

## ارزیابی صحت الگوریتم‌های طبقه‌بندی نظارت شده برای استخراج نقشه کاربری اراضی (مطالعه موردی: حوضه آبخیز تهم)

قباد رستمی‌زاد\*: استادیار بخش تحقیقات حفاظت خاک و آبخیزداری، مرکز تحقیقات و آموزش کشاورزی و منابع طبیعی استان زنجان، سازمان

تحقیقات، آموزش و ترویج کشاورزی، زنجان، ایران

زهرا خان‌بابایی: دانش‌آموخته دکتری ژئومورفولوژی از دانشگاه تهران و دبیر آموزش و پرورش استان زنجان

محمد طهمورث: استادیار بخش تحقیقات حفاظت خاک و آبخیزداری پژوهشکده حفاظت خاک و آبخیزداری، سازمان تحقیقات، آموزش و ترویج

کشاورزی، زنجان، ایران

نوع مقاله: پژوهشی

تاریخ پذیرش: ۱۴۰۱/۰۴/۲۲

تاریخچه مقاله (تاریخ دریافت): ۱۴۰۰/۱۰/۳۰



20.1001.1.22517812.1401.12.4.7.8

### چکیده

یکی از عوامل مهم در مطالعات منابع طبیعی و محیط زیست، شناخت کاربری‌ها و تعیین سطح گسترش آنها در منطقه است. داده‌های ماهواره‌ای، یکی از سریع‌ترین و کم هزینه‌ترین روش‌های تهیه نقشه کاربری اراضی است. پژوهش حاضر با هدف تعیین بهترین الگوریتم طبقه‌بندی تصاویر ماهواره‌ای لندست ۸ مرداد ماه سال ۱۴۰۰، از بین چهار روش طبقه‌بندی نظارت شده به منظور استخراج نقشه کاربری اراضی حوضه آبخیز سد تهم انجام شد. در این مطالعه از روش طبقه‌بندی نظارت شده تصاویر ماهواره‌ای و چهار الگوریتم ماشین‌بردار پشتیبان استفاده شد و حداکثر احتمال، حداقل فاصله از میانگین و فاصله ماکسیمی، پنج کلاس کاربری برای منطقه شامل اراضی کشاورزی (۳/۲۶٪)، اراضی مرتع (۸۹/۸۵٪)، مسکونی (۰/۱۵٪)، پهنه آبی (۱/۱۹٪) و اراضی بایر (۵/۵۵٪) استخراج شد. همچنین نتایج نشان داد که الگوریتم ماشین‌بردار پشتیبان با دقت کلی ۹۸/۲۴٪ و ضریب کاپا ۰/۹۶۹، به عنوان دقیق‌ترین الگوریتم مشخص شد و الگوریتم حداقل فاصله از میانگین با دقت کلی ۹۴/۳۶٪ و ضریب کاپا ۰/۹، دقت کمتری داشت. با استفاده از یافته‌های این پژوهش می‌توان فرسایش منطقه را تا حدود زیادی کاهش داد و از منابع موجود در منطقه به طور بهینه استفاده کرد که برآیند آن جلوگیری از کاهش کیفیت آب پشت سد برای مصرف آب شرب شهر زنجان و جلوگیری از ورود رسوبات بیشتر به دریاچه سد و کاهش حجم آن خواهد بود.

واژگان کلیدی: استفاده از زمین، تصاویر ماهواره‌ای، دقت کلی، سنجش از دور، ضریب کاپا.

## ۱- مقدمه

کاربری اراضی به عنوان نحوه استفاده از اراضی توسط مردم و زیستگاه آنها و معمولاً با تأکید بر نقش عملکردی اراضی برای فعالیت‌های اقتصادی تعریف می‌شود؛ در حالی که پوشش اراضی یک ویژگی فیزیکی سطح اراضی است (Mariye et al, 2020 & Watson et al, 2000). دینامیک پوشش اراضی / کاربری اراضی<sup>1</sup> (LULC)، فرایندی شناخته شده، شتاب‌دهنده و اساسی است که عمدتاً توسط فعالیت‌های انسانی هدایت می‌شود و به طور قابل توجهی به تکه تکه شدن جنگل‌ها و مراتع، تخریب اراضی و از دست دادن تنوع زیستی کمک می‌کند (Abbas and Jaber, 2020 & Haregeweyn et al, 2015 & Maitima et al, 2009). بررسی تغییرات پوشش اراضی / کاربری اراضی (LULC)، یکی از مفیدترین روش‌ها برای درک چگونگی استفاده از زمین در سال‌های گذشته، برای استفاده‌های قابل انتظار در آینده و شناسایی نیروهای محرک این تغییرات است (Haregeweyn et al, 2015 & Maitima et al, 2009). علاوه بر تغییرات طبیعی، افزایش جمعیت انسان نیز تغییراتی در سطح زمین به وجود می‌آورد که بی‌سابقه است (Goldewijk and Ramankutty, 2004). در واقع، تغییرات عمده LULC در سراسر جهان می‌تواند به توسعه شدید کشاورزی و جمعیت رو به رشد مربوط باشد (Tolessa et al, 2020). LULC‌ها مستقیماً بر تنوع زیستی در سراسر جهان تأثیر می‌گذارند، به تغییرات آب و هوایی محلی و منطقه‌ای و گرم شدن آب و هوای جهانی کمک می‌کنند، بر توانایی سیستم‌های بیولوژیکی برای پشتیبانی و انطباق با نیازهای انسان تأثیر می‌گذارند و یکی از منابع اولیه تخریب خاک هستند (Alemayehu et al, 2019 & Chang et al, 2018 & Lambin et al, 2001). بنابراین، نیاز به ارزیابی بهتر تغییرات پوشش اراضی (یعنی ویژگی‌های بیوفیزیکی سطح زمین) و کاربری اراضی برای اهداف انسانی، برای درک تغییرات گذشته و ترسیم روندهای آتی برای دهه‌های آینده ضروری می‌باشد. البته این کار بدون استفاده از فناوری‌های جدید و سنجش از دور بسیار هزینه‌بر و زمان‌بر خواهد بود.

