسگرهای چندطیفی نوری است که تصویرهای با رزولوشن مکانی تا ۱۰ متر در باندهای طیفی مرئی، نزدیک به مادون قرمز نزدیک (NIR) و متوسط فروسرخ (SWIR) ارائه می‌دهد و به دلیل پوشش جهانی رایگان و بازبینی تناوبی کوتاه، به طور گسترده برای نظارت بر وضعیت گیاهان و طبقه‌بندی پوشش‌های زمینی به کار گرفته می‌شود (Eisfelder et al., 2024). داده‌های سنیتل-۲ به دلیل توانایی نشان‌دادن اطلاعات طیفی گیاهان و پاسخ آن‌ها در طول رشد، پایه مهمی برای استخراج نماگرهای طیفی مانند NDVI و EVI هستند که در این مطالعه برای متمایزسازی گونه‌های پوشش استفاده شده‌اند (Dobrinic et al., 2021). از سوی دیگر، سنیتل-۱ یک سنجنده راداری با فناوری SAR^۲ در باند C است که می‌تواند در هر شرایط آب و هوایی و زمان روز داده‌های تصویری ارائه کند. مزیت بارز SAR این است که برخلاف سنجش‌های نوری، از ذرات اتمسفر و پوشش ابر مستقل است؛ بنابراین می‌تواند ویژگی‌های ساختار سطح، رطوبت خاک و تابش دستگاه را در تحلیل‌های چندمنبعی بهبود بخشد (Dobrinic et al., 2021). در این تحقیق، اندازه‌گیری‌های قطبش‌های VV و VH از سنیتل-۱ به منظور استخراج اطلاعات ساختاری و بازگشتی ساختار گیاهان برای طبقه‌بندی بهتر محصول‌ها

مورد استفاده قرار گرفت. تلفیق داده‌های سنیتل-۱ و سنیتل-۲ در پژوهش‌های اخیر به عنوان رویکردی قوی برای بهبود دقت طبقه‌بندی محصولات کشاورزی نشان داده شده است، زیرا داده‌های SAR می‌توانند در مواقعی که داده‌های نوری به دلیل ابر و پوشش ابر محدود می‌شوند، جبران اطلاعات از دست‌رفته را انجام دهند (Kordi & Yousefi, 2022). این ترکیب چندمنبعی در مطالعه حاضر با افزودن شاخص‌های طیفی، داده‌های SAR و اطلاعات کاربری اراضی به مدل طبقه‌بندی جنگل تصادفی اعمال شد. در شکل ۲، مراحل انجام کار نشان داده شده است.

داده‌های زمینی

در سال‌های زراعی ۱۴۰۳-۱۴۰۰، داده‌های میدانی از مزارع گندم در شهرستان‌های زیرحوضه مارون-جراحی خوزستان توسط کارشناسان جمع‌آوری شد. در جدول ۱، مزارع منتخب حاصل از بازدیدهای میدانی، با بالاترین مساحت زیرکشت نشان داده شده است.

جدول ۱- مشخصات کلی پایلوت‌های منتخب

شهرستان	نوع محصول	X	Y
امیدیه	گندم	۳۳۸۹۰۹۸	۴۳۱۹۷۰
بهبهان		۳۴۲۰۶۹۵	۳۸۸۵۶۴
رامشیر		۳۴۱۷۷۶۹	۳۳۸۸۷۸
شادگان		۳۳۹۶۶۳۹	۲۷۸۸۱۰
بندر ماهشهر		۳۴۰۶۷۹۸	۳۳۳۴۸۳

1 Copernicus
2 Synthetic Aperture Radar

سطحی کاربرد گسترده در مطالعات محیطی دارند. LST می‌تواند نمایانگر تغییرات فیزیولوژیکی محصولات باشد، چون دما بر تبخیر، تنش گرمایی و نیاز آبی گیاهان تاثیر می‌گذارد و افزودن این متغیر به مدل‌های طبقه‌بندی یا پیش‌بینی عملکرد محصولات می‌تواند دقت را افزایش دهد (Pande et al., 2024). محاسبه LST از داده‌های مادون قرمز حرارتی با تبدیل روشنایی سنجش شده^۱ به دمای سطحی انجام می‌شود که در این پژوهش، این فرآیند با اعمال ضرایب کالبدی و تصحیحات اومیژن انجام گرفت. دمای سطحی به عنوان یک ویژگی ضمنی در مدل یادگیری ماشین افزوده شد تا بتواند اثرات تنش‌های گرمایی و تغییرات دما در پوشش‌های کشاورزی را نیز مدل‌سازی کند، به طوری که ترکیب داده‌های طیفی، راداری و دما، عملکرد نهایی طبقه‌بندی را بهبود بخشید (Z. L. Li et al., 2023; Pande et al., 2024).

محصولات کاربری اراضی

در بخش تحلیل^۲ LULC، از چندین محصول داده جغرافیایی برای ارزیابی و همپوشانی با نتایج طبقه‌بندی استفاده شد. این محصولات شامل داده‌های LGRIP30 برای ارزیابی محدوده‌های آبیاری و غیرآبیاری، ESA WorldCover برای تعیین کلاس‌های کلی پوشش زمین و Dynamic World برای برآورد احتمال حضور پوشش‌های زراعی بودند.

LGRIP30^۳

LGRIP30 یک محصول نقشه‌برداری گسترده از زمین‌های زراعی سراسر جهان با تفکیک به دو گروه اصلی - آبیاری شده و دیم است که با استفاده از تصاویر Landsat 8 و با تفکیک مکانی ۳۰ متر تهیه می‌شود. این محصول بخشی از پروژه Global Food Security-support Analysis Data (GFSAD) است که برای پاسخ‌گویی به چالش‌های امنیت غذایی و آب در قرن ۲۱ طراحی شده است. در این محصول، هر پیکسل ۳۰ متر به چهار حالت کلی تخصیص داده می‌شود: زمین‌های زراعی آبیاری شده (irrigated)، زمین‌های زراعی دیم (rainfed)، غیرزراعی، و تالاب‌ها (مثلاً دریاچه‌ها و رودخانه‌ها) که اطلاعات دقیق مکانی در سطح جهانی را فراهم می‌کند و برای تحلیل کمی مساحت کلاس‌های زراعی مفید است. کاربرد اصلی این محصول در تفکیک انواع سیستم‌های کشاورزی آبی و دیم است؛ زیرا بسیاری از محصولات LULC

برای انجام طبقه‌بندی نظارت‌شده محصولات کشاورزی، نمونه‌های آموزشی و آزمون از داده‌های میدانی انتخاب شدند. نمونه‌های آموزشی صرفاً برای ساخت مدل اولیه طبقه‌بندی مورد استفاده قرار گرفتند، در حالی که نمونه‌های آزمون برای ارزیابی دقت، صحت و قابلیت اعتماد روش‌های طبقه‌بندی به کار رفتند. در این مطالعه، ۷۰٪ از داده‌های میدانی به صورت تصادفی انتخاب و به عنوان داده‌های آموزشی مورد استفاده قرار گرفت و بخش باقی‌مانده برای آزمون اختصاص یافت.

شاخص‌های طیفی و راداری مورد استفاده

شاخص‌های طیفی که از تصاویر سنتینل-۲ استخراج شدند، بر اساس ترکیب باندهای طیفی مختلف تعریف می‌شوند و برای تشخیص وضعیت پوشش گیاهی بسیار کارآمد هستند. مهم‌ترین شاخص‌ها با استفاده از اختلاف بازتاب مادون قرمز نزدیک و قرمز محاسبه می‌شود و سنج‌های از سلامت و تراکم پوشش سبز گیاهان فراهم می‌کند (Rokni & Musa, 2019). همچنین، شاخص‌های ترکیبی دیگر نیز به منظور بهبود جداسازی پوشش‌های متفاوت افزوده شدند که هر کدام اطلاعات متفاوتی از تراکم، رشد و ساختار بیوفیزیکی گیاهان ارائه می‌دهند. از داده‌های راداری سنتینل-۱ نیز علاوه بر باندهای قطبش VV و VH، نسبت VV/VH به عنوان شاخص راداری استخراج شد که می‌تواند به تفکیک بهتر کلاس‌های سطحی کمک کند، به ویژه در شرایطی که شدت بازتاب رادار بسته به آرایش ساختاری گیاهان و آب در گیاه تفاوت می‌کند (Vreugdenhil et al., 2018).

ترکیب داده‌های چندمنبعی و شاخص‌های متنوع مانند آنچه در این مطالعه استفاده شد، باعث بهبود عملکرد مدل‌های طبقه‌بندی مبتنی بر یادگیری ماشین نسبت به استفاده از تنها یک منبع داده شده است، زیرا هر شاخص نمایانگر جنبه متفاوتی از پوشش سطح است که باعث افزایش ظرفیت تفکیک نمونه‌های آموزشی و نتیجه‌گیری دقیق‌تر می‌شود (Eisfelder et al., 2024).

