



رابطه بین حجم نمونه آموزشی و دقت طبقه‌بندی برنج با استفاده از سنتینل ۱ و ۲  
مطالعه موردی: مازندران، ایران

تاریخ پذیرش: ۱۴۰۲/۰۳/۲۹

تاریخ دریافت: ۱۴۰۲/۰۲/۲۵

رحمان باریده<sup>\*</sup>، فرشته نسیمی<sup>۲</sup>

۱- دکتری تخصصی آبیاری و زهکشی دانشگاه ارومیه، ارومیه، ایران

۲- دانشجوی دکتری آبیاری و زهکشی دانشگاه ارومیه، ارومیه، ایران

#### چکیده

در چند دهه اخیر با پیشرفت علم و تنوع ماهواره‌های نوری و راداری امکان پایش و تفکیک محصولات کشاورزی در سطح وسیع فراهم شده است. اما یکی از چالش‌های اصلی در طبقه‌بندی تصاویر ماهواره‌ای، تعداد نمونه‌های آموزشی مورد نیاز است. بنابراین هدف از این تحقیق بررسی ارتباط بین تعداد نمونه‌های آموزشی و دقت طبقه‌بندی اراضی برنج و غیربرنج بود. برای این منظور از تصاویر ماهواره‌های سنتینل ۱ و ۲ و طبقه‌بندی کننده جنگل تصادفی در پلتفرم گوگل ارث انجین استفاده گردید. در مجموع ۲۵۰۰ نمونه برنج و ۹۵۰۰ نمونه غیربرنج در محدوده مطالعاتی تهیه گردید و تعداد ۱۰۰ اجرای متفاوت با تعداد نمونه‌های آموزشی مختلف انجام شد. نتایج نشان داد که سری زمانی تصاویر سنتینل ۱ امکان متمایز کردن مزارع برنج از غیر برنج را با دقت بالایی امکان‌پذیر می‌کند. بنابراین ورودی‌های طبقه‌بندی کننده جنگل تصادفی شامل شیب ضریب باز پراکنش، اختلاف حداکثر و حداقل ضریب باز پراکنش و شاخص NDVI بودند. نتایج نشان داد که افزایش تعداد نمونه‌های آموزشی با افزایش دقت طبقه‌بندی رابطه غیرخطی دارد و با افزایش تعداد نمونه‌ها دقت کاهش پیدا می‌کند. بیشترین دقت کلی (۸۹ درصد) و ضریب کاپا (۰/۸۶) زمانی به دست آمد که از یک نمونه آموزشی در هر ۱۸۱ هکتار از مزارع برنج و یک نمونه آموزشی در هر ۶۶۹ هکتار از کاربری‌های غیر برنج استفاده گردید. با افزایش دو برابری تعداد نمونه‌ها دقت کلی و ضریب کاپا به ترتیب برابر ۰/۸۷٪ و ۰/۸۴ به دست آمد. بنابراین نتایج نشان می‌دهد که صرفاً افزایش تعداد نمونه‌های آموزشی باعث افزایش دقت طبقه‌بندی نمی‌شود.

**کلمات کلیدی:** سری زمانی، جنگل تصادفی، ماهواره راداری، گوگل ارث انجین

## مقدمه

برنج یکی از مهم‌ترین محصولات غذایی جمعیت جهان به شمار می‌آید. این محصول ۲۰ درصد از کالری موجود در جهان را تشکیل می‌دهد (۱، ۲). کشت برنج در ۱۱۴ کشور صورت می‌گیرد و ۱۲ درصد از اراضی کشاورزی جهان به کشت برنج اختصاص داده شده است (۳). تقریباً ۹۰ درصد برنج جهان در منطقه آسیا تولید و مصرف می‌شود. بنابراین مؤسسات دولتی زمان و هزینه زیادی را برای پایش و تعیین سطح زیر کشت این محصول صرف می‌کنند (۱، ۴). برآورد سطح زیر کشت محصولات کشاورزی معمولاً از طریق تخمین کارشناسی و برآورد از طریق فهرست برداری انجام می‌گیرد. بررسی‌ها نشان می‌دهد که این روش‌ها دارای خطاهای زیادی هستند بنابراین به‌کارگیری روش‌های دقیق‌تر و سریع‌تر، مانند سنجش از دور، برای کمک به برنامه‌ریزی‌های کلان در حوزه کشاورزی ضروری است (۵، ۶).

در چند دهه اخیر، پیشرفت علوم کامپیوتری و افزایش تنوع ماهواره‌های نوری و راداری باعث افزایش استفاده از روش سنجش از دور در همه علوم به ویژه کشاورزی شده است. پردازش تصاویر ماهواره‌ای طیف وسیعی از اطلاعات لازم برای تصمیم‌گیری در حوزه کشاورزی و محیط زیست ارائه می‌دهد. بنابراین استفاده از تصاویر ماهواره‌ای برای تشخیص مزارع برنج علاوه بر حذف خطاهای انسانی، باعث کاهش هزینه و زمان می‌شود (۱، ۵).

در گذشته، پژوهش‌های زیادی برای تعیین و شناسایی سطح زیر کشت برنج با استفاده از تصاویر ماهواره‌های نوری مانند سنتینل ۲، لندست و مودیس صورت گرفته است. جین و همکاران (۷) برای تعیین سطح زیر کشت برنج در شمال شرقی چین از تصاویر لندست و شاخص پوشش گیاهی نرمال شده (NDVI)، شاخص ارتقا یافته پوشش گیاهی (EVI) و شاخص آب سطح زمین (LSWI) استفاده کردند. ساکاموتو و همکاران (۸) برای تشخیص مزارع برنج در ژاپن از تصاویر لندست و شاخص پوشش آب نرمال شده (NDWI) استفاده کردند. همچنین ژانگ و همکاران (۹) داده‌های دمای سطح زمین (LST) ماهواره مودیس را برای شناسایی مزارع برنج بکار بردند. با توجه به اینکه تولید برنج به آب و دمای زیادی نیاز دارد، اکثر شالیزارهای برنج در مناطق گرم و مرطوب قرار دارند. اما این‌گونه مناطق در اکثر اوقات پوشیده از ابر هستند و این موضوع امکان استفاده از تصاویر ماهواره‌های نوری را با مشکل مواجه می‌کند. در مقابل داده‌های راداری به دلیل عدم نیاز به نور خورشید قابلیت تصویر برداری در هر نوع آب و هوایی از جمله شرایط ابری را دارند. بنابراین استفاده از تصاویر ماهواره‌های راداری در تشخیص مزارع برنج به طور فزاینده‌ای مورد توجه قرار گرفته است (۱، ۱۰).

چانگ و همکاران (۱) برای تشخیص مزارع برنج از تصاویر ماهواره سنتینل ۱ استفاده کردند. در این پژوهش از تفاوت‌های ضریب باز پراکنش در مزارع برنج و سایر کشت‌ها استفاده گردید و نتایج این روش را با روش‌های طبقه‌بندی درخت تصمیم (DT)، ماشین بردار پشتیبان (SVM)، K-نزدیک‌ترین همسایه (KNN) و تجزیه و

تحلیل متمایز درجه دوم (QDA) مقایسه کردند. نتایج آن‌ها نشان داد که استفاده از تصاویر راداری دقت بالایی در تشخیص مزارع برنج دارد. سادات و همکاران (۱۰) از تصاویر ماهواره سنتینل ۱ و الگوریتم جنگل تصادفی (RF) برای شناسایی مزارع برنج استفاده کردند. در این پژوهش با استفاده از پولاریزاسیون VH مزارع برنج با دقت کلی ۸۰ درصد شناسایی شد. با این حال، در تحقیقات اخیر بیان شده است، روش‌هایی که فقط از تصاویر ماهواره‌های راداری استفاده می‌کنند دارای دقت کمتری نسبت به روش‌های مبتنی بر استفاده همزمان از تصاویر ماهواره‌های نوری و راداری هستند، زیرا سنسورهای راداری حساسیت کمتری به پوشش گیاهی دارند (۱۱)، (۱۲).