سنجش از دور و سیستم اطلاعات جغرافیایی با استفاده از تصاویر ماهواره‌ای با وضوح رادیومتری، طیفی، مکانی و زمانی متفاوت، می‌تواند به شناسایی یا پایش عوارض سطح زمین بپردازند. این فناوری‌ها در به حداقل رساندن زمان و هزینه استخراج نقشه‌های پوشش زمین و کاربری آن، چندین مزیت دارند. همچنین سنجش از راه دور، انواع مهم و متفاوتی از منابع داده از راه دور را برای استخراج اطلاعات LULC فراهم می‌کند (Dibs and Al-Hedny, 2019). استفاده از روش سنجش از دور، به طور گسترده‌ای رو به افزایش است و از آن برای اجرای طبقه‌بندی LULC در سراسر جهان استفاده می‌شود (Dibs, Al-Hedny et al, 2018). با استفاده از این روش می‌توان از اطلاعات به روز شده در مورد تمام ویژگی‌هایی که در سطح زمین قرار دارند، بهره برد (Dibs, Al-Janabi et al, 2018). Huang و همکاران (2002)، از الگوریتم SVM<sup>2</sup> برای طبقه‌بندی LULC از تصاویر نقشه‌بردار موضوعی Landsat (TM) استفاده کردند. نتایج آنها نشان داد که طبقه‌بندی SVM نسبت به روش‌هایی مانند حداکثر احتمال و درخت تصمیم، دقت بیشتری دارد. Yousefi و همکاران (2011)، ۹ الگوریتم طبقه‌بندی نظارت شده برای نقشه LULC در استان مازندران، ایران و با استفاده از تصاویر Landsat ETM<sup>+</sup> ارزیابی کردند. نتایج نشان داد که الگوریتم SVM نسبت به

<sup>1</sup> Land use and Land cover

<sup>2</sup> Support Vector Machine

سایر الگوریتم‌ها دقت بهتری دارد. Khazaei و همکاران (2019)، مقایسه دقت انواع روش‌های طبقه‌بندی (هشت روش) را در تهیه نقشه کاربری اراضی در شهر یزد بررسی کردند. نتایج آنها نشان داد که الگوریتم‌های حداکثر احتمال، فاصله ماهالانویی و ماشین‌بردار پشتیبان، نسبت به دیگر روش‌ها دقت بیشتری دارند. Shafiee و همکاران (2016) با استفاده از الگوریتم‌های حداکثر احتمال و فازی، نقشه پوشش اراضی شهرستان اسلام‌شهر را تهیه کردند. نتایج این پژوهش نشان داد که الگوریتم حداکثر احتمال نسبت به روش فازی، دقت کلی بیشتر و ضریب کاپای بالاتری دارد. کارایی روش‌های طبقه‌بندی تصاویر ماهواره‌ای در ارزیابی تغییرات کاربری اراضی شهری در شهر اردبیل نیز بررسی شد (Jafarzadeh and Nazmfar, 2019). نتایج این پژوهش نشان داد که روش شبکه عصبی با دقت کلی ۹۴/۴۷ و ضریب کاپای ۰/۹۲، نسبت به شش روش دیگر دقت بالاتری دارد.

مقایسه الگوریتم‌های طبقه‌بندی نوع کاربری در تهیه نقشه کاربری اراضی در حوضه آبخیز زنوز چای استان آذربایجان شرقی نیز بررسی شد (Moradi and Rezaei, 2021). در این پژوهش، بررسی صحت طبقه‌بندی تصویر TM نشان داد که الگوریتم حداکثر احتمال با صحت کلی ۷۳/۸۴٪ و ضریب کاپای ۰/۶۵، در طبقه‌بندی تصویر در چهار کلاس کاربری اراضی کارایی بالاتری نسبت به دیگر روش‌ها دارد. Davooupour and Torangzar (2021)، از تصاویر ماهواره‌ای لندست ۷ برای تعیین پوشش و کاربری اراضی حوضه آبخیز هندو در استان مرکزی استفاده کردند. در این مطالعه، طبقه‌بندی به روش نظارت شده و با چهار الگوریتم مختلف طبقه‌بندی بر روی گروه‌های سه بانندی صورت گرفت که شامل حداکثر احتمال، حداقل فاصله تا میانگین، فاصله ماهالانویی و متوازی السطوح بود. در نهایت، بهترین نتیجه طبقه‌بندی در مقایسه با نقشه واقعیت زمینی، با دقت کلی ۸۱/۲۵٪ در الگوریتم طبقه‌بندی حداکثر احتمال به دست آمد که بیانگر قابلیت و کارایی نسبتاً خوب داده‌های ETM<sup>+</sup> در تهیه نقشه نهایی کاربری اراضی است.

بنابراین با توجه به اهمیت نحوه استفاده از اراضی و حفظ پوشش سطحی آن، حفظ منابع آب و خاک و جلوگیری از تخریب این منابع، تهیه نقشه کاربری اراضی برای کمک به مدیران و تصمیم‌گیران امری ضروری به نظر می‌رسد. علاوه بر این، حوضه آبخیز تهم به دلیل وجود سد تهم در خروجی آن و تأمین آب شرب شهر زنجان از طریق این سد، موقعیت استراتژیک و مهمی داد. بر همین اساس، این مطالعه با هدف تهیه نقشه کاربری اراضی و تعیین دقیق‌ترین الگوریتم طبقه‌بندی برای استخراج نقشه کاربری اراضی در حوضه آبخیز تهم در شهر زنجان انجام شد.

## ۲- منطقه مورد مطالعه

حوضه آبخیز تهم با مساحت ۲۰۹۲۵ هکتار، در استان زنجان و شهرستان زنجان و بین طول شرقی ۱۷° ۴۸ تا ۳۷° ۴۸ و عرض شمالی ۳۶° ۴۶ تا ۳۶° ۴۶ قرار دارد (شکل ۱). حداکثر ارتفاع حوضه، ۲۶۴۴ متر و حداقل ارتفاع در خروجی حوضه و معادل ۱۴۸۰ متر از سطح دریا است. از نظر هیدرولوژیکی، حوضه مطالعاتی تهم یکی از سرشاخه‌های رودخانه زنجان رود، از زیر حوضه‌های قزل‌اوزن است. مراکز جمعیتی حوضه عبارتند از: روستاهای تهم، تازه‌کند، خشکه رود، گله‌رود، طاهرآباد، پالتو و سارمساقلو. شیب متوسط وزنی این حوضه ۴۷/۲٪، متوسط دمای سالانه ۹/۷ درجه سانتی‌گراد و بارندگی متوسط سالانه آن معادل ۳۸۲/۴ میلی‌متر است. حوضه تهم در رشته کوه طارم قرار دارد و عمدتاً از طبقات ولکانیک دوره ائوسن تشکیل شده است. با توجه به وضعیت مرفولوژیکی کوهستانی منطقه و شیب نسبتاً بالا،

کاربری عمده و مناطق اطراف آن از نوع مرتع است. وسعت کاربری‌های دیگر مثل زراعت و باغ در منطقه نیز نسبت به کاربری مرتع بسیار اندک است.

### ۳- مواد و روش

این مقاله در دو بخش تهیه شده است: بخش اول شامل طبقه‌بندی کاربری / پوشش اراضی با استفاده از روش‌های طبقه‌بندی نظارت شده<sup>۱</sup> و بخش دوم شامل ارزیابی دقت مدل‌ها است که با استفاده از ضرایب صحت کلی<sup>۲</sup>، ضریب کاپا<sup>۳</sup>، دقت تولیدکننده<sup>۴</sup>، دقت کاربر<sup>۵</sup>، خطای اضافه<sup>۶</sup> و خطای حذف<sup>۷</sup> انجام شد.