پروداکت‌های سنجش از دوری مورد استفاده

دمای سطح زمین (LST)

دمای سطح زمین (LST) یکی از متغیرهای کلیدی در سنجش از دور است که نشان‌دهنده انرژی تابشی سطح زمین بوده و می‌تواند تأثیر قابل توجهی در رشد و سلامتی محصولات زراعی داشته باشد. در این تحقیق، از تصاویر لندست ۸ (باندهای حرارتی TIRS) برای محاسبه LST استفاده شده است، چرا که این داده‌ها به دلیل کیفیت رادیومتریک بالا و طول دوره زمانی مشاهدات، برای استخراج دمای

1- brightness temperature

2- Landuse-Landcover

3- Landsat-Derived Global Rainfed and Irrigated-Cropland Product 30 m

آن به صورت یک ImageCollection در Google Earth Engine قابل دسترسی هستند (Brown et al., 2022). این محصول با استفاده از شبکه‌های عصبی کانولوشنال و یادگیری عمیق روی داده‌های سنتینل-۲ توسعه داده شده و نه تنها برچسب کلاس نهایی را برای هر پیکسل تولید می‌کند، بلکه احتمالات مربوط به دسته‌بندی کلاس‌ها را نیز ارائه می‌دهد که این امکان را برای تحلیل‌های کمی‌تر و دقیق‌تر فراهم می‌سازد (Small & Sousa, 2023). Dynamic World به دلیل تولید مداوم و به روزرسانی منظم در زمان نزدیک به «زمان واقعی»، برای تحلیل‌های پویا و پایش دوره‌ای تغییرات پوشش زمین بسیار مناسب است. به بیان دیگر، در پروژه‌های کشاورزی که نیاز به پیگیری تغییرات فصلی یا ماهانه پوشش گیاهی دارند، Dynamic World می‌تواند اطلاعات متغیر و زمانی-مکانی را بهتر انعکاس دهد (Brown et al., 2022).

مدل یادگیری ماشین

جنگل تصادفی

مدل جنگل تصادفی یکی از روش‌های پرکاربرد در طبقه‌بندی داده‌های سنجش از دور است که بر اساس مجموعه‌ای از درخت‌های تصمیم‌گیری عمل می‌کند و با ایجاد یک جنگل از درخت‌های مستقل، توانایی تعمیم بسیار خوبی حتی با داده‌های با ابعاد بالا دارد. مزیت اصلی این مدل مقاومت آن در برابر بیش‌برازش و توانایی کار با داده‌های پیچیده و غیرخطی است، به طوری که تعداد زیادی از تحقیقات اخیر در حوزه نقشه‌برداری پوشش زمین و محصولات کشاورزی از آن بهره گرفته‌اند (He et al., 2024). جنگل تصادفی برای ساخت درخت‌ها از مدل درخت تصمیم طبقه‌بندی و رگرسیون کارت^۱ استفاده می‌کند (Breiman, 2002). در این مدل، تقسیم در هر گره بر اساس معیارهایی مانند شاخص جینی^۲ انجام می‌شود. بهره‌گیری از چنین رویکردی موجب افزایش دقت طبقه‌بندی، کاهش خطای مدل و پایداری نتایج در داده‌های پیچیده و ناهمگون می‌شود. به کارگیری این روش موجب شده است که جنگل تصادفی در مقایسه با بسیاری از مدل‌های متداول مانند تحلیل تشخیصی، ماشین‌های بردار پشتیبان و شبکه‌های عصبی، از دقت بالاتری برخوردار باشد (Cutler et al., 2007; Liaw & Wiener, 2002). افزون بر این، جنگل تصادفی از سرعت بالای محاسباتی برخوردار است، در برابر بیش‌برازش مقاومت دارد و امکان ساخت تعداد دلخواهی درخت را برای کاربر فراهم می‌کند (Cutler et al., 2007).

معمول، صرفاً «زمینه زراعی» را نمایش می‌دهند بدون اینکه نوع سیستم آبیاری آن‌ها را تفکیک کنند. LGRIP30 به طور شفاف نشان می‌دهد کدام مناطق به روش‌های آبیاری وابسته‌اند و کدام‌ها به بارش مستقیم، که برای مطالعات بهره‌وری آب، مدیریت منابع و برنامه‌ریزی کشاورزی بسیار ارزشمند است (Teluguntla et al., 2024). در پژوهش‌های اخیر نیز بر اهمیت این نوع تفکیک تأکید شده است، چرا که وضعیت آبیاری در عملکرد محصول، نیازهای آبی و پایداری کشاورزی تأثیر مستقیم دارد و می‌تواند نشانگر حساسیت اکولوژیکی و اقتصادی مناطق مختلف باشد. چنین داده‌هایی به پژوهشگران امکان می‌دهند تا به طور جدی به موضوعاتی مانند بازده محصولات، مصرف آب، و پاسخ به تغییرات اقلیمی بپردازند (مثلاً مطالعاتی در زمینه مدیریت آب و توسعه پایدار)؛ اگرچه خود LGRIP30 صرفاً نقشه‌برداری را انجام می‌دهد و تحلیل ثانویه باید بر اساس آن صورت پذیرد.

ESA WorldCover

محصول WorldCover از سازمان فضایی اروپا یکی از نقشه‌های کاربری اراضی با وضوح مکانی بسیار بالا (۱۰ متر) است که برای سال‌های ۲۰۲۰ و ۲۰۲۱ منتشر شده و از ترکیب داده‌های سنتینل-۱ (رادار) و سنتینل-۲ استفاده می‌کند (Van De Kerchove et al., 2022). این نقشه شامل مجموعه‌ای از کلاس‌های اصلی پوشش زمین مانند:

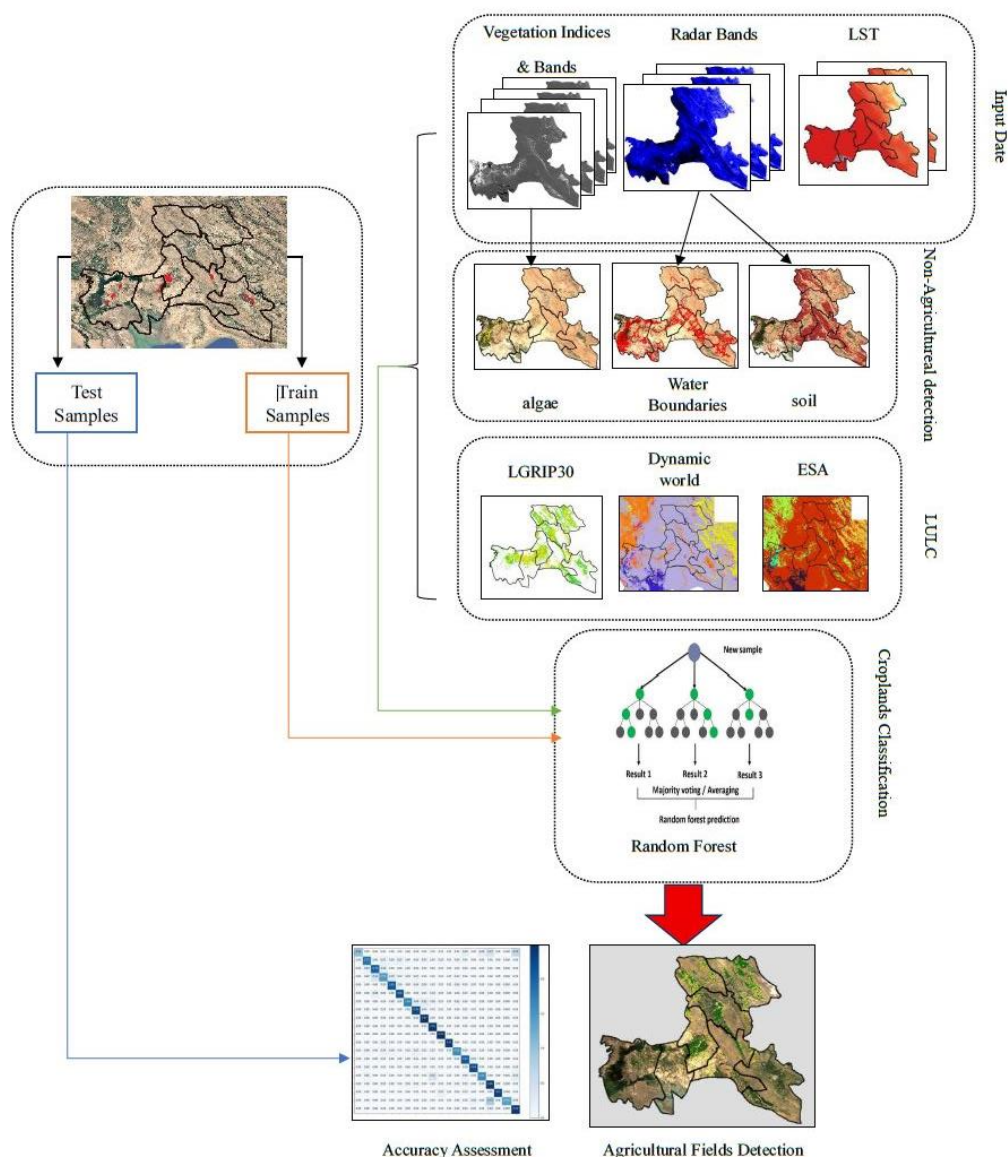
پوشش درختی (Tree Cover)، چمن‌زار، اراضی زراعی (Cropland)، مناطق مسکونی، سطوح آب، پوشش‌های خالی/تنک گیاه و غیره است، که به صورت کلاسه‌بندی صریح در سطح جهان ارائه شده‌اند. این محصول به دلیل دقت مکانی بالا و پشتیبانی از داده SAR (رادار) در کنار سنجش نوری، برای بسیاری از کاربردها از جمله مطالعات تغییر کاربری اراضی، کشاورزی، و محیط‌زیست بسیار کاربرد دارد (Van De Kerchove et al., 2022). در مقالات مقایسه‌ای اخیر، WorldCover در بررسی‌های عملکرد LULC نسبت به نقشه‌های دیگر مانند Dynamic World و Esri Land Cover عملکرد رقابتی و در برخی موارد بهتر در تفکیک «اراضی زراعی» نشان داده است، هرچند که دقت آن بسته به کلاس کاربردی می‌تواند متفاوت باشد و برای برخی اقلیم‌ها لازم است با داده‌های محلی اعتبارسنجی شود (Venter et al., 2022).