از محدودیت‌های استفاده از تصاویر ماهواره‌ای می‌توان به حجم بالای تصاویر و محاسبات نرم‌افزاری به خصوص در سطح وسیع اشاره کرد. این محدودیت باعث محدود شدن مطالعات به وسعت کم و عدم استفاده از سری زمانی شده است. بنابراین اخیراً در بسیاری از تحلیل‌های سنجش از دور از پلتفرم‌های مبتنی بر ابر استفاده می‌کنند (۱۱). Google Earth Engine (GEE) یک پلت فرم مبتنی برابر برای تجزیه و تحلیل تصاویر ماهواره‌ای در مقیاس وسیع است. این پلتفرم با پردازش سریع حجم زیادی از تصاویر ماهواره‌های مختلف و ترکیب آن‌ها امکان انجام مطالعات در سطح وسیع را مهیا ساخته است.

علاوه بر این، یکی از مسائل مهم در طبقه‌بندی و تفکیک محصولات کشاورزی در مقیاس وسیع با استفاده از تصاویر ماهواره‌ای، تعداد نمونه‌های مورد نیاز است. امینی و همکاران (۱۳) برای طبقه‌بندی محصولات کشاورزی در کانادا با استفاده از تصاویر ماهواره‌های سنتینل ۱ و ۲ از ۱۹۷۰۰۰ نمونه آموزشی استفاده کردند و ۱۷ محصول مختلف را در محدوده مطالعاتی با دقت کلی ۷۷٪ طبقه‌بندی کردند. همچنین تیواری و همکاران (۱۴) برای تعیین سطح زیر کشت گندم در افغانستان از ۱۶۰۰۰ نمونه آموزشی استفاده کردند. اینو و همکاران (۱۱) برای پهنه‌بندی مزارع برنج در ژاپن با استفاده از تصاویر سنتینل ۱ و ۲ از ۱۴۰۰ نمونه برنج و ۳۶۰۰۰ نمونه غیر برنج استفاده کردند. همچنین تالما و هایلو (۱۵) برای تشخیص مزارع برنج در اتیوپی از ۱۶۹ نمونه برنج و ۲۲۶ نمونه غیر برنج استفاده کردند.

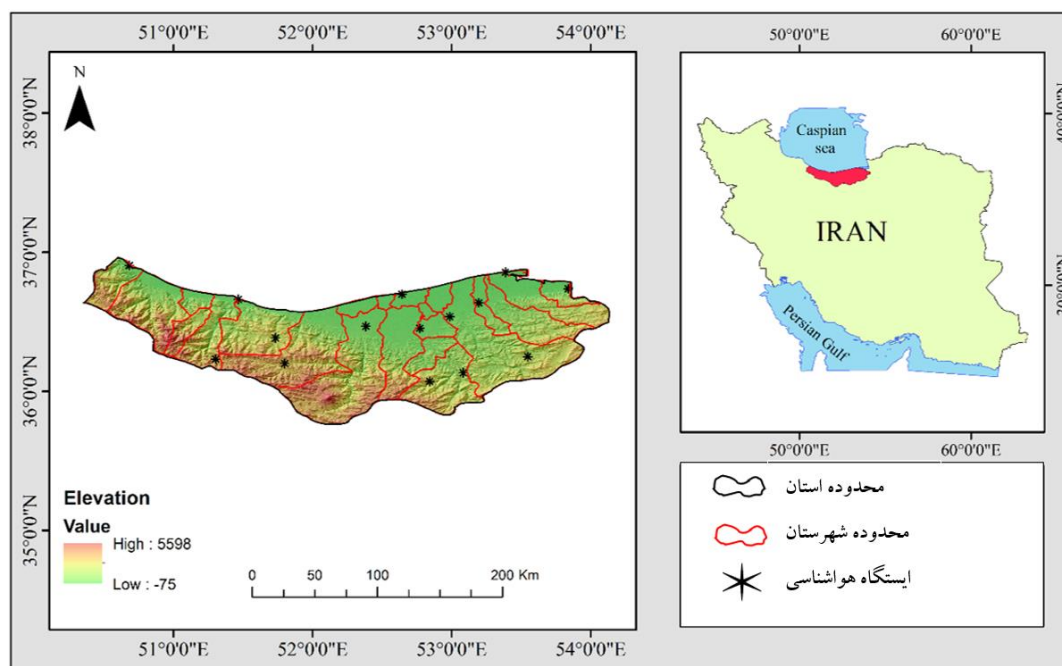
بررسی منابع نشان می‌دهد که در مطالعات مختلف، اعم از طبقه‌بندی مزارع برنج یا سایر محصولات، از تعداد نمونه‌های آموزشی متفاوتی استفاده شده است و ارتباط بین دقت نقشه‌های تولید شده و تعداد نمونه‌های آموزشی در واحد سطح به طور واضح مشخص نشده است. با توجه به اینکه جمع‌آوری نمونه‌های آموزشی هزینه‌بر و زمان‌بر است و همچنین تشخیص تعداد نمونه‌های مورد نیاز از دغدغه‌های اصلی هر محقق به شمار می‌رود، بنابراین در این پژوهش به بررسی دقت طبقه‌بندی محصول برنج و ارتباط آن با تعداد نمونه‌های آموزشی

مختلف پرداخته شده است. هدف از این بررسی‌ها تعیین حداقل نمونه آموزشی مورد نیاز برای تشخیص محصول برنج از غیر برنج با دقت قابل قبول بود.

## مواد و روش‌ها

### محدوده مطالعاتی

این پژوهش در استان مازندران واقع در شمال ایران انجام شد. این استان با مساحت ۲۳۸۴۲ کیلومتر مربع در محدوده ۳۵ درجه و ۴۷ دقیقه تا ۳۶ درجه و ۵۸ دقیقه عرض شمالی و ۵۰ درجه و ۳۴ دقیقه تا ۵۴ درجه و ۱۴ دقیقه طول شرقی قرار دارد. ارتفاع از سطح آب‌های آزاد استان مازندران از ۷۵- تا ۵۵۹۸ متر متغیر است. میانگین بارندگی سالانه استان مازندران از ۳۰۰ تا ۱۰۰۰ میلی‌متر متغیر است. بر اساس طبقه‌بندی دومارتن، نواحی غربی مازندران بسیار مرطوب، نواحی مرکزی مرطوب، نواحی شرقی مدیترانه‌ای و نواحی کوهستانی نیمه مرطوب است. استان مازندران به عنوان یکی از اصلی‌ترین مناطق تولید برنج در ایران به شمار می‌آید و در اکثر مزارع کشاورزی آن برنج کشت می‌شود (۱۰، ۱۶، ۱۷، ۱۸). موقعیت مکانی محدوده مطالعاتی در شکل (۱) نشان داده شده است.

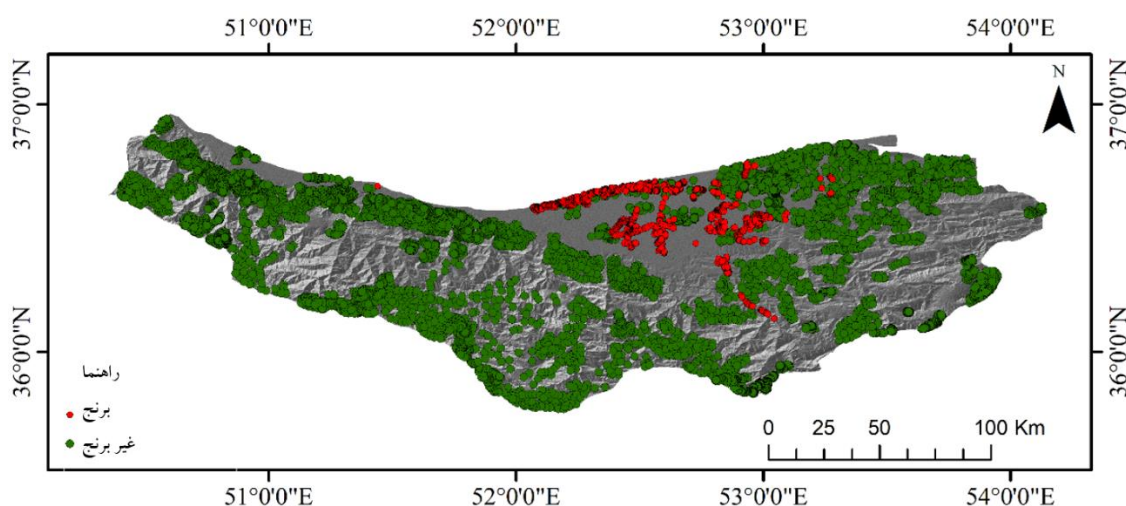


شکل (۱): محدوده مطالعاتی و محل قرارگیری آن در شمال ایران

Fig. (1): Location of the study area in north Iran.