فرایند طبقه‌بندی و تجزیه و تحلیل کلاس‌های مختلف کاربری اراضی، با استفاده از تصویر ماهواره‌ای لندست ۸ (مسیر ۱۶۷، ردیف ۰۳۴) مرداد ماه ۱۴۰۰ انجام شد. تصاویر لندست، از کاوشگر زمین‌شناسی ایالات متحده (USGS) بارگیری شده است (<https://earthexplorer.usgs.gov/>). انتخاب تاریخ تصاویر ماهواره‌ای لندست، تحت تأثیر کیفیت تصویر به ویژه برای تصویری است که پوشش ابری محدود یا کم دارند. سپس تصویر در نرم‌افزار ENVI 5.3 پردازش و با برش منطقه مورد مطالعه با استفاده از نرم‌افزار استخراج شد. از نرم‌افزار GIS 10.5<sup>۸</sup> برای ایجاد نقشه نقاط تعلیمی و ارزیابی، نقشه واقعیت زمینی و تبدیل فرمت لایه استفاده شد (شکل ۲). همچنین علاوه بر برداشت‌های زمینی، از نرم‌افزار گوگل ارث پرو ۹ برای برداشت نقاط استفاده شد. با توجه به هدف پژوهش فقط از باندهای مرئی، مادون قرمز نزدیک، مادون قرمز و طول موج کوتاه (با اندازه سلول سی متر) برای استخراج ارزهای طیفی متناظر با قطعه نمونه-های زمینی و تجزیه و تحلیل آماری استفاده شد. تصاویر خام معمولاً خیلی تاریک یا خیلی روشن است و توانایی تفکیک عوارض برای آن به خوبی میسر نمی‌باشد. بنابراین، بارسازی به روش توازن هیستوگرام انجام می‌شود که توزیع مقادیر پیکسل‌ها بر اساس فراوانی آنها است. این روش اغلب برای افزایش کنتراست عوارض خاص - که معمولاً فراوانی بیشتری دارند - مؤثر می‌باشد. انتخاب بهترین ترکیب باندی شاخص مطلوبیت، شاخص آماری است که بر اساس انحراف معیار و ضریب همبستگی بین باندها می‌باشد.

<sup>1</sup> Supervised classification

<sup>2</sup> Overall Accuracy

<sup>3</sup> Kappa coefficient

<sup>4</sup> Producer Accuracy

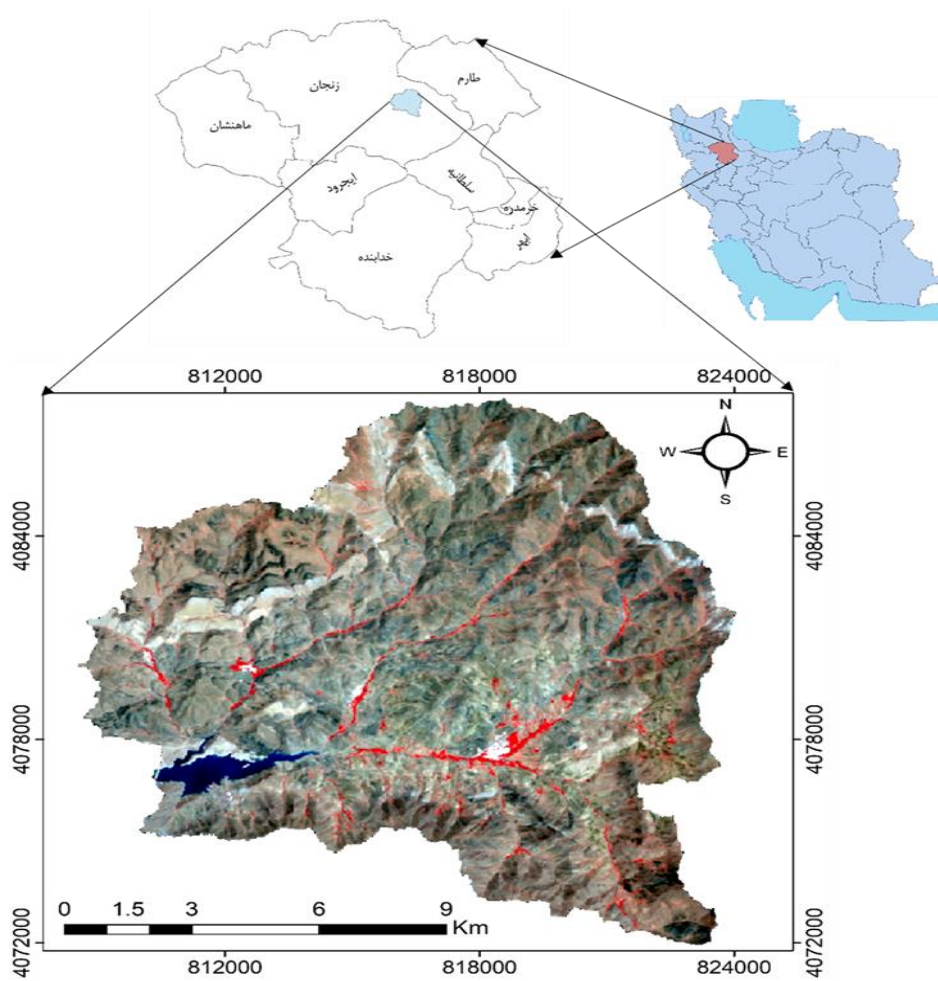
<sup>5</sup> User Accuracy

<sup>6</sup> Commission

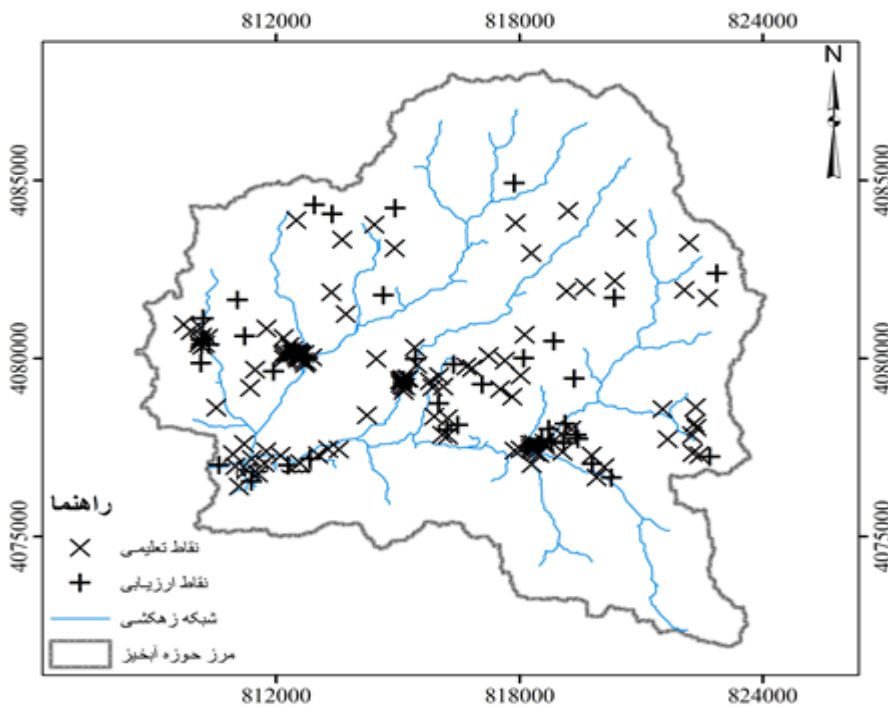
<sup>7</sup> Omission

<sup>8</sup> ARC GIS 10.5

<sup>9</sup> Google Earth Pro



شکل ۱: موقعیت جغرافیایی منطقه مورد مطالعه



شکل ۲: موقعیت نقاط تعلیمی و ارزیابی برداشت‌شده در منطقه مورد مطالعه

## روش‌های طبقه‌بندی

در این تحقیق برای تهیه نقشه کاربری اراضی از روش‌های طبقه‌بندی نظارت شده حداکثر احتمال<sup>۱</sup>، حداقل فاصله از میانگین<sup>۲</sup>، فاصله ماهالونویی<sup>۳</sup> و ماشین بردار پشتیبان<sup>۴</sup> استفاده شد.