Dynamic World

محصول Dynamic World یک داده جهانی، با وضوح ۱۰ متر و به روزرسانی نزدیک به زمان واقعی است که برای کلاس‌بندی کاربری و پوشش زمین (LULC) توسعه یافته است و خروجی‌های

1- CART

2- GINI Index



شکل ۲- فلوچارت روند انجام کار

اهمیت زیادی برخوردار است، زیرا باید میان کاهش همبستگی درخت‌ها و افزایش قدرت پیش‌بینی تعادل برقرار شود (Horning, 2010). بر اساس پیشنهاد بریمن (۲۰۰۲)، انتخاب تعداد متغیرها برابر با جذر تعداد کل متغیرها ($\sqrt{M}$) معمولاً نتایج بهینه‌ای به همراه دارد (Breiman, 2002).

ارزیابی صحت مدل‌های یادگیری ماشین در طبقه‌بندی تصاویر صحت کلی

صحت کلی یکی از شاخص‌های استاندارد برای ارزیابی دقت طبقه‌بندی است و نشان می‌دهد چه درصدی از کل نمونه‌ها به درستی

در اجرای مدل جنگل تصادفی، تعیین دو پارامتر اولیه ضروری است: تعداد درخت‌ها (N) و تعداد متغیرهایی که در هر گره برای انتخاب بهترین تقسیم در نظر گرفته می‌شوند. معمولاً از دو سوم داده‌های آموزشی برای ساخت مدل و از یک سوم باقی‌مانده، موسوم به داده‌های خارج از نمونه^۱، برای برآورد خطای مدل استفاده می‌شود. برای هر نمونه بوت‌استرپ، یک درخت بدون هرس ساخته می‌شود که در هر گره آن، m متغیر به صورت تصادفی انتخاب شده و بهترین تقسیم از میان آن‌ها انجام می‌گیرد. انتخاب تعداد مناسب متغیرها از

1 Out-of-Bag or OOB

می دهد و بیان می کند چه درصدی از نمونه های متعلق به کلاس هدف، توسط مدل به درستی تشخیص داده شده اند:

$$\text{Recall} = \frac{TP}{TP+FN} \quad (۴)$$

FN: تعداد نمونه هایی است که به کلاس هدف تعلق داشته اند اما توسط مدل شناسایی نشده اند (Foody, 2023).

F1-Score

برای ارزیابی جامع تر عملکرد مدل، شاخص F1-Score نیز محاسبه شد که میانگین هارمونیک دقت و بازیابی است. دقت بیانگر درصد نمونه های مثبت پیش بینی شده که واقعاً مثبت بودند و بازیابی نسبت نمونه های واقعی مثبت شناسایی شده توسط مدل را نشان می دهد (Zheng et al., 2024).

معیار F1-Score به صورت میانگین هارمونیک دقت و بازیابی تعریف می شود:

$$F1 = \frac{Precision \times Recall}{Precision + Recall} \quad (۵)$$

ضریب کاپا^۲

ضریب کاپا معیاری آماری است که میزان توافق بین نتایج طبقه بندی و داده های مرجع را اندازه گیری می کند و اثر توافق تصادفی را کاهش می دهد (Feizizadeh et al., 2022a). مقدار کاپا بین ۱- و ۱+ نوسان می کند، به گونه ای که $\kappa=1$ نشان دهنده توافق کامل، $\kappa=0$ نشان دهنده توافق معادل با شانسی و مقادیر منفی نشان دهنده توافق کمتر از شانسی است.

$$Kappa = \frac{P_0 - P_e}{1 - P_e} \quad (۶)$$

طبق رابطه ۶، در آن:

P_0 : برابر است با تعداد نمونه های به درستی طبقه بندی شده تقسیم بر کل تعداد نمونه ها

P_e : برابر است با مجموع حاصل ضرب های تعداد واقعی نمونه ها و تعداد پیش بینی شده در هر دسته، تقسیم بر مجذور کل تعداد نمونه ها (Zhan et al., 2024).

نتایج و بحث

ارزیابی اهمیت شاخص ها در مدل های طبقه بندی زمین کشاورزی، به ویژه برای شناسایی مزارع گندم، یکی از بخش های کلیدی تحلیل داده های سنجش از دور است. نتایج نشان می دهد که شاخص های مختلف شامل شاخص های طیفی، راداری و حرارتی نه تنها اطلاعات متمایزی درباره ویژگی های سطح زمین ارائه می دهند بلکه در کنار هم موجب افزایش دقت و پایداری مدل می شوند. همانطور که در

طبقه بندی شده اند. این شاخص با جمع مقادیر قطر اصلی ماتریس ابهام تقسیم بر تعداد کل نمونه ها محاسبه می شود (Aziz et al., 2024). صحت کلی به صورت رابطه ۱، تعریف می شود.

$$\text{Overall Accuracy} = \frac{N_{\text{Correct}}}{N_{\text{total}}} \quad (۱)$$

N_{Correct} : تعداد نمونه های درستی است که به درستی طبقه بندی شده اند، N_{total} : کل تعداد نمونه ها است.

ماتریس درهم ریختگی^۱

ماتریس ابهام، ماتریسی است که تعداد نمونه های پیش بینی شده برای هر کلاس را در مقابل مقادیر واقعی آن ها نمایش می دهد. در این ماتریس، معمولاً سطرها نشان دهنده کلاس های پیش بینی شده و ستون ها نمایانگر کلاس های واقعی هستند، و مقادیر سلول های قطر اصلی (i, i) تعداد نمونه هایی را نشان می دهند که به درستی طبقه بندی شده اند (Nicolau et al., 2024). دلیل الگوی خطاها در این ماتریس امکان شناسایی کلاس هایی که اغلب با یکدیگر اشتباه گرفته شده اند را فراهم می کند (Feizizadeh et al., 2022b). ساختار کلی ماتریس ابهام به صورت رابطه ۲، تعریف می شود:

$$\begin{bmatrix} \text{False Positive (FP)} & \text{True Positive (TP)} \\ \text{True Negative (TN)} & \text{False Negative (FN)} \end{bmatrix} \quad (۲)$$

که در آن:

True Positive (TP): تعداد نمونه هایی که به درستی به کلاس مورد نظر اختصاص یافته اند.

False Positive (FP): تعداد نمونه هایی که به نادرست به کلاس مثبت اختصاص یافته اند.

False Negative (FN): تعداد نمونه هایی که متعلق به کلاس هدف بوده اند اما توسط مدل شناسایی نشده اند.

True Negative (TN): تعداد نمونه هایی که به درستی به کلاس نادرست طبقه بندی شده اند.

Precision

دقت معیاری برای سنجش توانایی مدل در پیش بینی درست نمونه های مثبت است و نشان می دهد از میان نمونه هایی که مدل به عنوان متعلق به یک کلاس خاص پیش بینی کرده است، چه نسبتی واقعاً به آن کلاس تعلق دارند (Foody, 2023).

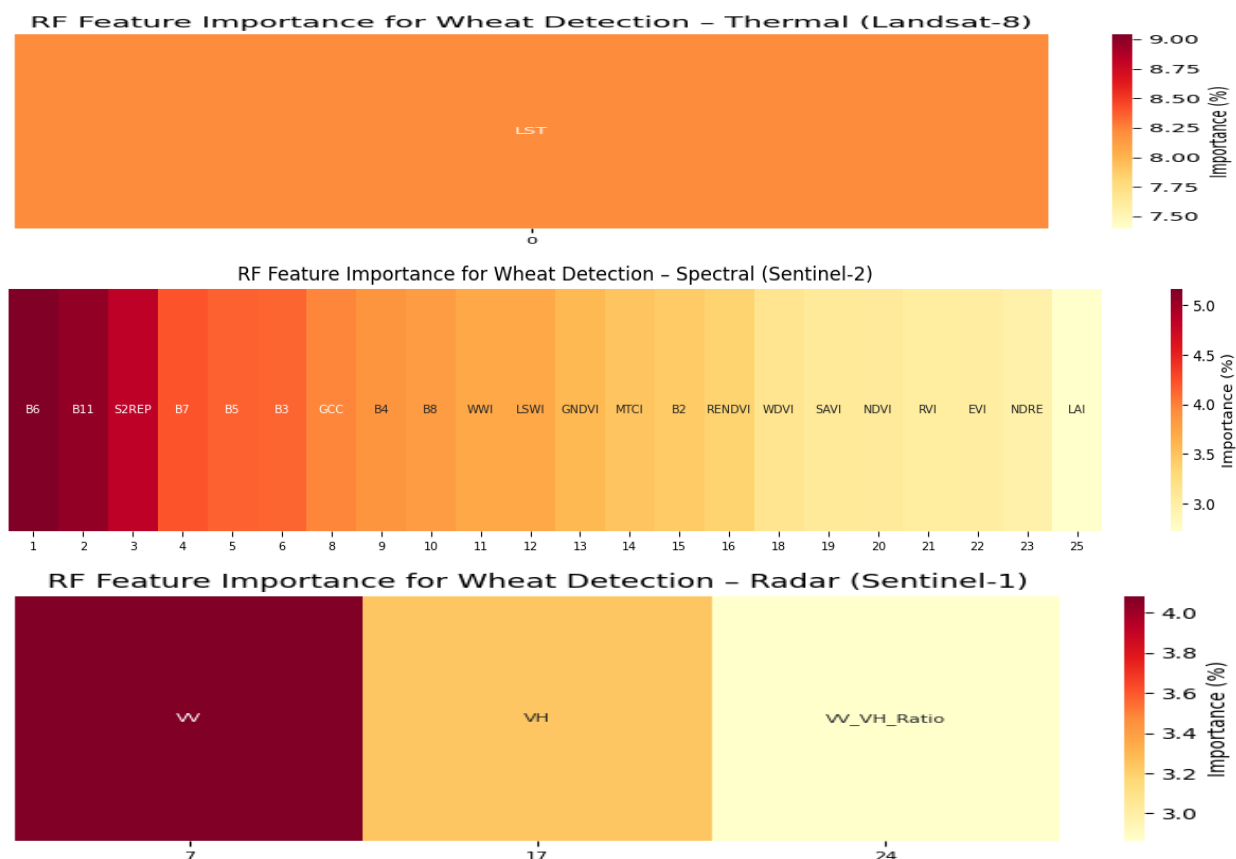
$$\text{Precision} = \frac{TP}{TP+FP} \quad (۳)$$

Recall

توانایی مدل در شناسایی نمونه های واقعی یک کلاس را نشان

که در شکل ۳، نشان داده شده است، شاخص‌های طیفی (Spectral Indices)، راداری (Radar Backscatter) و حرارتی (Land Surface Temperature) هر یک نقش متفاوت اما مکملی در تشخیص گندم داشته‌اند. در ادامه هر یک از این گروه‌ها به صورت مجزا تحلیل می‌شود.