### برداشت نمونه

برای انجام این پژوهش، تعداد ۱۲۰۰۰ نمونه از کاربری‌های مختلف جمع‌آوری گردید. با توجه به شکل ۱ مساحت کل محدوده مطالعاتی برابر ۲۳۸۴۲۰۰ هکتار است. بنابراین در هر ۲۰۰ هکتار یک نمونه با استفاده از پایش میدانی و تصاویر گوگل ارث تهیه شد. ۲۵۰۰ نمونه از مزارع برنج با استفاده از GPS برداشت گردید. همچنین ۹۵۰۰ نمونه دیگر که شامل کاربری‌های شهری، جنگل، پهنه‌های آبی، مراتع و کشت‌های غیربرنج بود با استفاده از GPS و تصاویر گوگل ارث استخراج گردید. برای ایجاد پراکندگی نمونه‌ها، سعی گردید که از همه کاربری‌های محدوده مطالعاتی نمونه جمع‌آوری شود. پراکندگی نمونه‌ها در شکل ۲ نشان داده شده است.



شکل (۲): پراکندگی نمونه‌های برنج و غیربرنج در محدوده مطالعاتی

Fig. (2): Geographical distribution of samples of rice and other covers in the study area

### تصاویر ماهواره‌ای

با توجه به اینکه منطقه مورد مطالعه در شمال ایران و منتهی به دریای خزر است، در اکثر روزهای سال آسمانی ابری و بارانی دارد. این موضوع باعث می‌شود که تصاویر ماهواره‌های نوری دارای ابر زیادی باشند که استفاده از آن‌ها را با مشکل مواجه می‌کند. لذا ترکیب ماهواره‌های سنتینل ۱ و ۲ می‌تواند این مشکل را برطرف کند. بنابراین برای شناسایی مزارع برنج از ترکیب تصاویر این دو ماهواره استفاده گردید. برای این منظور از پلتفرم گوگل ارث انجین استفاده شد.

### نقشه کشت برنج

برای تعیین سطح زیر کشت برنج از طبقه‌بندی کننده جنگل تصادفی (RF) استفاده گردید. برای آموزش RF از ۷۰ درصد نمونه‌ها استفاده شد. ارزیابی طبقه‌بندی‌های انجام شده نیز با نمونه‌های باقیمانده صورت گرفت. به منظور بررسی ارتباط بین دقت تصاویر طبقه‌بندی شده با استفاده از تعداد نمونه‌های متفاوت ۱۰۰ اجرای مختلف انجام شد. در اجرای اول از ۱۰ درصد نمونه‌های برنج و غیربرنج استفاده گردید و در اجراهای دوم تا دهم تعداد نمونه‌های برنج ثابت و در هر مرحله ۱۰ درصد به نمونه‌های غیربرنج اضافه گردید. این اجراها برای ۲۰، ۳۰، ۴۰، ۵۰، ۶۰، ۷۰، ۸۰، ۹۰ و ۱۰۰ درصد نمونه‌های آموزشی برنج تکرار شدند. در هر مرحله ماتریس خطا تشکیل گردید و مقدار دقت کلی و ضریب کاپا محاسبه شدند.

### متغیرهای ورودی به RF

برای تعیین نقشه کشت برنج ابتدا سری زمانی باز پراکنش سنتینل ۱ برای کشت‌ها و کاربری‌های مختلف ترسیم گردید. با توجه به اینکه مزارع برنج به صورت غرق آب هستند بنابراین ضریب باز پراکنش در ابتدای رشد برای این مزارع کم است و با رشد گیاه این ضریب افزایش پیدا می‌کند. اما این موضوع برای سایر محصولات صادق نیست. پس می‌توان از آن به عنوان یک ویژگی متمایزکننده کشت برنج از سایر محصولات استفاده نمود (۱).

در محدوده مطالعاتی تالاب‌ها، آبگیرها و استخرهای زیادی وجود دارد که در ابتدای دوره رشد محصولات زراعی دارای عمق زیادی هستند ولی با گذشت زمان و رشد محصولات زراعی آب آن‌ها نیز کم می‌شود و در نتیجه پوشش گیاهی قابل توجهی نمایان می‌شود که شرایط مزارع برنج را تداعی می‌کند. برای جلوگیری از این خطا، تصاویر سنتینل ۱ در دوره غیرزراعی مورد استفاده قرار گرفتند. در این دوره تالاب‌ها و استخرها دارای آب قابل توجهی هستند و ضریب باز پراکنش آن‌ها کم است در حالی که در سایر کاربری‌ها تجمع آب صورت نمی‌گیرد. بنابراین این موضوع به عنوان یک متغیر متمایز کننده کشت برنج از پهنه‌های آبی در نظر گرفته شد. همچنین با توجه به اینکه در مطالعات اخیر به این نتیجه رسیدند که پولاریزاسیون VH تفاوت بین مزارع برنج و غیر برنج را بهتر نشان می‌دهد لذا در این پژوهش از پولاریزاسیون VH سنتینل ۱ استفاده گردید (۱، ۱۰).

برای ایجاد تفاوت بیشتر بین مزارع برنج و سایر کشت‌ها از شاخص پوشش گیاهی نرمال شده NDVI (معادله ۱) استفاده گردید. سری زمانی این شاخص در دوره رشد برنج با استفاده از تصاویر سنتینل ۲ برای کشت‌های مختلف مورد بررسی قرار گرفت. با توجه به اینکه دوره رشد برنج و سایر محصولات در منطقه مورد مطالعه متفاوت است لذا استفاده از شاخص NDVI در دوره حداکثر پوشش گیاهی (میانگین ماه حداکثر رشد) می‌تواند باعث افزایش دقت طبقه‌بندی کننده RF شود.

$$NDVI = \frac{NIR-RED}{NIR+RED} \quad (1)$$

که در آن، NIR بازتاب در باند فرسرخ نزدیک و RED بازتاب در باند سرخ است. شاخص NDVI به شدت وابسته به پوشش گیاهی است. مقادیر بالای NDVI مربوط به مناطقی است که انعکاس بیشتری در طیف مادون قرمز نزدیک دارند. انعکاس بالاتر در مادون قرمز نزدیک مربوط به پوشش گیاهی سالم‌تر و متراکم‌تر است (۲).