- روش طبقه‌بندی حداکثر احتمال: در این روش، میزان کمی واریانس و همبستگی ارزش‌های طیفی باندهای مختلف برای مناطق نمونه محاسبه و از همین خاصیت، برای ارتباط یک پیکسل طبقه‌بندی نشده به یکی از گروه‌ها یا نمونه‌های طیفی نیز استفاده می‌شود (shafiee et al, 2016).

- روش طبقه‌بندی حداقل فاصله: اولین مرحله در این روش، تغییر مقادیر طیفی هر باند برای هر کلاس است. پس از تعیین متوسط ارزش طیفی پیکسل، فاصله هر پیکسل طبقه‌بندی نشده با پیکسل‌های میانگین مقایسه می‌شود و پیکسل مورد نظر به کلاسی اختصاص می‌یابد که کمترین فاصله را با میانگین دارد؛ به عبارت دیگر، در این روش پیکسل مجهول به کلاسی واگذار می‌شود که بردار میانگین آن به کلاس موردنظر نزدیک‌ترین فاصله اقلیدسی را داشته باشد (yousefi et al, 2011).

- روش فاصله ماهالانویی: فاصله ماهالانویی، طبقه‌بندی دیگری نظیر روش کمترین فاصله است. در این روش، اساس باز هم شبیه روش کمترین فاصله است اما نه کمترین فاصله اقلیدسی، بلکه از کمترین فاصله ماهالانویی استفاده می‌کند (khazaei et al, 2019).

- روش ماشین بردار پشتیبان: روش ماشین بردار پشتیبان، مرزی است که با معیار قرار دادن بردارهای پشتیبان، بهترین طبقه‌بندی و تفکیک بین داده‌ها را مشخص می‌کند. در SVM فقط داده‌های قرار گرفته در بردارهای پشتیبان، مبنای یادگیری ماشین و ساخت مدل است. این الگوریتم به سایر نقاط داده حساس نیست و هدف آن هم یافتن بهترین مرز در بین داده‌ها است؛ به گونه‌ای که بیشترین فاصله ممکن را از تمام دسته‌ها (بردارهای پشتیبان آنها) داشته باشد (yousefi et al, 2011).

به این ترتیب که برای کاربری‌های موجود، ابتدا در منطقه با استفاده از بازدیدهای میدانی و دستگاه موقعیت‌یاب جهانی (GPS)، نمونه‌های تعلیمی (۱۶۰ نقطه تعلیمی) به تفکیک هر کاربری در منطقه برداشت شد. سپس این نمونه‌ها به دو دسته تقسیم‌بندی شد که از هفتاد درصد آنها برای استفاده در طبقه‌بندی و از سی درصد برای بررسی صحت روش-های طبقه‌بندی استفاده شد. پس از اعمال تصحیحات تصاویر با توجه به الگوریتم‌های مورد مطالعه برای هر الگوریتم با استفاده از نرم‌افزار EVVI 5.3 و نمونه‌های تعلیمی، نقشه کاربری اراضی برای منطقه مطالعاتی تهیه شد. با توجه به نمونه‌برداری انجام شده، پنج کلاس کاربری در منطقه مورد مطالعه مشخص شد که شامل زمین کشاورزی، مرتع، زمین بایر، مناطق مسکونی و پهنه‌های آبی است.

<sup>1</sup> Maximum likelihood

<sup>2</sup> Minimum distance

<sup>3</sup> Mahalanobis distance

<sup>4</sup> Support Vector Machine

## ارزیابی دقت

ارزیابی دقت و صحت نقشه‌های طبقه‌بندی شده به روش‌های مختلف، با ماتریس خطا<sup>۱</sup> و با استفاده از نرم‌افزار ENVI محاسبه شد. برای این کار، در تحقیق حاضر از ضرایب صحت کلی<sup>۲</sup>، ضریب کاپا<sup>۳</sup>، دقت تولیدکننده<sup>۴</sup>، دقت کاربر<sup>۵</sup>، خطای اضافه<sup>۶</sup> و خطای حذف<sup>۷</sup> استفاده شد. دقت کلی، از جمع عناصر قطر اصلی ماتریس خطا تقسیم بر تعداد کل پیکسل‌ها طبق رابطه<sup>۱</sup> به دست می‌آید (Abbas and Jaber, 2020):

$$OA = 1/N \sum P_{ii} \quad \text{رابطه ۱}$$

در این رابطه OA دقت کلی، N تعداد پیکسل‌های آزمایشی،  $\sum P_{ii}$  جمع عناصر قطر اصلی ماتریس خطا است. به دلیل ایرادات وارده بر دقت کلی، غالباً در کارهای اجرایی که مقایسه دقت طبقه‌بندی قابل توجه است، از شاخص کاپا استفاده می‌شود؛ چون شاخص کاپا پیکسل‌های نادرست طبقه‌بندی شده را مدنظر قرار می‌دهد. این شاخص از رابطه<sup>۲</sup> محاسبه می‌شود (Abbas and Jaber, 2020):

$$Kappa = (P_o - P_c / 1 - P_c) * 100 \quad \text{رابطه ۲}$$

که در رابطه بالا،  $P_o$  درستی مشاهده شده و  $P_c$  توافق مورد انتظار است. دقت تولیدکننده عبارت است از احتمال اینکه یک پیکسل در تصویر کلاسه‌بندی در همان کلاس در روی زمین قرار بگیرد و دقت کاربر، احتمال اینکه یک کلاس مشخص در روی زمین در همان کلاس بر روی تصویر طبقه‌بندی شده قرار بگیرد که از روابط ۳ و ۴ محاسبه می‌شود (Abbas and Jaber, 2020):

$$PA = t_a / g_a * 100 \quad \text{رابطه ۳}$$

$$UA = t_a / n_1 * 100 \quad \text{رابطه ۴}$$

در این روابط PA درصد دقت کلاس a برای دقت تولیدکننده،  $t_a$  تعداد پیکسل‌های صحیح طبقه‌بندی شده به عنوان کلاس a،  $g_a$  تعداد پیکسل‌های کلاس a در واقعیت زمینی، UA درصد دقت کلاس a برای دقت کاربر و  $n_1$  تعداد پیکسل‌های کلاس a در نتیجه طبقه‌بندی است. بر اساس دو دقت ذکر شده، خطای لحاظ شده و حذف شده به صورت روابط (۵ و ۶) تعریف می‌شود (Abbas and Jaber, 2020):