مطالعات پیشین نیز نشان داده شده است، استفاده از ترکیب منابع داده‌های سنجنش از دور، مخصوصاً زمانی که این داده‌ها نمایانگر جنبه‌های مختلف فیزیولوژیکی و ساختاری پوشش گیاهی باشند، می‌تواند تفکیک کلاس‌ها را بهبود بخشد و دقت کلی مدل را افزایش دهد (Eisfelder et al., 2024; Soltani et al., 2026; Zaka & Samat, 2024; Zheng et al., 2024). در این تحقیق، همان‌گونه



شکل ۳- درصد عملکرد شاخص برتر در گروه طیفی، راداری و دمایی در شناسایی

(Tsuchiya & Sonobe, 2025). نتایج ارزیابی اهمیت متغیرها نشان می‌دهد که شاخص‌های راداری سهم قابل توجهی در تشخیص مزارع گندم دارند. در این میان، باند VV با درصد اهمیت حدود ۴۱٪ بیشترین نقش را به خود اختصاص داده است و پس از آن باند VH با ۳۳٪ و نسبت VV/VH با ۲۹٪ قرار دارند. حساسیت بالای باند VV به تغییرات ساختاری سطح زمین، زبری سطح، تراکم پوشش گیاهی و رطوبت خاک موجب شده است که این باند در مراحل مختلف رشد گندم، به‌ویژه در فازهای پنجه‌زنی و ساقه‌دهی، بازپراکنش بالاتری ثبت کند. ساختار عمودی گیاه گندم در این مراحل سبب تقویت سیگنال VV شده و امکان تفکیک بهتر این محصول از سایر کاربری‌های اراضی را فراهم می‌سازد. باند VH نیز که به

شاخص‌های راداری

شاخص‌های راداری استخراج شده از داده‌های سنتینل-۱، از جمله باندهای VV و VH و نسبت‌های قطبی‌سازی شده مانند VV/VH، اطلاعات ارزشمندی درباره ساختار فیزیکی، زبری سطح، رطوبت خاک و معماری پوشش گیاهی فراهم می‌کنند. مزیت اصلی داده‌های راداری، استقلال آن‌ها از شرایط نوری و پوشش ابری است؛ به‌گونه‌ای که حتی در شرایط ابری یا بارندگی، امکان پایش پیوسته سطح زمین و مزارع کشاورزی فراهم می‌شود. داده‌های SAR با ثبت الگوهای بازپراکنش امواج راداری، قادرند ویژگی‌هایی را آشکار سازند که در تصاویر نوری به‌وضوح قابل تشخیص نیستند و از این رو نقش مهمی در تمایز مزارع گندم از سایر کلاس‌های پوشش زمینی ایفا می‌کنند

نمایش ویژگی‌های بیوفیزیکی و بیوشیمیایی گیاه، ستون فقرات سامانه‌های پایش و طبقه‌بندی محصولات کشاورزی محسوب می‌شوند.

شاخص‌های حرارتی

شاخص‌های حرارتی، به‌ویژه دمای سطح زمین (LST)، اطلاعات تکمیلی ارزشمندی را فراهم می‌کنند که در داده‌های نوری و راداری به‌طور مستقیم قابل مشاهده نیست. دمای سطح زمین به فرآیندهایی نظیر تبخیر و تعرق، رطوبت خاک و وضعیت آبی گیاه وابسته است و می‌تواند به‌عنوان شاخصی از استرس محیطی و شرایط فیزیولوژیکی گیاه مورد استفاده قرار گیرد (Chen et al., 2020). اگرچه LST به‌تنهایی برای شناسایی نوع پوشش گیاهی کافی نیست، اما ترکیب آن با شاخص‌های طیفی و راداری موجب بهبود تفکیک‌پذیری بین کلاس‌هایی می‌شود که از نظر طیفی شباهت زیادی دارند اما از لحاظ حرارتی متفاوت‌اند؛ مانند تمایز گندم آبی و گندم دیم در دوره‌های خشک و مرطوب.

بر اساس نتایج به‌دست‌آمده، شاخص دمای سطح زمین با درصد اهمیت حدود ۸٪ سهم قابل توجهی در مدل طبقه‌بندی داشته است. این میزان اهمیت نشان‌دهنده نقش مؤثر اطلاعات حرارتی در تشخیص مزارع گندم است. در مزارع فعال گندم، به دلیل تعرق بالا و پوشش گیاهی متراکم، دمای سطح زمین معمولاً کمتر از زمین‌های بایر یا مناطق شهری است و همین اختلاف حرارتی، LST را به شاخصی قدرتمند برای جداسازی اراضی زراعی تبدیل می‌کند (Xiao et al., 2024). افزون بر این، تغییرات دمایی در طول فصل رشد گندم الگوی زمانی مشخصی ایجاد می‌کند که می‌تواند برای تشخیص مراحل فیزیولوژیک مورد استفاده قرار گیرد. بنابراین، ادغام شاخص‌های حرارتی با داده‌های طیفی و راداری، نقش بسزایی در افزایش دقت و پایداری نتایج طبقه‌بندی دارد و LST به‌عنوان نماینده شرایط انرژی و وضعیت آبی گیاه، نقش مکمل مهمی در شناسایی مزارع گندم ایفا می‌کند.

بررسی دقت مدل جنگل تصادفی در شناسایی مزارع گندم و برآورد سطح زیر کشت

در این پژوهش، عملکرد الگوریتم جنگل تصادفی در تشخیص و تفکیک مزارع گندم از دیگر کلاس‌های پوشش زمین مورد ارزیابی قرار گرفت. نتایج حاصل از ماتریس درهم‌ریختگی (شکل ۴) و شاخص‌های صحت‌سنجی (جدول ۲) برای مجموعه داده‌های آموزش و تست به‌طور جداگانه استخراج شد تا قابلیت تعمیم‌پذیری و صحت مدل بررسی گردد.

پراکندگی حجمی درون تاج گیاه حساس است، اطلاعات ارزشمندی درباره زیست‌توده و معماری پوشش گیاهی ارائه می‌دهد. ترکیب همزمان باندهای VV و VH توانایی مدل را در تشخیص تفاوت‌های فیزیولوژیک گندم افزایش داده و دقت طبقه‌بندی را بهبود می‌بخشد (Yang et al., 2024). نسبت VV/VH نیز به‌عنوان شاخصی مکمل، با کاهش اثرات زاویه دید، تغییرات رطوبتی و ناهمگنی سطح زمین، موجب تقویت سیگنال‌های مرتبط با پوشش گیاهی و کاهش نویزهای محیطی می‌شود (H. Y. Li et al., 2023). به‌طور کلی، داده‌های راداری به دلیل پایداری زمانی و استقلال از شرایط جوی، نقش کلیدی در پایش و شناسایی مزارع گندم، به‌ویژه در مناطق ابری و مرطوب، ایفا کرده و موجب افزایش معنادار دقت طبقه‌بندی می‌شوند.

شاخص‌های طیفی

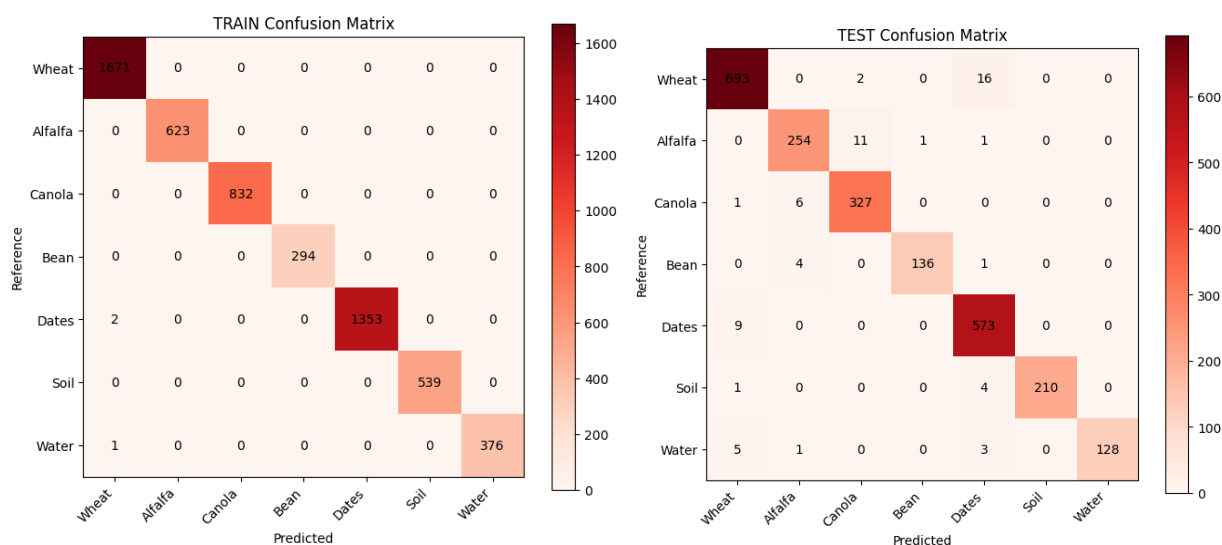
شاخص‌های طیفی استخراج‌شده از داده‌های سنتینل-۲، از جمله EVI، GNDVI، NDVI و سایر شاخص‌های پوشش گیاهی، اصلی‌ترین مؤلفه‌های مورد استفاده در مدل‌های طبقه‌بندی محصولات کشاورزی به شمار می‌روند. این شاخص‌ها بر پایه بازتاب طیفی گیاه در باندهای مرئی، فروسرخ نزدیک و فروسرخ میانی طراحی شده‌اند و ارتباط مستقیمی با وضعیت زیستی، تراکم برگ، میزان کلروفیل و فعالیت فیزیولوژیکی گیاه دارند (Luo, 2025).