#### اعتبارسنجی نقشه کشت برنج

در هر مرحله برای ارزیابی دقت طبقه‌بندی ماتریس خطا ایجاد شد و ضریب کاپا (معادله ۲) و دقت کلی (معادله ۳) محاسبه شدند.

$$Kappa = \frac{N \sum_{i=1}^r X_{ii} - \sum_{i=1}^r (X_{i+} \cdot X_{+i})}{N^2 - \sum_{i=1}^r (X_{i+} \cdot X_{+i})} \quad (2)$$

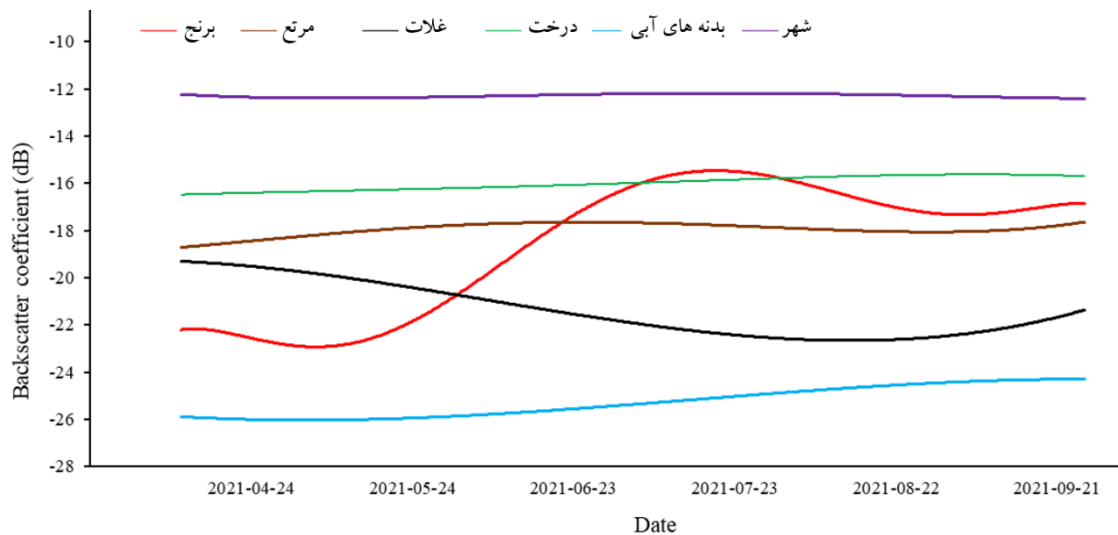
$$OA = \frac{\text{تعداد کل پیکسل های طبقه بندی شده صحیح}}{\text{تعداد کل پیکسل های مرجع}} \times 100 \quad (3)$$

که در آن،  $r$  تعداد سطرها و ستون‌ها در ماتریس خطا،  $N$  تعداد کل مشاهدات،  $X_{ii}$  مشاهده در ردیف  $i$  و ستون  $i$ ،  $X_{i+}$  مجموع ردیف  $i$ ، و  $X_{+i}$  مجموع ستون  $i$  است. ضریب کاپا دقت بین داده‌های مرجع و نقشه طبقه‌بندی شده نشان می‌دهد. اگر مقدار کاپا  $0/4$  یا کمتر باشد، دقت پایین نقشه طبقه‌بندی شده را نشان می‌دهد. مقدار بین  $0/4$  و  $0/8$  نشان‌دهنده دقت متوسط و مقادیر بیشتر از  $0/8$  نشان‌دهنده دقت بالا است. OA میزان موفقیت طبقه‌بندی را در همه کلاس‌ها نشان می‌دهد (۱، ۱۰، ۱۹، ۲۰).

#### نتایج و بحث

##### سری زمانی ضریب باز پراکنش سنتینل ۱

در شکل ۳ سری زمانی ضریب باز پراکنش تصاویر ماهواره سنتینل ۱ برای کاربری‌های مختلف منطقه مورد مطالعه نشان داده شده است. با توجه به اینکه مزارع برنج در ابتدای دوره رشد کاملاً پوشیده از آب هستند و با رشد محصول سطح مزرعه از گیاه پوشیده می‌شود لذا مشاهده می‌شود که تغییرات ضریب باز پراکنش برای محصول برنج بیشتر از سایر کاربری‌ها است. کمترین مقدار ضریب باز پراکنش در پهنه‌های آبی و بیشترین مقدار آن در کاربری شهری مشاهده می‌شود. در پوشش‌های مرتعی، جنگلی و پهنه‌های آبی به دلیل ثابت بودن پوشش آن‌ها، روند تغییرات ناچیز است. بنابراین تفاوت تغییرات ضریب باز پراکنش در مزارع برنج نسبت به سایر پوشش‌ها به عنوان وردی به طبقه‌بندی‌کننده RF مورد استفاده قرار گرفت.



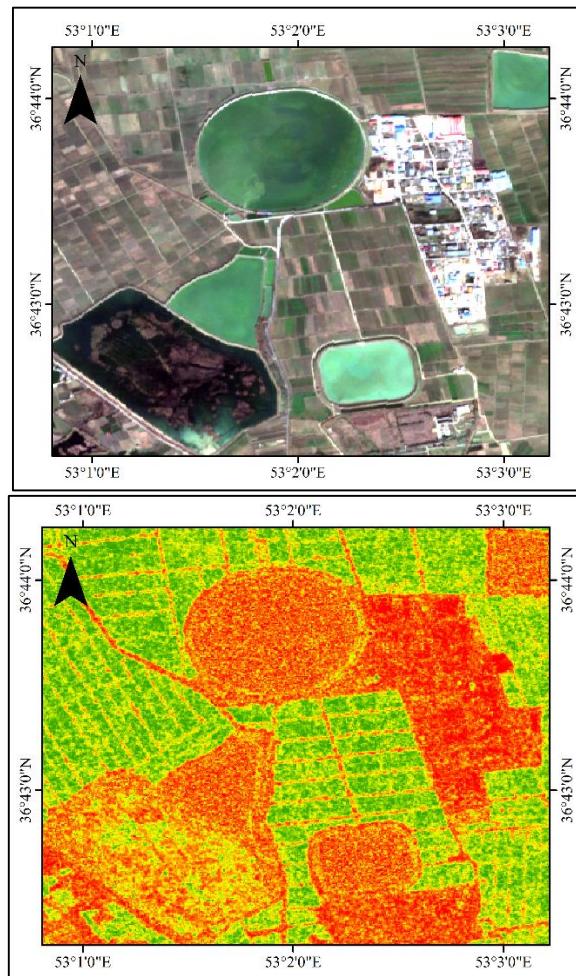
شکل (۳): سری زمانی ضریب باز پراکنش در کشت‌ها و کاربری‌های مختلف محدوده مطالعاتی

Fig. (3): The backscatter coefficient time series for different crops and land uses

#### تفاوت حداکثر و حداقل ضریب باز پراکنش (BD)

برای ایجاد تمایز بین مزارع برنج و سایر پوشش‌ها از تفاوت بین حداکثر و حداقل ضریب باز پراکنش (BD) در دوره رشد برنج استفاده شد. برای این منظور میانگین تصاویر سنتینل ۱ در بازه‌ی زمانی ۲۰۲۱/۰۴/۱۰ تا ۲۰۲۱/۰۵/۱۰ به عنوان حداقل و میانگین تصاویر سنتینل ۱ در بازه‌ی زمانی ۲۰۲۱/۰۷/۱۰ تا ۲۰۲۱/۰۸/۱۰ به عنوان حداکثر ضریب باز پراکنش در نظر گرفته شد. شکل ۴ اختلاف تصاویر حداکثر و حداقل را نشان می‌دهد. با توجه به شکل ۴ مشاهده می‌شود که مزارع برنج به خوبی از سایر کاربری‌ها از جمله پوشش‌های آبی و شهری متمایز شده است.