$$C_a = 1 - UA \quad \text{رابطه ۵}$$

$$O_e = 1 - PA \quad \text{رابطه ۶}$$

<sup>1</sup> Confusion matrix

<sup>2</sup> Overall Accuracy

<sup>3</sup> Kappa coefficient

<sup>4</sup> Producer Accuracy

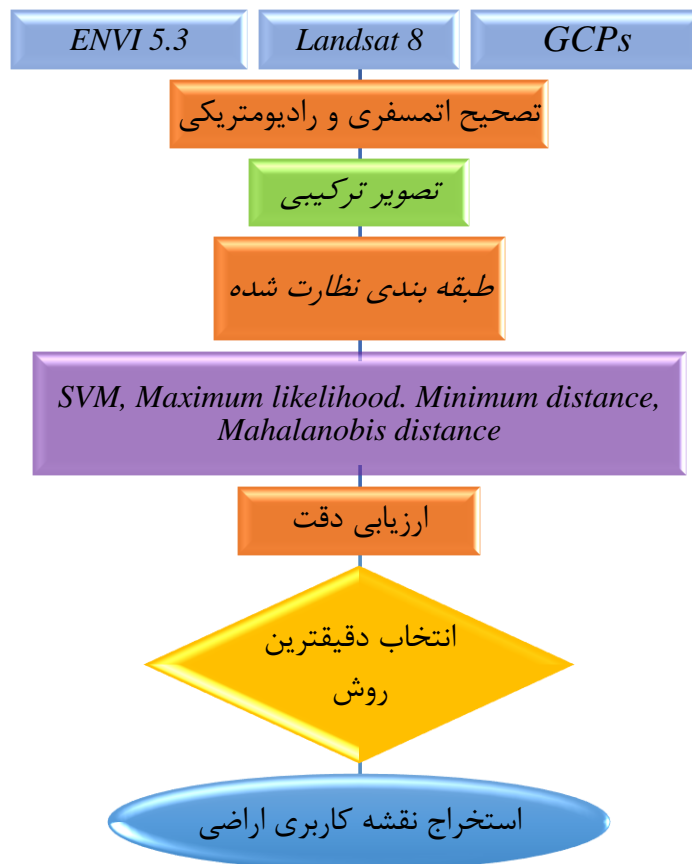
<sup>5</sup> User Accuracy

<sup>6</sup> Commission

<sup>7</sup> Omission



خطای لحاظ شده ( $C_e$ ) که براساس دقت کاربر محاسبه می‌شود، معادل آن درصد از پیکسل‌هایی است که در واقع به کلاس مورد نظر تعلق ندارد، ولی طبقه‌بندی کننده آنها را جزء آن کلاس خاص در نظر گرفته‌است. خطای حذف شده ( $O_e$ )، مربوط به آن درصد از پیکسل‌هایی است که در واقعیت زمینی به کلاس مورد نظر مربوط است، ولی جزء کلاس‌های دیگر طبقه‌بندی شده‌است. نمودار جریانی مراحل تحقیق، در شکل ۳ آورده شده‌است.

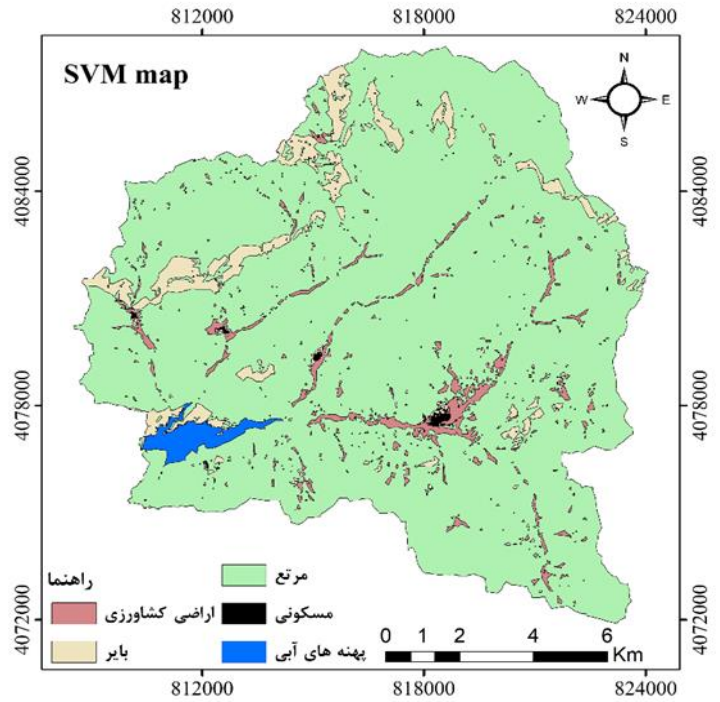


شکل ۳: شماتیک روند تهیه نقشه کاربری اراضی

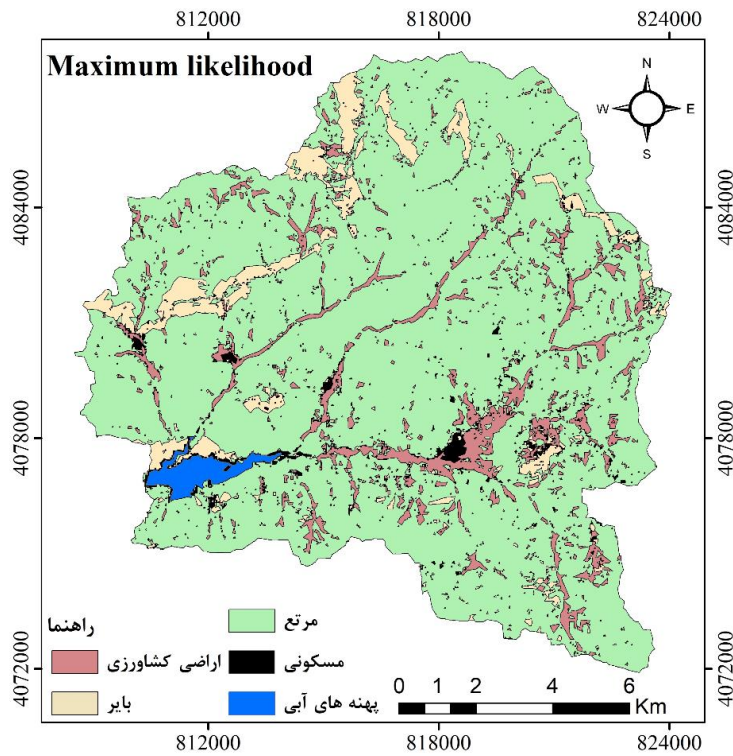
#### ۴- یافته‌ها (نتایج)

استفاده از تصاویر ماهواره‌ای، یکی از منابع مهم مدیریت عرصه‌های منابع طبیعی است. استفاده و کاربرد این تصاویر در زمینه‌های مختلف، می‌تواند قابلیت و محدودیت‌های این داده‌ها را مشخص کند. برای ارزیابی دقت و صحت نقشه‌های طبقه‌بندی شده به روش‌های مختلف، با قطع دادن هر یک از نقشه‌های طبقه‌بندی شده با نقشه واقعیت زمینی و تشکیل ماتریس خطا، از دقت کلی و صحت متوسط استفاده شد. روش‌های مورد استفاده شامل حداکثر احتمال، حداقل فاصله از میانگین، فاصله ماهالونوبی و ماشین بردار پشتیبان است. بر همین اساس، نقشه کاربری اراضی منطقه مورد مطالعه در پنج کلاس تهیه شد که شامل اراضی کشاورزی، مرتع، زمین بایر، مناطق مسکونی و پهنه‌های آبی است. نقشه حاصل از هر یک از این روش‌ها، به تفکیک در شکل‌های (۴-۷) آورده شده‌است. علاوه بر این، مساحت هر یک از کلاس‌های طبقه‌بندی شده کاربری‌ها نیز برآورد شد که نتایج آن در جدول ۱ ذکر شده‌است.