نتایج نشان می‌دهد که شاخص‌ها و باندهای طیفی بیشترین سهم را در افزایش دقت تشخیص مزارع گندم داشته‌اند. در این میان، باندهای B6 (Red Edge)، B11 (SWIR) و شاخص S2REP با درصد اهمیت بالاتر از ۴۵٪ در صدر قرار گرفته‌اند و پس از آن باندهای B3، B5، B7 و شاخص GCC قرار دارند. باندهای B6، B7 و Rededge به تغییرات کلروفیل و وضعیت فیزیولوژیک گیاه بسیار حساس هستند و توان بالایی در تشخیص تفاوت‌های ظریف میان محصولات زراعی مختلف دارند. اهمیت بالای این باندها نشان می‌دهد که متغیرهای مرتبط با فتوسنتز و فعالیت زیستی گیاه، نقش اساسی در تفکیک مزارع گندم ایفا می‌کنند. باند SWIR (B11) نیز به رطوبت خاک و پوشش گیاهی حساس بوده و تغییرات محتوای آب برگ و خاک را در طول فصل رشد گندم به‌خوبی منعکس می‌کند. از این رو، این باند نقش مهمی در تشخیص مزارع فعال زراعی و تفکیک آن‌ها از اراضی بایر یا کم‌پوشش دارد (Dash et al., 2010). اگرچه شاخص‌های پوشش گیاهی مانند NDVI، EVI، SAVI، RVI و GNDVI به‌صورت جداگانه درصد اهمیت کمتری نسبت به باندهای خام نشان داده‌اند، اما در مجموع نقش مکمل بسیار مؤثری در پایش وضعیت زیستی گیاه و شناسایی الگوهای فیزیولوژیک گندم دارند و به‌ویژه در تمایز مراحل مختلف رشد این محصول مؤثر هستند (Bade et al., 2025). شاخص‌های طیفی به دلیل توانایی بالا در

ماتریس درهم‌ریختگی

نمونه‌های گندم تقریباً به‌صورت کامل و صحیح در کلاس مورد نظر قرار گرفته‌اند. چنین الگوی طبقه‌بندی، نشان‌دهنده توانمندی بالای جنگل تصادفی در شناسایی الگوهای طیفی و فضایی گندم است. ماتریس‌های درهم‌ریختگی چون در شکل‌های مربوطه قابل مشاهده‌اند، به‌وضوح نشان می‌دهند که اغلب اشتباهات طبقه‌بندی در حد بسیار اندکی میان کلاس‌هایی مانند گندم، یونجه و سایر محصولات بوده‌اند، ولی این اشتباهات برای گندم در سطح پایین و قابل قبول قرار دارند.

ماتریس درهم‌ریختگی برای مجموعه تست نشان می‌دهد که بیشترین تعداد پیکسل‌های واقعی گندم توسط مدل به‌درستی تشخیص داده شده‌اند، در حالی که تعداد خطاهای طبقه‌بندی برای این کلاس کم بوده است. به طور خاص، از مجموع نمونه‌های واقعی گندم، تنها تعداد اندکی به کلاس‌های دیگر منتسب شده‌اند و بخش عمده نمونه‌های گندم در موقعیت صحیح خود قرار گرفته‌اند. این الگو در ماتریس مربوط به مجموعه آموزش نیز دیده می‌شود، جایی که



شکل ۴- ماتریس درهم‌ریختگی داده‌های آموزش و آزمون پیکسل‌های مزارع گندم

مرجع فراتر از تصادف است. این نتایج نشان می‌دهند که مدل جنگل تصادفی توانسته است به‌صورت همزمان هم حساسیت بالا در تشخیص کلاس گندم و هم دقت بالا در جلوگیری از اشتباهات طبقه‌بندی داشته باشد. به عبارتی، مدل نه فقط در تشخیص نمونه‌های واقعی گندم دقیق بوده است، بلکه اشتباه اندکی در تخصیص برچسب «گندم» به نمونه‌های غیرگندم انجام داده است.

مقادیر به‌دست‌آمده برای دقت کلی و ضریب کاپا در حدود ۰/۹۷ با نتایج گزارش‌شده در مطالعات اخیر در زمینه طبقه‌بندی محصولات زراعی مطابقت دارد. به عنوان نمونه، Yao et al. (2022) نشان دادند که ترکیب RF با روش‌های دیگر می‌تواند به دقت کلی ۰/۹۸ و کاپا ۰/۹۷ دست یابد، که نشان‌دهنده توانمندی بالای الگوریتم در استخراج کلاس‌های محصولات است (Yao et al., 2022). در یک مطالعه دیگر که با استفاده از سنتینل-۲ و الگوریتم‌های یادگیری ماشین طبقه‌بندی گندم انجام شد، Overall Accuracy $\approx$ ۰/۹۲۵ و Kappa $\approx$ ۰/۹۲۵ برای نقشه‌برداری گندم گزارش شده است (Luo, 2025). همچنین در مطالعاتی دیگر نشان داده شد که با RF و ترکیب ویژگی‌های چندمنبع شاخص‌های گیاهی و ویژگی‌های طیفی (، برای

به‌منظور سنجش دقیق عملکرد مدل در شناسایی گندم، سه شاخص اصلی دقت کاربر، بازیابی و امتیاز F1 استخراج شد. همچنین دقت کلی و ضریب کاپا برای سنجش عملکرد کلی مدل گزارش گردید. Precision برای کلاس گندم برابر با ۰/۹۷۷۴ است، که نشان می‌دهد از میان تمام پیکسل‌هایی که مدل به‌عنوان گندم پیش‌بینی کرده، تقریباً ۹۷/۷٪ واقعاً متعلق به گندم بوده‌اند. این مقدار بالای Precision نشان‌دهنده خطای کمسیون بسیار کم (False Positives) در تشخیص گندم است. Recall (بازیابی/حساسیت) برای گندم برابر با ۰/۹۷۴۷ بوده است، که بیانگر آن است که از میان تمام پیکسل‌های واقعی گندم، بیش از ۹۷٪ به‌درستی توسط مدل تشخیص داده شده‌اند. این موضوع نشان می‌دهد که مدل به‌خوبی از پدیدارشناسی گندم بهره برده و خطای حذف (False Negatives) ناچیزی داشته است. F1-score که میانگین هارمونیک Precision و Recall است، برابر با ۰/۹۷۶۱ بوده که نمایانگر تعادل بسیار مطلوب بین حساسیت و دقت برای کلاس گندم است. دقت کل برای مجموعه تست برابر با ۰/۹۷۲۴ و ضریب کاپا برابر با ۰/۹۶۵۶ بوده است، که نشان‌دهنده توافق بسیار خوب بین طبقه‌بندی مدل و داده

تشخیص گندم در دوره رشد و شاخص‌های زمانی مختلف ۰/۹۶
 $\geq$ Overall Accuracy و $\geq$ Kappa قابل دستیابی می‌باشد
 که با نتایج بدست آمده در این پژوهش همسو است. تفاوت اندک
 بین نتایج مجموعه آموزش و مجموعه تست در شاخص‌های دقت،
 نشان‌دهنده این است که مدل نه بیش‌برازش شده است و نه دچار
 افت عملکرد زیادی در داده‌های جدید و مستقل شده است. این رفتار
 تداعی‌کننده پایداری و تعمیم‌پذیری مدل در شرایط عملی/آزمایشی

است. در مجموع، مدل جنگل تصادفی به‌طور مؤثر توانسته است
 مزارع گندم را از دیگر کلاس‌های پوشش زمین در تصاویر ماهواره‌ای
 تفکیک کند. مقادیر بالای Precision، Recall، F1-score،
 Overall Accuracy و Kappa برای گندم بیانگر این است که مدل
 نه فقط در تشخیص درست پیکسل‌های گندم عملکرد قابل قبولی
 دارد، بلکه تعمیم‌پذیری آن نیز در داده‌های تست مستقل حفظ
 می‌شود.