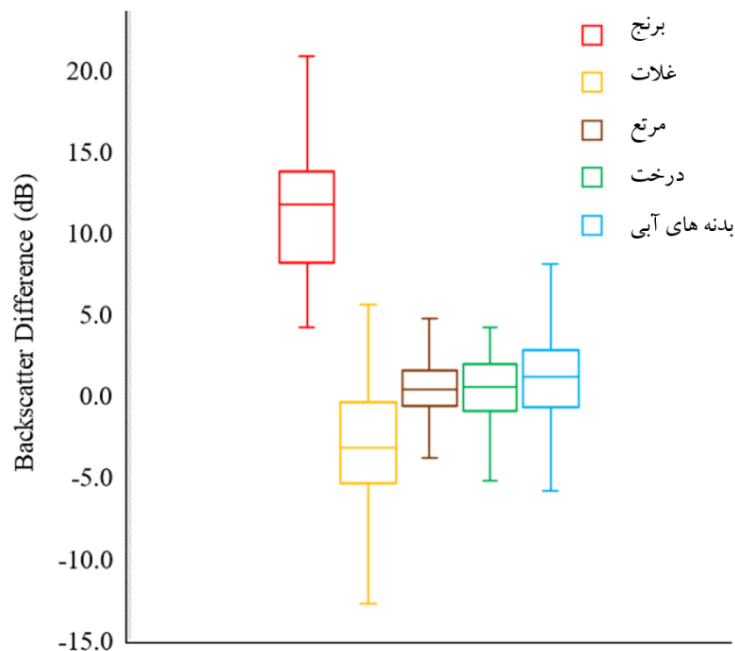




شکل ۴. اختلاف تصاویر حداکثر و حداقل ضریب باز پراکنش (BD) (سمت چپ) و تصویر RGB سنتینل ۲ (سمت راست)  
**Fig4. The difference between the minimum and maximum backscatter coefficient (left) and sentinel 2 RGB image (right)**

شکل ۵ تغییرات مقادیر تصویر BD را در کاربری‌های مختلف نشان می‌دهد. برای این منظور از تصویر BD برای هر یک از کاربری‌ها نمونه جمع‌آوری گردید. با توجه به شکل ۵ مشاهده می‌شود که حداکثر مقدار مربوط به کشت برنج و کمترین مقدار مربوط به کشت غلات است. محدوده تغییرات BD در مزارع برنج از ۴/۲ تا ۲۰/۰۸ است، در حالی که محدوده تغییرات آن در سایر کاربری‌ها کمتر از ۸ است. در برخی نقاط مشاهده می‌شود که بین مقادیر تصویر BD برای محصول برنج، غلات و پهنه‌های آبی اشتراک وجود دارد (محدوده ۴/۲dB تا ۸dB). بنابراین تمایز بین این کاربری‌ها در برخی نقاط با تصویر BD امکان‌پذیر نیست. این می‌تواند به دلیل آبیاری

غرق آبی مزارع باشد. همچنین در طول فصل رشد برنج و با گرم شدن هوا سطح پهنه‌های آبی کاهش پیدا می‌کند و علفزارها به وجود می‌آیند که شرایط مزارع برنج را تداعی می‌کند. این باعث عدم تمایز این نوع نقاط با محصول برنج می‌شود. بنابراین علاوه بر تصویر BD نیاز به ورودی‌های بیشتری به طبقه‌بندی کننده RF است.



شکل ۵. نمودار جعبه‌ای مقادیر تصویر BD در کاربری‌های مختلف

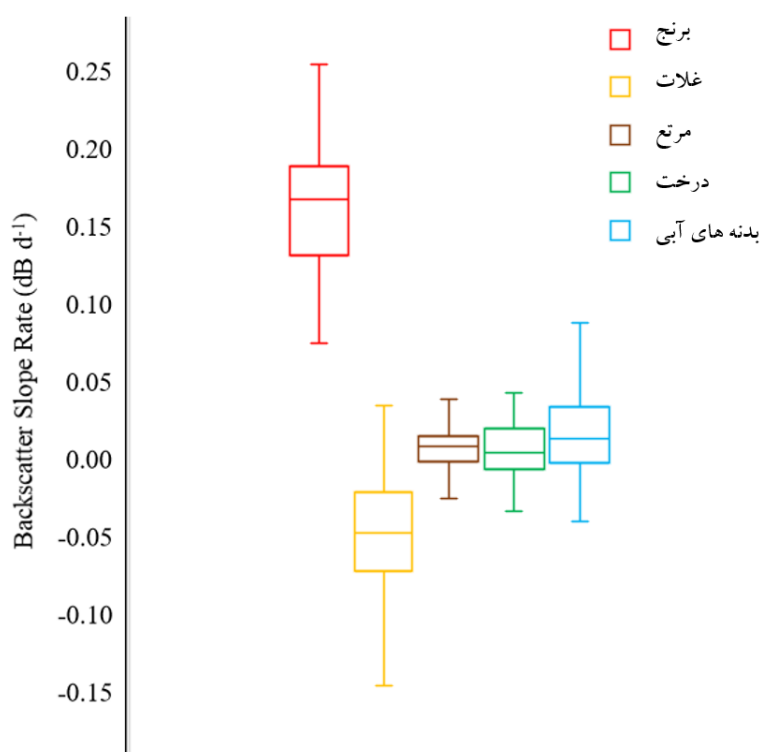
**Fig5. Box plot of BD values in different crops and land uses**

#### شیب ضریب باز پراکنش

با توجه به شکل (۳) مشاهده می‌شود که شیب نمودار سری زمانی ضریب باز پراکنش در مزارع برنج نسبت به سایر کاربری‌ها زیاد است. در طول دوره رشد برنج پوشش کاربری‌های جنگل، شهر و مراتع تقریباً بدون تغییر است و لذا شیب سری زمانی نمودار ضریب باز پراکنش در آن‌ها تقریباً ثابت است. اما با رشد محصول برنج، سایر محصولات برداشت می‌شوند و همین موضوع باعث کاهش ضریب باز پراکنش می‌شود و شیب منفی را شکل می‌دهد. در شکل ۶ مقادیر شیب ضریب باز پراکنش در بازه‌ی زمانی ۲۰۲۱/۰۴/۱۰ تا ۲۰۲۱/۰۸/۱۰ در کاربری‌های مختلف نشان داده شده است. با توجه به شکل ۶ مشاهده می‌شود که شیب ضریب باز پراکنش به خوبی تمایز بین مزارع برنج و سایر کاربری‌ها را نمایان می‌سازد. شیب ضریب باز پراکنش مزارع برنج در بازه

۰/۰۷ تا ۰/۲۶ متغیر است و در مزارع غلات به طور کلی کمتر از صفر است. بنابراین برای تمایز بین مزارع برنج و غلات از این ویژگی استفاده گردید.

با توجه به شکل (۶) مشاهده می‌شود که در برخی از مناطق امکان تفکیک پهنه‌های آبی وجود ندارد. زیرا در پهنه‌های آبی به دلیل کاهش سطح ایستابی و نمایان شدن علفزارها ضریب باز پراکنش افزایش می‌یابد و شیب مثبتی را به وجود می‌آورد. بنابراین برای تفکیک پهنه‌های آبی که کاهش سطح آب آن‌ها باعث رشد علفزارها و تداعی مزارع برنج می‌شود از تصاویر سنتینل ۱ در بازه زمانی که پوشیده از آب هستند (فصل غیر زراعی) استفاده گردید.

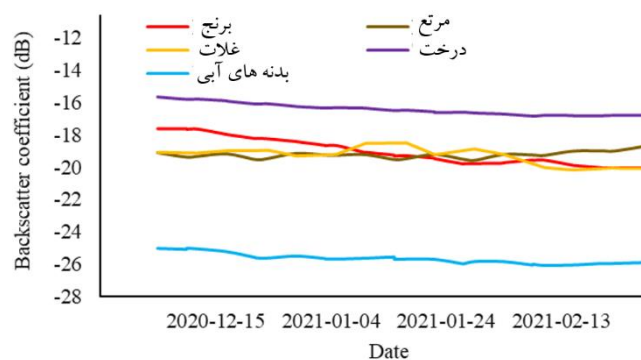


شکل (۶): شیب ضریب باز پراکنش در طول دوره رشد برنج برای کاربری‌های مختلف

Fig. (6): Box plot of BS values in different crops and land uses

آبگیرها و تالاب‌ها در زمستان پوشیده از آب هستند و بنابراین با سایر کاربری‌ها متفاوت هستند. برای بررسی این موضوع تصاویر سنتینل ۱ در بازه زمانی ۲۰۲۰/۱۲/۰۸ تا ۲۰۲۱/۰۲/۲۸ فراخوانی و سری زمانی این تصاویر برای کاربری‌های مختلف ترسیم گردید (شکل ۷). بر اساس شکل مشاهده می‌شود که مقدار ضریب باز پراکنش

برای پهنه‌های آبی کمترین مقدار است و میانگین این تصاویر امکان تفکیک آبگیرها و تالاب‌ها را از مزارع برنج مهیا می‌کند.