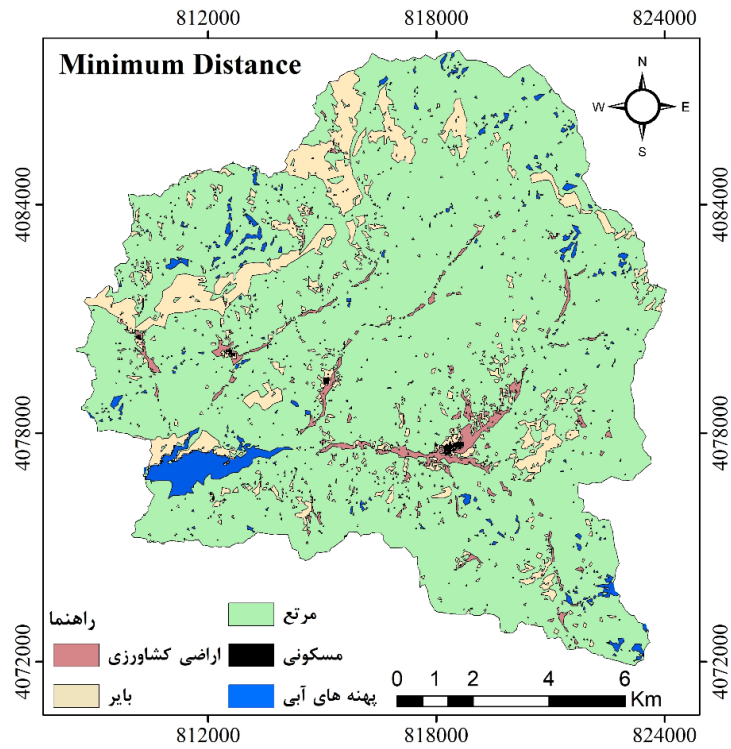




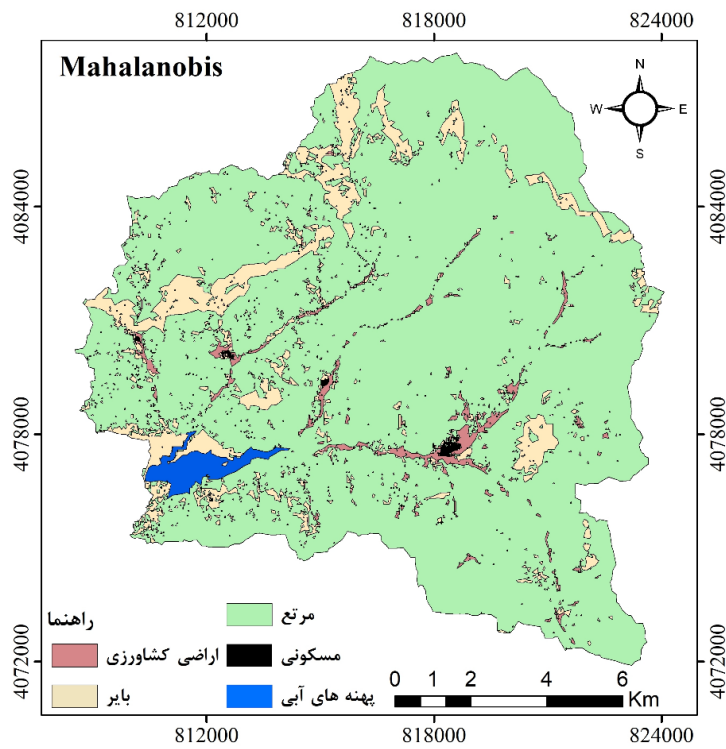
شکل ۴: نقشه طبقه‌بندی شده کاربری اراضی با استفاده از الگوریتم ماشین بردار پشتیبان



شکل ۵: نقشه طبقه‌بندی شده کاربری اراضی با استفاده از الگوریتم حداکثر احتمال



شکل ۶: نقشه طبقه‌بندی شده کاربری اراضی با استفاده از الگوریتم حداقل فاصله از میانگین



شکل ۷: نقشه طبقه‌بندی شده کاربری اراضی با استفاده از الگوریتم فاصله ماحالانویی

جدول ۱: درصد مساحت کاربری‌های استخراج شده با استفاده از الگوریتم‌های طبقه‌بندی نظارت شده

کاربری اراضی	ماشین بردار پشتیبان	فاصلهٔ ماهالانوبی	حداکثر احتمال	حداقل فاصله از میانگین
مسکونی	۰/۱۵	۰/۱۵	۱/۲۲	۰/۱۱
پهنه‌های آبی	۱/۱۹	۱/۲۳	۱/۰۹	۲/۵۹
اراضی کشاورزی	۳/۲۶	۱/۷۸	۷/۹۸	۱/۹۸
مرتع	۸۹/۸۵	۸۷/۰۸	۸۳/۳۴	۸۵/۸۶
بایر	۵/۵۵	۹/۷۶	۶/۳۸	۹/۴۶

براساس نقشه کاربری استخراج شده با استفاده از الگوریتم ماشین بردار پشتیبان، کاربری مرتع و کاربری مسکونی به ترتیب بیشترین (۸۹/۸۵٪) و کمترین (۰/۱۵٪) سطح از حوضه آبخیز سد تهم را به خود اختصاص دادند. در دیگر الگوریتم‌های طبقه‌بندی نیز به ترتیب کاربری مرتع، اراضی بایر، اراضی کشاورزی، پهنه آبی و مناطق مسکونی به ترتیب بیشترین سطح از منطقه را در بر گرفتند. دقت روش‌های طبقه‌بندی نیز با استفاده از تابع ماتریس خطا ۱ در محیط نرم‌افزار ENVI 5.3 ارزیابی شد. برای این کار، دقت کلی<sup>۲</sup> و ضریب کاپا<sup>۳</sup> به تفکیک برای هر یک از روش‌های طبقه‌بندی محاسبه و برآورد شد (جدول ۲).