جدول ۲- نتایج صحت مدل‌های یادگیری ماشین در شناسایی مزارع گندم

مدل یادگیری ماشین	داده‌های آموزش					داده‌های تست				
	Precision	Recall	F1 score	Overall Accuracy	Kappa Coefficient	Precision	Recall	F1 score	Overall Accuracy	Kappa Coefficient
جنگل تصادفی	۰/۹۷	۰/۹۷	۰/۹۷	۰/۹۹	۰/۹۹	۰/۹۷	۰/۹۷	۰/۹۷	۰/۹۷	۰/۹۶

مطابق جدول ۳، نتایج برآورد سطح زیرکشت مزارع گندم دیم و
 آبی در حوضه مارون-جراحی نشان داد که مدل جنگل تصادفی
 توانسته است سطح مزارع گندم را با دقت مطلوبی تخمین بزند. مقدار
 واقعی سطح مزارع گندم طبق آمارنامه وزارت جهادکشاورزی،
 ۱۳۵۶۵۸ هکتار و مقدار برآوردی ۱۴۰۵۴۳ هکتار به دست آمد که
 منجر به بایاس ۰/۰۳۴+ (معادل بیش‌برآوردی حدود ۳ درصد) شد.
 این مقدار خطا در مقیاس حوضه نشان‌دهنده دقت بالا و قابلیت اعتماد
 مدل در شناسایی نواحی کشت گندم است. برآوردهای جداگانه نیز
 نشان می‌دهد که سطح گندم دیم حدود ۴۵،۱۷۴ هکتار و گندم آبی
 حدود ۹۵،۳۶۹ هکتار برآورد شده است که مقادیر برآوردی بسیار
 نزدیک به مقادیر واقعی هستند و توانایی مدل در تمایز بین دو نوع

کشت را تأیید می‌کند، در شکل ۵ جانمایی مجزای مزارع دیم و آبی
 نشان داده شده است. دقت بالای مدل را می‌توان به قدرت تفکیک
 بالای الگوریتم جنگل تصادفی در شناسایی الگوهای پیچیده و
 غیرخطی داده‌های چندمنبعی، استفاده از شاخص‌های طیفی، راداری و
 دمای سطح زمین، و انتخاب نمونه‌های آموزشی و آزمون نماینده
 نسبت داد. بایاس مثبت اندک مشاهده شده ممکن است ناشی از
 تشخیص برخی زمین‌های کوچک یا پراکنده به عنوان مزارع گندم
 باشد که در مقیاس حوضه تأثیر قابل توجهی ندارد. نتایج نشان
 می‌دهند که مدل جنگل تصادفی با استفاده از داده‌های چندمنبعی،
 ابزار موثری برای پایش و برآورد سطح زیرکشت مزارع گندم است و
 قابلیت کاربرد در برنامه‌ریزی کشاورزی و مدیریت منابع آب را دارد.

جدول ۳- مقایسه سطح زیر کشت برآورده شده براساس مدل‌های یادگیری ماشین و مقدار واقعی

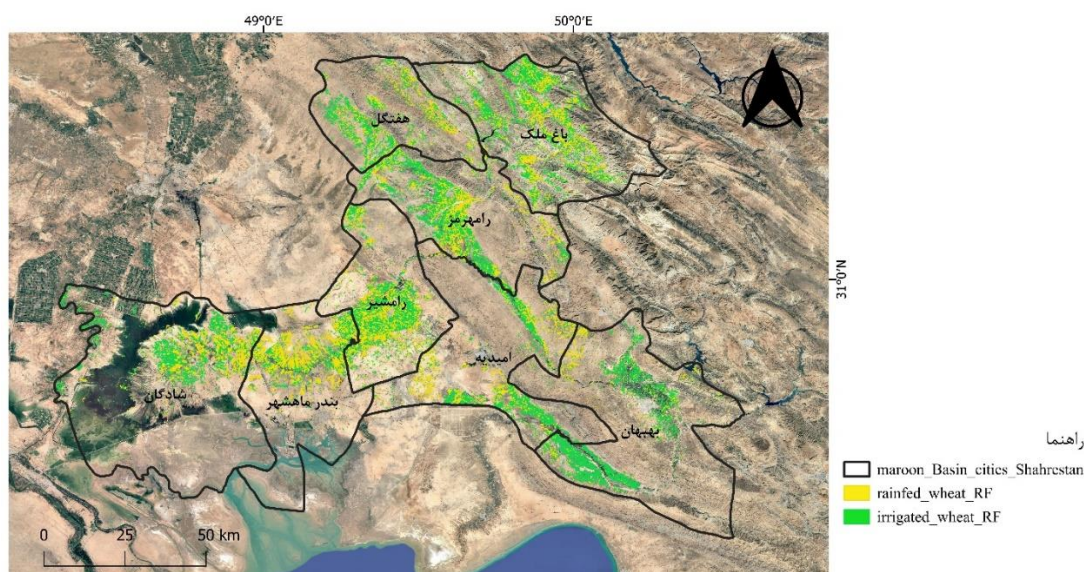
بایاس	گندم	سطح زیرکشت محاسباتی
	۱۳۵۶۵۸	سطح زیرکشت واقعی
۰/۰۳۴	گندم دیم: ۴۵۱۷۴/۷۹	سطح زیرکشت براساس مدل جنگل تصادفی
	گندم آبی: ۹۵۳۶۹/۱۵	
	مجموع: ۱۴۰۵۴۳	

نتیجه‌گیری

نتایج این پژوهش نشان داد که چارچوب تلفیقی مبتنی بر
 داده‌های اپتیکی سنتینل ۲، داده‌های راداری سنتینل ۱- و متغیر دمای
 سطح زمین، در ترکیب با الگوریتم جنگل تصادفی، توانایی بالایی در
 شناسایی و تفکیک مزارع گندم دیم و آبی در حوضه مارون دارد.
 عملکرد بالای مدل با دستیابی به دقت کلی و ضریب کاپای حدود ۹۷
 و ۹۶ درصد و همچنین بایاس اندک (حدود ۳ درصد) در برآورد سطح

زیرکشت، برتری الگوریتم جنگل تصادفی را در پردازش داده‌های
 چندمنبعی و استخراج الگوهای غیرخطی تأیید می‌کند. با این حال،
 وابستگی نتایج به کیفیت نمونه‌های آموزشی، وجود پیکسل‌های
 مختلط در مزارع کوچک و محدودیت‌های تفسیری مدل‌های یادگیری
 ماشین از مهم‌ترین چالش‌های این تحقیق محسوب می‌شوند. بر این
 اساس، پیشنهاد می‌شود در پژوهش‌های آتی از داده‌های با تفکیک
 مکانی و زمانی بالاتر، مدل‌های یادگیری عمیق مبتنی بر سری‌های

زمانی و تلفیق اطلاعات سنجش از دور با داده‌های میدانی و هواشناسی استفاده شود تا دقت، پایداری و تفسیرپذیری نقشه‌های



شکل ۵- موقعیت مزارع گندم دیم و آبی حوضه مارون-جراحی براساس مدل جنگل تصادفی

آمارنامه کشـاورزی جـلد اول.

<https://dpe.maj.ir/%D8%A2%D9%85%D8%A7%D8%B1%D9%86%D8%A7%D9%85%D9%87-%D9%87%D8%A7>

Abbaspour, K. C., Faramarzi, M., Ghasemi, S. S. and Yang, H. 2009. Assessing the impact of climate change on water resources in Iran. *Water Resources Research*. 45(10).

<https://doi.org/10.1029/2008WR007615>

Adib, A., Haidari, B., Lotfifard, M. and Sasani, H. 2023. Evaluating climatic change effects on EC and runoff in the near future (2020–2059) and far future (2060–2099) in arid and semi-arid watersheds. *Applied Water Science*. 13(6): 122.

<https://doi.org/10.1007/s13201-023-01926-1>

Aobpaet, A. 2022. Monitoring of crop growth stages using Sentinel-1 synthetic aperture radar data. *Agriculture and Natural Resources*. 56(2): 307–320.

<https://li01.tci-thaijo.org/index.php/anres/article/view/254483>

Araghi, A. and Daccache, A. 2025. Remote sensing and TerraClimate datasets for wheat yield prediction using machine learning. *Smart Agricultural Technology*. 11: 100909.

<https://doi.org/10.1016/j.atech.2025.100909>

Aziz, G., Minallah, N., Saeed, A., Frnda, J. and Khan, W. 2024. Remote sensing based forest cover classification using machine learning. *Scientific Reports*. 14(1): 69. <https://doi.org/10.1038/s41598-023-50863-1>

Bade, B., Schulthess, U., Blasch, G., Sherpa, S. R.,

منابع

سلطانی، م. و بهمن آبادی، ب. ۱۴۰۴. تحلیل عملکرد الگوریتم‌های یادگیری ماشین در طبقه‌بندی تصاویر چند زمانه در مدیریت کشاورزی (مطالعه موردی: دشت قزوین). مدل سازی و مدیریت آب و خـ. اک. (۳): ۵۴–۷۳.

<https://doi.org/10.22098/mmws.2025.16425.1532>

سلطانی، م. و بهمن آبادی، ب. مختاران، ع. (۱۴۰۴). طبقه‌بندی محصولات کشاورزی و برآورد سطح کشت در مقیاس حوضه‌ای با رویکرد یادگیری ماشین. تحقیقات آب و خاک ایران. ۵۶(۱۱): ۳۱۲۹–۳۱۵۶.

<https://doi.org/10.22059/ijswr.2025.402830.670012>

محمدی منور، ح، زیبازاده، س. (۱۴۰۱). تفکیک مزارع کشت دیم و آبی در استان همدان با استفاده از شاخص‌های طیفی تصاویر ماهواره‌ای. ماشین‌های کشاورزی. ۱۲(۴): ۵۲۹–۵۴۲.

<https://doi.org/10.22067/jam.2021.69074.1022>

موسوی، س. تهرانی، ن. جانعلی‌پور، م. (۱۳۹۹). برآورد سطح زیر کشت گندم دیم با استفاده از تصاویر ماهواره‌ای سنتینل-۲، مطالعه موردی: منطقه سجاس رود شهرستان خدابنده، استان زنجان. پژوهش و فناوری محیط زیست. ۵(۷): ۷۷–۹۰.

<https://sid.ir/paper/373696/fa>

وزارت جهاد کشاورزی. (۱۴۰۳). آمار نامه سال زراعی ۱۴۰۲–۱۴۰۳.