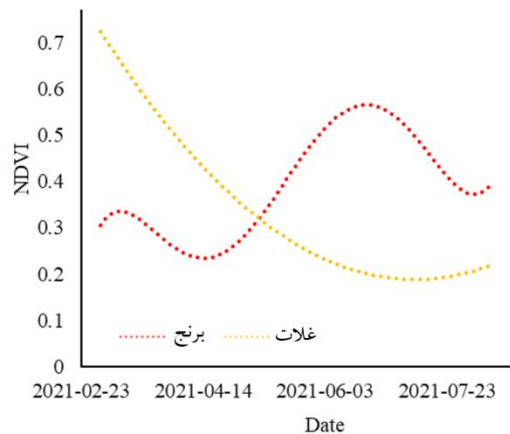


شکل (۷): سری زمانی ضریب باز پراکنش در فصل غیر زراعی برای تفکیک پهنه‌های آبی از مزارع برنج

Fig. (7): The backscatter coefficient time series during the non-cropping season

سری زمانی شاخص پوشش گیاهی نرمال شده

برای افزایش دقت و تفکیک مزارع برنج از غیر برنج از شاخص NDVI استفاده گردید. سری زمانی شاخص NDVI با استفاده از تصاویر سنتینل ۲ در بازه زمانی ۲۰۲۱/۰۳/۰۱ تا ۲۰۲۱/۰۸/۱۰ تهیه گردید. در شکل ۸ روند این شاخص برای کشت غلات و برنج نشان داده شده است. بر اساس شکل مشاهده می‌شود که مقدار NDVI در تابستان برای برنج حداکثر است. در این زمان غلات برداشت و پوشش آن‌ها کم شده است. بنابراین تصویر میانگین NDVI در بازه‌ی ۲۰۲۱/۶/۲۰ تا ۲۰۲۱/۷/۲۰ به عنوان ورودی به طبقه‌بندی کننده RF در نظر گرفته شد.