جدول ۲: محاسبه شاخص‌های ارزیابی دقت برای الگوریتم‌های طبقه‌بندی

الگوریتم	کاربری	دقت تولیدکننده	دقت کاربر	دقت اضافه	دقت حذف	دقت کلی	ضریب کاپا
ماشین بردار پشتیبان	مسکونی	۸۹/۶۶	۱۰۰	۰	۱۰/۳۴	۹۸/۲۴	۰/۹۶۹
	پهنه‌های آبی	۱۰۰	۱۰۰	۰	۰		
	اراضی کشاورزی	۹۸/۹۱	۱۰۰	۰	۱/۰۹		
	مرتع	۱۰۰	۹۷/۷	۲/۳	۰		
حداقل فاصله از میانگین	بایر	۹۳/۱۴	۹۸/۳۳	۱/۶۷	۶/۸۶	۹۴/۳۶	۰/۹۰۴
	مسکونی	۸۱/۰۳	۱۰۰	۰	۱۸/۹۷		
	پهنه‌های آبی	۱۰۰	۷۳/۱۴	۲۶/۸۶	۰		
	اراضی کشاورزی	۹۴/۵۴	۱۰۰	۰	۵/۴۶		
حداکثر احتمال	بایر	۹۷/۶۳	۸۹/۸۹	۱۰/۴۱	۲/۳۷	۹۷/۴۵	۰/۹۶۱
	مسکونی	۹۲/۱۹	۸۲/۸۶	۱۷/۱۴	۰		
	پهنه‌های آبی	۱۰۰	۱۰۰	۰	۷/۸۱		
	اراضی کشاورزی	۹۹/۴۱	۹۶/۳۲	۳/۶۸	۰		
	مرتع	۹۶/۵۷	۹۹/۰۸	۰/۹۲	۰/۵۹		

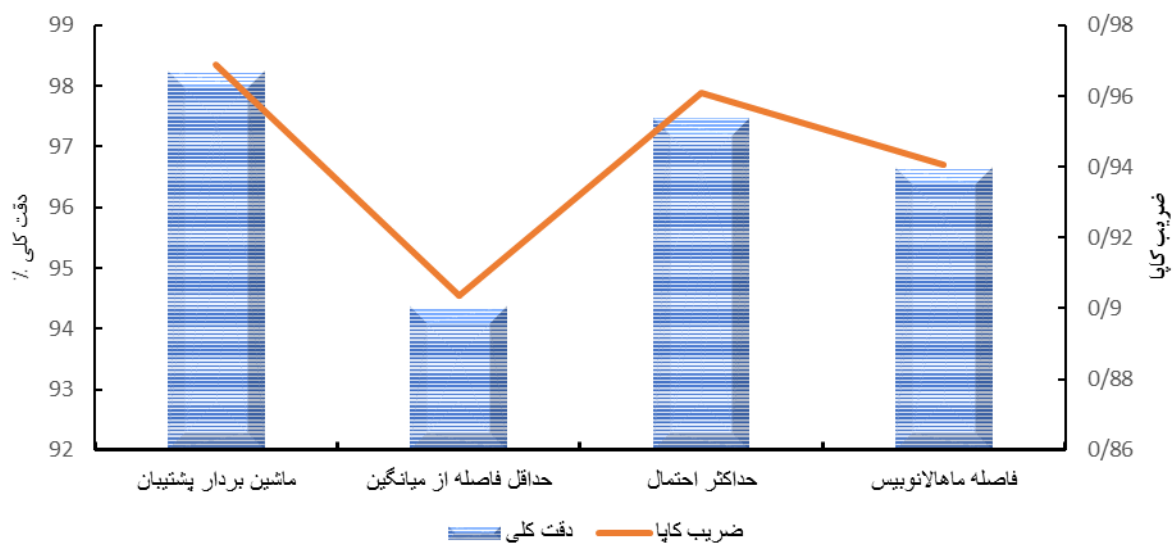
<sup>۱</sup> Confusion matrix

<sup>۲</sup> Overall accuracy

<sup>۳</sup> Kappa coefficient

					۹۲/۱۹	بایر
		۳/۴۳	۰	۱۰۰		مسکونی
		۱۵/۵۲	۰	۱۰۰	۸۴/۴۸	پهنه‌های آبی
		۰	۰	۱۰۰	۱۰۰	اراضی کشاورزی
۰/۹۴	۹۶/۶۴	۴/۳۷	۰	۱۰۰	۹۵/۶۳	فاصله ماهالانویی
		۱/۲۶	۳/۱۴	۹۶/۸۶	۹۸/۷۴	مرتع
		۸/۷۱	۷/۲۴	۹۲/۷۶	۹۱/۲۹	بایر

با توجه به نتایج جدول ۲، الگوریتم ماشین بردار پشتیبان با دقت کلی ۹۸/۲۴ درصد و ضریب کاپا معادل ۰/۹۶۹، در بین دیگر الگوریتم‌ها دقت بالایی داشت. البته الگوریتم حداکثر احتمال نیز با دقت کلی ۹۷/۴۵ درصد و ضریب کاپا معادل ۰/۹۶۱، در استخراج نقشه کاربری اراضی دارای صحت بالایی بود و با الگوریتم ماشین بردار پشتیبان اختلاف ناچیزی داشت؛ به عبارتی، می‌توان گفت دو الگوریتم ماشین بردار پشتیبان و حداکثر احتمال، در استخراج نقشه کاربری اراضی دقت بالایی داشتند. دو الگوریتم فاصله ماهالانویی و حداقل فاصله از میانگین، به ترتیب در رتبه‌های بعدی قرار گرفتند (شکل ۸). علاوه بر این، دو پارامتر صحت تولیدکننده و صحت کاربر نیز به تفکیک هر کلاس کاربری و برای هر یک از روش‌های طبقه‌بندی برآورد شد.



شکل ۸: ارزیابی الگوریتم‌های طبقه‌بندی بر اساس دقت کلی و ضریب کاپا

با توجه به نتایج جدول‌های ۱ و ۲ مشاهده می‌شود که درصد‌های مربوط به پنج کلاس کاربری اراضی (مناطق مسکونی، پهنه‌های آبی، اراضی کشاورزی، مرتع و اراضی بایر) اکثراً دقت بالایی دارد. این موضوع بیانگر این مطلب است که درصد بالایی از پیکسل‌های مربوط به کاربری‌های ذکر شده، درست طبقه‌بندی شده‌است.

## ۵- بحث و نتیجه‌گیری

امروزه تهیه نقشه LULC، مسئله‌ای حیاتی برای جمع‌آوری اطلاعات در زمینه برنامه‌ریزی شهری، مدیریت جامع حوضه‌های آبخیز، شناسایی عوامل مؤثر و نظارت بر محیط است. استفاده از فناوری سنجنش از دور با تصاویر ماهواره‌ای مختلف، بهترین استراتژی توصیه شده برای تهیه نقشه LULC است. با استفاده از سیستم‌های مختلف می‌توان محدودیت و قابلیت این تصاویر را تشخیص داد. هدف از این تحقیق، تولید نقشه‌های موضوعی LULC با مقایسه بین الگوریتم‌های مختلف، با استفاده از حداکثر احتمال، حداقل فاصله از میانگین، فاصله ماه‌الونویی و ماشین بردار پشتیبان در حوضه آبخیز تهم در استان زنجان، با استفاده از تصویر لندست ۸ مرداد ماه ۱۴۰۰ است. پس از اجرای همه اصلاحات لازم با مراحل پیش‌پردازش، پردازش و تحلیل تصاویر ماهواره‌ای، از چهار الگوریتم مختلف برای طبقه‌بندی تصویر ماهواره‌ای استفاده شد.