- wheat yield in fields using remotely sensed phenology, climate data and machine learning. *Remote Sensing*. 13(13): 2435. <https://doi.org/10.3390/rs13132435>
- FAO, F. a. A. O. o. t. U. N. 2010. Provincial Land Cover Atlas of Islamic Republic of Afghanistan. Rome, Italy.
- Feizizadeh, B., Darabi, S., Blaschke, T. and Lakes, T. 2022a. QADI as a New Method and Alternative to Kappa for Accuracy Assessment of Remote Sensing-Based Image Classification. *Sensors*. 22(12): 4506. <https://www.mdpi.com/1424-8220/22/12/4506>
- Feizizadeh, B., Darabi, S., Blaschke, T. and Lakes, T. 2022b. QADI as a New Method and Alternative to Kappa for Accuracy Assessment of Remote Sensing-Based Image Classification. *Sensors*. 22(12).
- Foody, G. M. 2023. Challenges in the real world use of classification accuracy metrics: From recall and precision to the Matthews correlation coefficient. *Plos one*. 18(10): e0291908.
- Ghosh, A., Sharma, R. and Joshi, P. K. 2014. Random forest classification of urban landscape using Landsat archive and ancillary data: Combining seasonal maps with decision level fusion. *Applied Geography*. 48: 31-41. <https://doi.org/10.1016/j.apgeog.2014.01.003>
- Guan, X., Chen, J. M., Shen, H. and Xie, X. 2021. A modified two-leaf light use efficiency model for improving the simulation of GPP using a radiation scalar. *Agricultural and Forest Meteorology*. 307: 108546. <https://doi.org/10.1016/j.agrformet.2021.108546>
- He, J., Zeng, W., Ao, C., Xing, W., Gaiser, T. and Srivastava, A. K. 2024. Cross-Regional Crop Classification Based on Sentinel-2. *Agronomy*. 14(5): 1084. <https://www.mdpi.com/2073-4395/14/5/1084>
- Heydari, N. and Taran, F. 2025. Effect of Climate Change on Wheat Yield and Water Productivity in Iran and the World. *Iranica Journal of Energy & Environment*. 16(2): 253-269. <https://doi.org/10.5829/ijee.2025.16.02.08>
- Horning, N. 2010. Random Forests: An algorithm for image classification and generation of continuous fields data sets. *Proceedings of the International Conference on Geoinformatics for Spatial Infrastructure Development in Earth and Allied Sciences*, Osaka, Japan.
- Huang, J., Ma, H., Sedano, F., Lewis, P., Liang, S., Wu, Q., . . . Zhu, D. 2019. Evaluation of regional estimates of winter wheat yield by assimilating three remotely sensed reflectance datasets into the coupled WOFOST-PROSAIL model. *European Journal of Agronomy*. 102: 1-13. <https://doi.org/10.1016/j.eja.2018.10.008>
- Kamir, E., Waldner, F. and Hochman, Z. 2020. Estimating wheat yields in Australia using climate Shrestha, H. L., Bajimaya, D., . . . Nepali, S. 2025. Temporal Sentinel-2 Imagery for Wheat mapping and monitoring: Analyzing Phenological Stages with Machine Learning to Improve Mapping Precision for Small Farms. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. X-G-2025: 143-149. <https://doi.org/10.5194/isprs-annals-X-G-2025-143-2025>
- Bahrami, H., McNairn, H., Mahdianpari, M. and Homayouni, S. 2022. A Meta-Analysis of Remote Sensing Technologies and Methodologies for Crop Characterization. *Remote Sensing*. 14(22): 5633. <https://www.mdpi.com/2072-4292/14/22/5633>
- Bourbour, H., Abdollahipour, M., Abdollahi, H., Abiar, E. and Mashal, M. 2025. Pre-harvest forecasting of rainfed wheat yield in Iran using multi-source remote sensing and machine learning. *Agricultural Water Management*. 322: 110005. <https://doi.org/10.1016/j.agwat.2025.110005>
- Breiman, L. 2002. Manual on setting up, using, and understanding random forests v3. 1. Statistics Department University of California Berkeley, CA, USA. 1(58): 3-42.
- Brown, C. F., Brumby, S. P., Guzder-Williams, B., Birch, T., Hyde, S. B., Mazzariello, J., . . . Tait, A. M. 2022. Dynamic World, Near real-time global 10 m land use land cover mapping. *Scientific Data*. 9(1): 251. <https://doi.org/10.1038/s41597-022-01307-4>
- Chen, X., Zhan, Y., Liu, Y., Gu, X., Yu, T., Wang, D., . . . Zhang, Y. 2020. Improving the Classification Accuracy of Annual Crops Using Time Series of Temperature and Vegetation Indices. *Remote Sensing*. 12(19): 3202. <https://www.mdpi.com/2072-4292/12/19/3202>
- Cutler, D. R., Edwards Jr, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J. and Lawler, J. J. 2007. Random forests for classification in ecology. *Ecology*. 88(11): 2783-2792.
- Dash, J., Curran, P., Tallis, M., Llewellyn, G., Taylor, G. and Snoeij, P. 2010. Validating the MERIS Terrestrial Chlorophyll Index (MTCI) with ground chlorophyll content data at MERIS spatial resolution. *International Journal of Remote Sensing*. 31(20): 5513-5532.
- Dobrinčić, D., Gašparović, M. and Medak, D. 2021. Sentinel-1 and 2 Time-Series for Vegetation Mapping Using Random Forest Classification: A Case Study of Northern Croatia. *Remote Sensing*. 13(12): 2321. <https://doi.org/10.3390/rs13122321>
- Eisfelder, C., Boemke, B., Gessner, U., Sogno, P., Alemu, G., Hailu, R., . . . Huth, J. 2024. Cropland and Crop Type Classification with Sentinel-1 and Sentinel-2 Time Series Using Google Earth Engine for Agricultural Monitoring in Ethiopia. *Remote Sensing*. 16(5): 866. <https://doi.org/10.3390/rs16050866>
- Evans, F. H. and Shen, J. 2021. Long-term hindcasts of

- <https://doi.org/10.1080/07038992.1993.10874543>
 Mohamadi-Monavar, H. and Zibazadeh, S. 2022. Distinguishing Rain-fed and Irrigated Crops in Hamadan Province Using Spectral Indices of Satellite Images. *Journal of Agricultural Machinery*. 12(4): 529-542. <https://doi.org/10.22067/jam.2021.69074.1022> [In Persian]
- Mousavi, S. A., Abasyadeh Tehrani, N. and Janalipour, M. 2020. Estimation of the train wheat under cultivation area using Sentinel-2 satellite images (Case study: Sojasroud region, Khodabandeh city, Zanjan province). *Journal of Environment and Research Technology*. 5(7): 77-99. <https://sid.ir/paper/373696/en> [In Persian]
- Nicolau, A. P., Dyson, K., Saah, D. and Clinton, N. 2024. Accuracy Assessment: Quantifying Classification Quality. In J. A. Cardille, M. A. Crowley, D. Saah, and N. E. Clinton (Eds.), *Cloud-Based Remote Sensing with Google Earth Engine: Fundamentals and Applications*. Springer International Publishing: 135-145. https://doi.org/10.1007/978-3-031-26588-4_7
- Obsie, E. Y., Qu, H. and Drummond, F. 2020. Wild blueberry yield prediction using a combination of computer simulation and machine learning algorithms. *Computers and Electronics in Agriculture*. 178: 105778. <https://doi.org/10.1016/j.compag.2020.105778>
- Ouaadi, N., Jarlan, L., Khabba, S., Le Page, M., Chakir, A., Er-Raki, S. and Frison, P.-L. 2023. Are the C-band backscattering coefficient and interferometric coherence suitable substitutes of NDVI for the monitoring of the FAO-56 crop coefficient? *Agricultural Water Management*. 282: 108276. <https://doi.org/10.1016/j.agwat.2023.108276>
- Pan, Y., Li, L., Zhang, J., Liang, S., Zhu, X. and Sulla-Menashe, D. 2012. Winter wheat area estimation from MODIS-EVI time series data using the Crop Proportion Phenology Index. *Remote Sensing of Environment*. 119: 232-242. <https://doi.org/10.1016/j.rse.2011.10.011>
- Pande, C. B., Egbueri, J. C., Costache, R., Sidek, L. M., Wang, Q., Alshehri, F., . . . Chandra Pal, S. 2024. Predictive modeling of land surface temperature (LST) based on Landsat-8 satellite data and machine learning models for sustainable development. *Journal of Cleaner Production*. 444: 141035. <https://doi.org/10.1016/j.jclepro.2024.141035>
- Pooya, M. R., Hasankhani, A., Fathololomi, S. and Karimi Firozjaei, M. 2025. A Spatial Multi-Criteria Decision-Making Approach to Evaluating Homogeneous Areas for Rainfed Wheat Yield Assessment. *Water*. 17(7): 1045. <https://www.mdpi.com/2073-4441/17/7/1045>
- Rahnama, S., Maharlooei, M., Rostami, M. and Maghsoudi, H. 2019. Determining the best classification algorithm in order to estimate the area under date palm cultivation using Landsat 8 satellite records, satellite image time series and machine learning methods. *ISPRS Journal of Photogrammetry and Remote Sensing*. 160: 124-135. <https://doi.org/10.1016/j.isprsjprs.2019.11.008>
- Kayad, A., Sozzi, M., Gatto, S., Marinello, F. and Pirotti, F. 2019. Monitoring Within-Field Variability of Corn Yield using Sentinel-2 and Machine Learning Techniques. *Remote Sensing*. 11(23): 2873. <https://doi.org/10.3390/rs11232873>
- Kordi, F. and Yousefi, H. 2022. Crop classification based on phenology information by using time series of optical and synthetic-aperture radar images. *Remote Sensing Applications: Society and Environment*. 27: 100812. <https://doi.org/10.1016/j.rsase.2022.100812>
- Li, H. Y., Lawrence, J. A., Mason, P. J. and Ghail, R. C. 2023. UNSUPERVISED WINTER WHEAT MAPPING BASED ON MULTI-SPECTRAL AND SYNTHETIC APERTURE RADAR OBSERVATIONS. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XLVIII-1/W2-2023: 1411-1416. <https://doi.org/10.5194/isprs-archives-XLVIII-1-W2-2023-1411-2023>
- Li, Z. L., Wu, H., Duan, S. B., Zhao, W., Ren, H., Liu, X., . . . Zhu, J. 2023. Satellite remote sensing of global land surface temperature: Definition, methods, products, and applications. *Reviews of Geophysics*. 61(1). <https://doi.org/10.1029/2022RG000777>
- Liaw, A. and Wiener, M. 2002. Classification and regression by randomForest. *R news*. 2(3): 18-22.
- Liu, J., Feng, Q., Gong, J., Zhou, J., Liang, J. and Li, Y. 2018. Winter wheat mapping using a random forest classifier combined with multi-temporal and multi-sensor data. *International Journal of Digital Earth*. 11(8): 783-802. <https://doi.org/10.1080/17538947.2017.1356388>
- Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, Z., Yang, L. and Merchant, J. W. 2000. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *International Journal of Remote Sensing*. 21(6-7): 1303-1330. <https://doi.org/10.1080/014311600210191>
- Luo, H. 2025. Rapid and Accurate Mapping of Winter Wheat Distribution in Jiaozuo City Using Sentinel-2 and Google Earth Engine. *Academic Journal of Science and Technology*. 14(1): 327-333. <https://doi.org/10.54097/ttcn1622>
- Mandal, D., Kumar, V., Ratha, D., Dey, S., Bhattacharya, A., Lopez-Sanchez, J. M., . . . Rao, Y. S. 2020. Dual polarimetric radar vegetation index for crop growth monitoring using sentinel-1 SAR data. *Remote Sensing of Environment*. 247: 111954. <https://doi.org/10.1016/j.rse.2020.111954>
- McNairn, H. and Protz, R. 1993. Mapping Corn Residue Cover on Agricultural Fields in Oxford County, Ontario, Using Thematic Mapper. *Canadian Journal of Remote Sensing*. 19(2): 152-159.