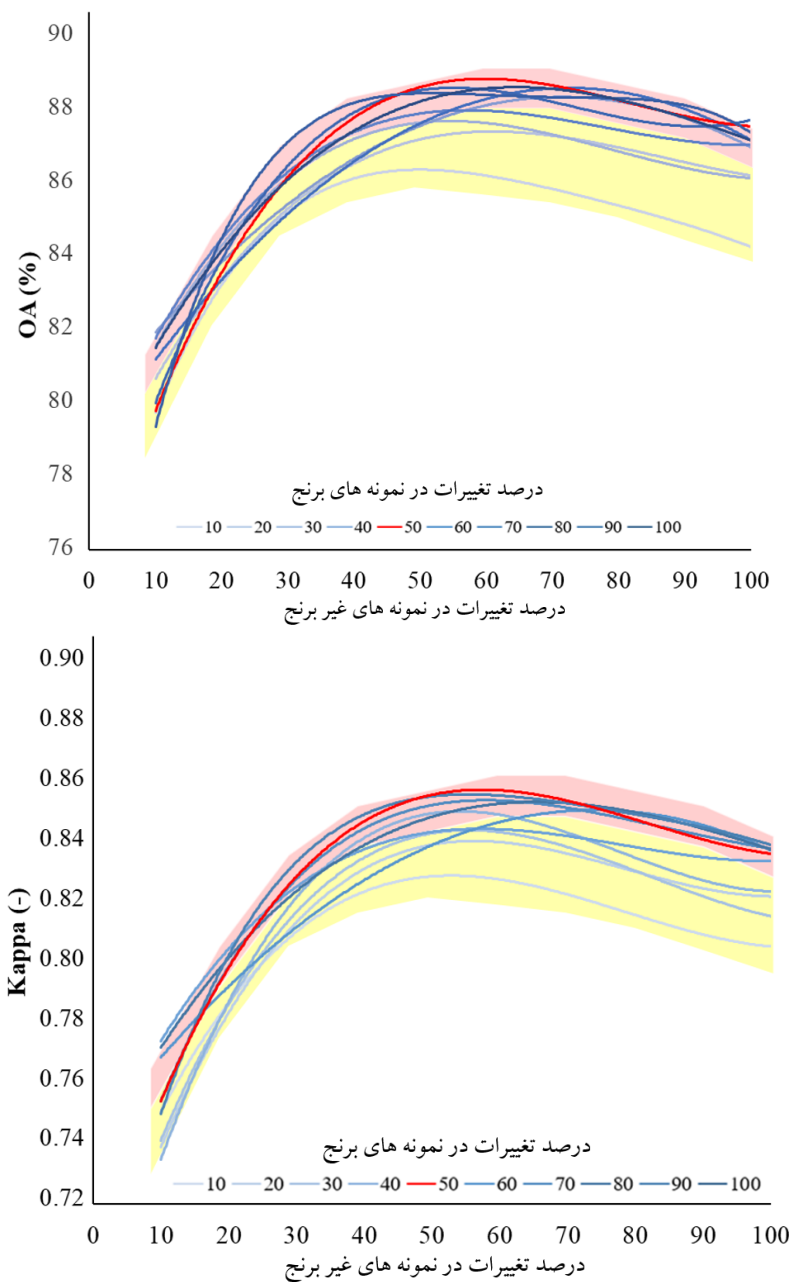


شکل (۸): سری زمانی NDVI کشت برنج و غلات

Fig. (8): The NDVI time series of rice and cereals

#### کلاس‌بندی با تعداد داده‌های مختلف

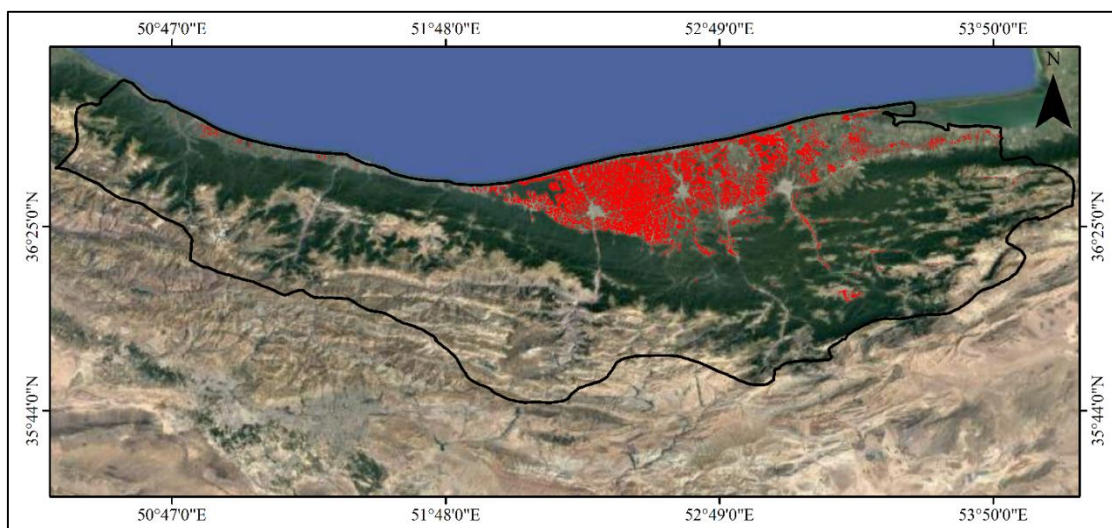
این مرحله با استفاده از ۷۰ درصد کل نمونه‌ها و طبقه‌بندی کننده جنگل تصادفی انجام شد. برای بررسی ارتباط بین تعداد نمونه‌ها و دقت طبقه‌بندی، نمونه‌های آموزشی برنج و غیربرنج به ۱۰ قسمت مساوی تقسیم شدند و در هر اجرا تعداد نمونه‌ها به صورت تجمعی افزایش یافت. بنابراین در مجموع ۱۰۰ اجرای مختلف صورت گرفت و در هر مرحله دقت کلی و ضریب کاپا با استفاده از ۳۰ درصد نمونه‌ها محاسبه شدند. شکل ۹ مقدار دقت کلی و ضریب کاپا را در هر اجرا نشان می‌دهد. با توجه به این شکل افزایش دقت کلی و ضریب کاپا با افزایش تعداد نمونه‌ها رابطه غیرخطی دارد و استفاده از همه نمونه‌ها باعث کاهش دقت کلی و ضریب کاپا شده است. با افزایش تعداد نمونه‌های غیر برنج از ۱۰ تا ۵۰ درصد، دقت کلی و ضریب کاپا افزایش می‌یابد. ولی در این بازه، افزایش تعداد نمونه‌های برنج از ۱۰ تا ۱۰۰ درصد تأثیر چندانی در افزایش دقت و ضریب کاپا نداشته است. همچنین مشاهده می‌شود که افزایش بیشتر از ۵۰ درصد تعداد نمونه‌های غیربرنج تأثیری در افزایش دقت طبقه‌بندی نداشته و باعث کاهش آن شده است. باید در نظر داشت زمانی که تعداد نمونه‌های غیر برنج بیشتر از ۴۰ درصد شده است افزایش تعداد نمونه‌های برنج تأثیر زیادی در تغییرات دقت کلی و ضریب کاپا نداشته است. در این بازه، زمانی که تعداد نمونه‌های برنج از ۱۰ تا ۴۰ درصد افزایش یافته است، دقت کلی از ۸۴ به ۸۷ درصد و مقدار ضریب کاپا از ۰/۸۱ به ۰/۸۵ رسیده است. اما افزایش تعداد نمونه‌های برنج از ۴۰ تا ۱۰۰ درصد تأثیر چندانی بر افزایش دقت نداشته است. بنابراین می‌توان نتیجه گرفت که افزایش تعداد نمونه‌ها لزوماً باعث افزایش دقت نمی‌شود. بیشترین دقت کلی (۸۹ درصد) و ضریب کاپا (۰/۸۶) زمانی اتفاق افتاده که از ۵۰ درصد نمونه‌های غیر برنج و برنج استفاده شده است.



شکل (۹): تغییرات دقت کلی و ضریب کاپا در رابطه با تعداد نمونه‌های آموزشی

Fig. (9): The relationship between the sample size and classification accuracy

بر اساس شکل ۹ بیشترین دقت کلی و ضریب کاپا با استفاده از ۵۰ درصد نمونه‌های آموزشی به دست آمده است. بنابراین برای تعیین سطح زیر کشت برنج از ۵۰ درصد نمونه‌های برنج و غیر برنج استفاده گردید. در شکل ۱۰ نقشه کشت برنج استان مازندران برای سال ۲۰۲۱ نشان داده شده است. بر اساس شکل مساحت زیر کشت برنج برابر ۱۵۸۳۸۴ هکتار محاسبه گردید. بیشترین سطح زیر کشت برنج در دشت‌های منتهی به دریای خزر و کمترین سطح زیر کشت در قسمت‌های جنوبی و غربی استان مازندران به دلیل وجود ارتفاعات رشته کوه البرز مشاهده می‌شود.



شکل (۱۰): نقشه کشت برنج در منطقه مورد مطالعه در سال ۲۰۲۱  
Fig. (10): The rice cultivation map in the study area in 2021

#### نمونه در واحد سطح

بر اساس نتایج بیشترین دقت زمانی به دست آمد که از ۵۰ درصد نمونه‌های برنج و غیربرنج استفاده گردید. بنابراین با این تعداد نمونه طبقه‌بندی صورت گرفت و مساحت زیر کشت برنج و غیربرنج به ترتیب برابر ۱۵۸۳۸۴ و ۲۲۲۵۸۱۶ هکتار به دست آمد. تعداد نمونه‌های برنج و غیربرنج به ترتیب برابر ۲۵۰۰ و ۹۵۰۰ نمونه بود که از ۷۰ درصد آن‌ها برای آموزش استفاده گردید. با در نظر گرفتن ۵۰ درصد نمونه‌های آموزشی هر گروه و مساحت به دست آمده، تعداد نمونه در واحد سطح محاسبه گردید. بنابراین با تعداد یک نمونه در هر ۱۸۱ هکتار از مزارع برنج و تعداد یک نمونه در هر ۶۶۹ هکتار از کاربری‌های غیر برنج بیشترین دقت کلی (۸۹ درصد) و ضریب کاپا (۰/۸۶) به دست آمد. با افزایش دو برابری تعداد نمونه‌ها (یک نمونه در هر ۹۰ هکتار از مزارع برنج و یک نمونه در هر ۳۳۵ هکتار از مزارع غیر برنج) دقت کلی و ضریب کاپا به ترتیب برابر ۰/۸۷ و



۰/۸۴ به دست آمد، که نشان دهنده کاهش دقت با افزایش تعداد نمونه است. به طور کلی نتایج نشان داد که با استفاده از یک نمونه در هر ۳۰۲ هکتار از مزارع برنج و یک نمونه در هر ۱۱۱۶ هکتار از مزارع غیر برنج ضریب کاپا بزرگتر از ۰/۸ است. بنابراین بسته به دقت مورد نیاز می‌توان تعداد نمونه‌های مختلفی استفاده کرد. برای تخمین دقت کلی و ضریب کاپا بر اساس تعداد نمونه در هکتار معادلات (۴) و (۵) ارائه گردید. این معادلات بر اساس ساده‌سازی اطلاعات شکل (۹) به دست آمدند.

$$OA = 7 \times 10^8 \times Sn^3 - 6 \times 10^6 \times Sn^2 + 14877 \times Sn + 4.0469 \times S + 78 \quad (۴)$$

$$K = -33666 \times Sn^2 + 123.99 \times Sn + 4.0469 \times S + 0.6986 \quad (۵)$$

در این معادلات OA دقت کلی (./)، K ضریب کاپا (-)، Sn یک نمونه غیر برنج در هکتار (به عنوان مثال  $\frac{1}{200 ha}$ ) یا  $\frac{1}{500 ha}$  و S یک نمونه برنج در هکتار است. مقادیر S و Sn از تقسیم تعداد نمونه‌های آموزشی بر مساحت تقریبی محصول مورد نظر به دست می‌آید. این روابط تخمینی از دقت کلی و ضریب کاپا را در بازه یک نمونه در ۲۰۰ هکتار تا ۱ نمونه در ۱۰۰۰۰۰ هکتار ارائه می‌دهد.

### نتیجه‌گیری

هدف از این تحقیق بررسی ارتباط تعداد نمونه‌های آموزشی با دقت تشخیص مزارع برنج از غیربرنج توسط طبقه‌بندی کننده جنگل تصادفی بود. علاوه بر این مساحت زیر کشت برنج منطقه مورد مطالعه برای سال ۲۰۲۱ با بیشترین دقت تعیین گردید. برای دستیابی به این هدف از سری زمانی تصاویر ماهواره‌های سنتینل ۱ و ۲ استفاده گردید. نتایج نشان داد که استفاده از پولاریزاسیون VH سنتینل ۱ می‌تواند تفاوت‌های چشمگیری در متمایز کردن کشت برنج از غیر برنج به وجود آورد. بنابراین علاوه بر استفاده از این باند از تصاویر اختلاف حداکثر و حداقل ضریب باز پراکنش و شیب ضریب باز پراکنش در طول دوره رشد محصول برنج به عنوان ورودی به طبقه‌بندی کننده استفاده گردید. همچنین برای افزایش دقت تفکیک‌پذیری سایر کلاس‌ها از محصول برنج از شاخص NDVI نیز استفاده گردید. نتایج نشان داد که افزایش تعداد نمونه‌های آموزشی با افزایش دقت طبقه‌بندی رابطه غیرخطی دارد و با افزایش تعداد نمونه‌ها دقت کاهش پیدا می‌کند. بیشترین دقت کلی (۸۹ درصد) و ضریب کاپا (۰/۸۶) زمانی به دست آمده که از یک نمونه در هر ۱۸۱ هکتار از مزارع برنج و یک نمونه در هر ۶۶۹ هکتار از کاربری‌های غیر برنج استفاده گردید. با افزایش دو برابری تعداد نمونه‌ها (یک نمونه در هر ۹۰ هکتار از مزارع برنج و یک نمونه در هر ۳۳۵ هکتار از مزارع غیر برنج) دقت کلی و ضریب کاپا به ترتیب برابر ۰/۸۷ و ۰/۸۴ به دست آمد. بنابراین می‌توان نتیجه گرفت که صرفاً افزایش تعداد نمونه باعث افزایش دقت طبقه‌بندی نمی‌شود. در نهایت معادلاتی برای تعیین تعداد نمونه مورد نیاز برای دستیابی به دقت موردنظر در



تفکیک مزارع برنج از غیر برنج در منطقه مورد مطالعه ارائه شد. همچنین پیشنهاد می‌شود در طبقه‌بندی تصاویر ماهواره‌ای، همه نمونه‌های آموزشی به یکباره مورد استفاده قرار نگیرد و این کار با تعداد نمونه‌های مختلف انجام شود.

## منابع

- 1- Chang, L., Chen, Y.T., Wang, J.H., Chang, Y.L. (2021). Rice-Field Mapping with Sentinel-1A SAR Time-Series Data. *Remote Sens.* 13, 103, doi:10.3390/RS13010103.
- 2- Mansaray, L.R., Huang, W., Zhang, D., Huang, J., Li, J. 2017. Mapping Rice Fields in Urban Shanghai, Southeast China, Using Sentinel-1A and Landsat 8 Datasets. *Remote Sens.* 9, 257, doi:10.3390/RS9030257.
- 3- Mohanty, S., Wassmann, R., Nelson, A., Moya, P., Jagadish, S.V.K. (2013). Rice and Climate Change: Significance for Food Security and Vulnerability. *Philipp. Int. Rice Res. Inst.*, 49, 1–14.
- 4- Bouvet, A., Le Toan, T. (2011). Use of ENVISAT/ASAR Wide-Swath Data for Timely Rice Fields Mapping in the Mekong River Delta. *Remote Sens. Environ.* 115, 1090–1101, doi:10.1016/J.RSE.2010.12.014.
- 5- Kamkar, B., Dashtimarvili, M., Kazemi, H. (2019). Detection of Rice and Soybean Grown Fields and Their Related Cultivation Area Using Sentinel-2 Satellite Images in Summer Cropping Patterns to Analyze Temporal Changes in Their Cultivation Area (Case Study: Four Watershed Basins of Golestan Province). *J. Water Soil Conserv.* 26, 151–167, doi:10.22069/jwsc.2019.15246.3044.
- 6- Ziaiean Firouzabadi, p., Sayyad Bidhendi, l., eskandari nodeh, m. (2010). Mapping and Acreage Estimating of Rice Agricultural Land Using RADARSAT a Satellite Images. *Phys. Geogr. Res. Q.*, 41.
- 7- Jin, C., Xiao, X., Dong, J., Qin, Y., Wang, Z. (2016). Mapping Paddy Rice Distribution Using Multi-Temporal Landsat Imagery in the Sanjiang Plain, Northeast China. *Front. Earth Sci.* 10, 49–62, doi:10.1007/S11707-015-0518-3.
- 8- Sakamoto, T., Sprague, D.S., Okamoto, K., Ishitsuka, N. (2018). Semi-Automatic Classification Method for Mapping the Rice-Planted Areas of Japan Using Multi-Temporal Landsat Images. *Remote Sens. Appl. Soc. Environ.* 10, 7–17, doi:10.1016/J.RSASE.2018.02.001.
- 9- Zhang, G., Xiao, X., Dong, J., Kou, W., Jin, C., Qin, Y., Zhou, Y., Wang, J., Menarguez, M.A., Biradar, C. (2015). Mapping Paddy Rice Planting Areas through Time Series Analysis of MODIS Land Surface Temperature and Vegetation Index Data. *ISPRS J. Photogramm. Remote Sens.* 106, 157–171, doi:10.1016/J.ISPRSJPRS.2015.05.011.
- 10- Saadat, M., Hasanlou, M., Homayouni, S. (2019). RICE CROP MAPPING USING SENTINEL-1 TIME SERIES IMAGES (CASE STUDY: MAZANDARAN, IRAN). *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* XLII-4/W18, 897–904, doi:10.5194/isprs-archives-XLII-4-W18-897-2019.
- 11- Inoue, S., Ito, A., Yonezawa, C. (2020). Mapping Paddy Fields in Japan by Using a Sentinel-1 SAR Time Series Supplemented by Sentinel-2 Images on Google Earth Engine. *Remote Sens.* 12, 1622, doi:10.3390/RS12101622.

- 12- Torbick, N., Chowdhury, D., Salas, W., Qi, J. (2017). Monitoring Rice Agriculture across Myanmar Using Time Series Sentinel-1 Assisted by Landsat-8 and PALSAR-2. *Remote Sens.* 9, 119, doi:10.3390/RS9020119.
- 13- Amani, M., Kakooei, M., Moghimi, A., Ghorbanian, A., Ranjgar, B., Mahdavi, S., Davidson, A., Fisette, T., Rollin, P., Brisco, B. (2020). Application of Google Earth Engine Cloud Computing Platform, Sentinel Imagery, and Neural Networks for Crop Mapping in Canada. *Remote Sens.* 12, 1–18, doi:10.3390/rs12213561.
- 14- Tiwari, V., Matin, M.A., Qamer, F.M., Ellenburg, W.L., Bajracharya, B., Vadrevu, K., Rushi, B.R., Yusafi, W. (2020). Wheat Area Mapping in Afghanistan Based on Optical and SAR Time-Series Images in Google Earth Engine Cloud Environment. *Front. Environ. Sci.* 8, doi:10.3389/fenvs.2020.00077.
- 15- Talema, T., Hailu, B.T. (2020). Mapping Rice Crop Using Sentinels (1 SAR and 2 MSI) Images in Tropical Area: A Case Study in Fogera Wereda, Ethiopia. *Remote Sens. Appl. Soc. Environ.* 18, doi:10.1016/j.rsase.2020.100290.
- 16- Darzi, A., Karandish, F. (2016). Rice Cultivation Management in Mazandaran Province under Climate Change. *J. Water Res. Agric.* 30, 333–346, doi:10.22092/jwra.2016.107154.
- 17- Farsadnia, F., Rostami Kamrod, M., Moghadam Nia, A. (2012). Rainfall Trend Analysis of Mazandaran Province Using Regional Mann-Kendall Test. *Iran-Water Resour. Res.* 8, 60–70.
- 18- Goli, I., Azadi, H., Miceikienė, A., Tanaskovik, V., Stamenkovska, I.J., Kurban, A., Viira, A.H. (2022). Training Needs Assessment: The Case of Female Rice Farmers in Northern Iran. *Agric.* 12, 390, doi:10.3390/AGRICULTURE12030390/S1.
- 19- Congalton, R.G., Green, K. (2008). *Assessing the Accuracy of Remotely Sensed Data*, CRC Press, ISBN 9780429143977.
- 20- Taheri, M., Emadzadeh, M., Gholizadeh, M., Tajrishi, M., Ahmadi, M., Moradi, M. (2019). Investigating the Temporal and Spatial Variations of Water Consumption in Urmia Lake River Basin Considering the Climate and Anthropogenic Effects on the Agriculture in the Basin. *Agric. Water Manag.* 213, 782–791, doi:10.1016/J.AGWAT.2018.11.013.