- Vegetation Dynamics: An Austrian Case Study. *Remote Sensing*. 10(9): 1396.
- Weng, Q., Hu, X. and Liu, H. 2009. Estimating impervious surfaces using linear spectral mixture analysis with multitemporal ASTER images. *International Journal of Remote Sensing*. 30(18): 4807-4830.
<https://doi.org/10.1080/01431160802665926>
- Xiao, G., Zhang, X., Niu, Q., Li, X., Li, X., Zhong, L. and Huang, J. 2024. Winter wheat yield estimation at the field scale using sentinel-2 data and deep learning. *Computers and Electronics in Agriculture*. 216: 108555.
<https://doi.org/10.1016/j.compag.2023.108555>
- Yang, M., Guo, B., Wang, J. and Tian, C. 2024. Phenology Index-Based Method for Mapping Winter Wheat and Summer Maize Rotation Cropping Pattern With Sentinel-2 Imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 17: 13795-13808.
<https://doi.org/10.1109/JSTARS.2024.3434438>
- Yao, J., Wu, J., Xiao, C., Zhang, Z. and Li, J. 2022. The Classification Method Study of Crops Remote Sensing with Deep Learning, Machine Learning, and Google Earth Engine. *Remote Sensing*. 14(12): 2758. <https://www.mdpi.com/2072-4292/14/12/2758>
- Zaka, M. M. and Samat, A. 2024. Advances in Remote Sensing and Machine Learning Methods for Invasive Plants Study: A Comprehensive Review. *Remote Sensing*. 16(20): 3781.
<https://www.mdpi.com/2072-4292/16/20/3781>
- Zhang, H., Zhang, Y., Tong, F. and Li, M. 2025. A novel winter wheat yield prediction framework using fused spatial-temporal-spectral (STS) information from Sentinel-2 and Landsat 8 via improved Pix2Pix network. *Computers and Electronics in Agriculture*. 231: 109982.
<https://doi.org/10.1016/j.compag.2025.109982>
- Zhao, Y., Potgieter, A. B., Zhang, M., Wu, B. and Hammer, G. L. 2020. Predicting Wheat Yield at the Field Scale by Combining High-Resolution Sentinel-2 Satellite Imagery and Crop Modelling. *Remote Sensing*. 12(6): 1024.
<https://doi.org/10.1080/01431160802665926>
- Zheng, Y., Dong, W., ZhipingYang, Lu, Y., Zhang, X., Dong, Y. and Sun, F. 2024. A new attention-based deep metric model for crop type mapping in complex agricultural landscapes using multisource remote sensing data. *International Journal of Applied Earth Observation and Geoinformation*. 134: 104204.
<https://doi.org/10.1016/j.jag.2024.104204>
- imagery.
- Rokni, K. and Musa, T. A. 2019. Normalized difference vegetation change index: A technique for detecting vegetation changes using Landsat imagery. *CATENA*. 178: 59-63.
<https://doi.org/10.1016/j.catena.2019.03.007>
- Sharifi Paichoon, M., Omidvar, K. and Motazaker, K. 2019. Assessment of flooding using cluster analysis and multivariable regression methods with emphasis on hydro geomorphological parameters (Case study: Maroon catchment). *Journal of Natural Environmental Hazards*. 8(21): 75-92.
<https://doi.org/10.22111/jneh.2018.22519.1336> [In Persian]
- Small, C. and Sousa, D. 2023. Spectral Characteristics of the Dynamic World Land Cover Classification. *Remote Sensing*. 15(3): 575.
<https://www.mdpi.com/2072-4292/15/3/575>
- Soltani, M., Bahmanabadi, B. and Mokhtaran, A. 2026. Crop Classification and Acreage Estimation at Basin Scale Using Different Machine Learning Approaches. *Iranian Journal of Soil and Water Research*. 56(11): 3129-3156.
<https://doi.org/10.22059/ijswr.2025.402830.670012> [In Persian]
- Teluguntla, P., Thenkabail, P., Oliphant, A., Gumma, M. K., Anece, I. and Foley, D. 2024. Landsat-derived global rainfed and irrigated area product @ 30m (LGRIP30): A study of irrigated and rainfed areas at the country, state, and county level for conterminous United States (CONUS). AGU Fall Meeting Abstracts.
- Thenkabail, P. S., Biradar, C. M., Noojipady, P., Dheeravath, V., Li, Y., Velpuri, M., . . . Dutta, R. 2009. Global irrigated area map (GIAM), derived from remote sensing, for the end of the last millennium. *International Journal of Remote Sensing*. 30(14): 3679-3733.
<https://doi.org/10.1080/01431160802698919>
- Tsuchiya, Y. and Sonobe, R. 2025. Crop Classification Using Time-Series Sentinel-1 SAR Data: A Comparison of LSTM, GRU, and TCN with Attention. *Remote Sensing*. 17(12): 2095.
<https://www.mdpi.com/2072-4292/17/12/2095>
- Van De Kerchove, R., Zanaga, D., Xu, P., Tsendbazar, N. and Lesiv, M. 2022. Product User Manual V 2.0.
- Venter, Z. S., Barton, D. N., Chakraborty, T., Simensen, T. and Singh, G. 2022. Global 10 m land use land cover datasets: A comparison of dynamic world, world cover and esri land cover. *Remote Sensing*. 14(16): 4101. <https://doi.org/10.3390/rs14164101>
- Vreugdenhil, M., Wagner, W., Bauer-Marschallinger, B., Pfeil, I., Teubner, I., Rüdiger, C. and Strauss, P. 2018. Sensitivity of Sentinel-1 Backscatter to

Intelligent Monitoring of Rainfed and Irrigated Wheat Cultivation in the Marun–Jarahi Basin Using Sentinel Data and the Random Forest Model

M. Soltani¹, B. Bahmanabadi^{2*}, A. Mokhtaran³

Recived: Feb.22, 2025

Accepted: May.18, 2026

Abstract

Wheat is one of the most strategic crops for food security and occupies a large proportion of agricultural lands in Iran. However, climate change, increasing water scarcity, and the limitations of traditional field-based statistics have made accurate monitoring of wheat cultivated areas a challenging task. In this study, an integrated framework based on multi-source remote sensing data and a machine learning approach was developed to identify rainfed and irrigated wheat fields and to estimate their cultivated area in the Maroon watershed in southwest Iran. A set of spectral, radar, and thermal features representing the biophysical, structural, and water-related characteristics of wheat was extracted and used as input variables for the model. The Random Forest algorithm was selected due to its robustness against overfitting and its ability to handle heterogeneous and non-linear data. Model performance was evaluated using a confusion matrix, overall accuracy, and the kappa coefficient. The results demonstrated a high classification performance, with an overall accuracy of approximately 97% and a kappa coefficient of about 0.96 for the test dataset. In addition, the estimated wheat cultivated area showed a low bias of around 3%, indicating a high level of reliability in distinguishing between rainfed and irrigated wheat fields.

Keywords: ESA WorldCover, LGRIP30, Dynamic World, wheat

1- Assistant Professor, Department of Water Science and Engineering, Faculty of Agriculture and Natural Resources, Imam Khomeini International University, Qazvin, Iran

2- Corresponding Author, Ph.D of water engineering, Department of Water Science and Engineering, Faculty of Agriculture and Natural Resources, Imam Khomeini International University, Qazvin, Iran

3- Assistant Professor, Agricultural Engineering Research Institue (AERI) Agricultural Research, Education and Extension Organization (AREEO); Karaj; Iran

(* - Corresponding Author Email: b.bahmanabadi@gmail.com)