الگوریتم ماشین بردار پشتیبان با دقت کلی ۹۸/۲۴٪ و ضریب کاپا معادل ۰/۹۶۹، در بین دیگر الگوریتم‌ها دقت بالایی داشتند. البته الگوریتم حداکثر احتمال نیز با دقت کلی ۹۷/۴۵٪ و ضریب کاپا معادل ۰/۹۶۱، در زمینه استخراج نقشه کاربری اراضی دارای صحت بالایی بود و با الگوریتم ماشین بردار پشتیبان اختلاف ناچیزی داشت؛ به عبارتی می‌توان گفت دو الگوریتم ماشین بردار پشتیبان و حداکثر احتمال، در استخراج نقشه کاربری اراضی دقت بالایی داشتند. دو الگوریتم فاصله ماه‌الونویی و حداقل فاصله از میانگین، به ترتیب در رتبه‌های بعدی قرار گرفتند. نتایج این مطالعه با نتایج سایر محققان از جمله (Huang et al, 2002 & Tolessa et al, 2020) - که روش‌های مختلفی را برای طبقه‌بندی نقشه LULC ارزیابی کردند - مطابقت دارد. همچنین مشخص شد که نتایج روش SVM و الگوریتم ML، در تهیه نقشه LULC نسبت به سایر روش‌ها بالاترین دقت ارزیابی را دارد و با نتایج تحقیقات سایر محققان از جمله (Deilmai et al, 2014 & Dibs and Al-Hedny, 2019 & Dibs, Al-Hedny, et al, 2018 & Dibs, Al-) (Janabi, et al, 2018 & Dibs et al, 2017 & Dibs et al, 2016 & Yousef et al, 2011) مطابقت دارد. مقایسه بین الگوریتم‌های طبقه‌بندی نظارت‌شده به منظور استخراج نقشه موضوعی LULC نشان داد که ضریب کاپا و دقت کلی در الگوریتم SVM، نسبت به سایر الگوریتم‌ها بالاتر است و در نتیجه، دقت بیشتری دارد.

براساس نقشه کاربری استخراج شده با استفاده از الگوریتم ماشین بردار پشتیبان، کاربری مرتع و کاربری مسکونی به ترتیب بیشترین (۸۹/۸۵٪) و کمترین (۰/۱۵٪) سطح حوضه آبخیز سد تهم را به خود اختصاص دادند. در دیگر الگوریتم‌های طبقه‌بندی نیز به ترتیب کاربری مرتع، اراضی بایر، اراضی کشاورزی، پهنه‌های آبی و مناطق مسکونی به ترتیب بیشترین سطح از منطقه را دربر گرفتند.

با توجه به نتایج این تحقیق می‌توان چنین نتیجه گرفت که روش SVM و حداکثر احتمال با پایه آماری مستحکمی که دارد، مرز بین کلاس‌ها را بهتر از سایر روش‌های طبقه‌بندی تفکیک می‌کند. همچنین با توجه به قدرت تفکیکی طیفی بالای سنجنده OLI، روش SVM و حداکثر احتمال بهترین نتایج را برای طبقه‌بندی نظارت شده داده‌های سنجنده OLI ارائه می‌دهد. با این روش‌ها می‌توان بهترین نقشه کاربری اراضی را برای منطقه مورد مطالعه و مناطقی با شرایط اقلیمی و اداکیکی مشابه تهیه و تنظیم کرد. علاوه بر این، تا حدود زیادی می‌توان فرسایش منطقه را کاهش داد و

از منابع موجود در منطقه به طور بهینه استفاده کرد که حاصل آن، جلوگیری از کاهش کیفیت آب پشت سد برای مصرف آب شرب شهر زنجان و جلوگیری از ورود رسوبات بیشتر به دریاچه سد و کاهش حجم آن خواهد بود.

## منابع

1. Abbas, Z., & H. S. Jaber., (2020). Accuracy assessment of supervised classification methods for extraction land use maps using remote sensing and GIS techniques. *IOP Conference Series: Materials Science and Engineering*, 745(1). 159-166
2. Alemayehu, F.; Tolera, M.; & G. Tesfaye, 2019. Land use land cover change trend and its drivers in somodo watershed south western, Ethiopia. *African Journal of Agricultural Research*, 14(2), 102-117 .
3. Chang, Y.; Hou, K.; Li, X.; Zhang, Y.; & P. Chen, 2018. Review of land use and land cover change research progress, *IOP Conference Series: Earth and Environmental Science* , 113 (1), p. 012087
4. Davoodpour, R., & H. Toranj Zar., (2021). Determining land use map using ETM + sensor (Case study of Hindudar watershed). *Journal of Environmental Science and Technology*. 22 (6), 379-389.
5. Deilmai, B. R.; Ahmad, B. B.; & H. Zabihi, 2014. Comparison of two classification methods (MLC and SVM) to extract land use and land cover in Johor Malaysia, *IOP conference series: Earth and environmental science* , 20(1), 348-354.
6. Dibs, H., & S. Al-Hedny., (2019). Detection wetland dehydration extent with multi-temporal remotely sensed data using remote sensing analysis and GIS techniques. *International Journal of Civil Engineering and Technology*. 10, 143-154 .
7. Dibs, H.; Al-Hedny, S.; & H. A. Karkoosh, 2018. Extracting Detailed Buildings 3D Model with Using High Resolution Satellite Imagery by Remote Sensing and GIS Analysis; Al-Qasim Green University a Case Study, *International Journal of Civil Engineering and Technology*, 9(7), 1097-1108 .
8. Dibs, H.; Al-Janabi, A.; & C. Gomes, 2018. Easy To Use Remote Sensing and GIS Analysis for Landslide Risk Assessment, *Journal of University of Babylon for Engineering Sciences*, 26(1), 42-54.
9. Dibs, H.; Idrees, M. O.; & G. B. A. Alsalthin, 2017. Hierarchical classification approach for mapping rubber tree growth using per-pixel and object-oriented classifiers with SPOT-5 imagery, *The Egyptian Journal of Remote Sensing and Space Science*, 20(1), 21-30 .
10. Dibs, H.; Idrees, M. O.; Saeidi, V.; & S. Mansor, 2016. Automatic Keypoints Extraction from UAV Image with Refine and Improved Scale Invariant Features Transform (RI-SIFT), *International Journal of Geoinformatics*, 12(3), 51-58.
11. Goldewijk, K. K., & N. Ramankutty., (2004). Land cover change over the last three centuries due to human activities: The availability of new global data sets. *GeoJournal*. 61(4), 335-344 .
12. Haregeweyn, N.; Tesfaye, S.; Tsunekawa, A.; Tsubo, M.; Meshesha, D. T.; Adgo, E.; & A. Elias, 2015. Dynamics of land use and land cover and its effects on hydrologic responses: case study of the Gilgel Tekeze catchment in the highlands of Northern Ethiopia, *Environmental monitoring and assessment*, 187(1), 1-14 .
13. Huang, C.; Davis, L.; & J. Townshend, 2002. An assessment of support vector machines for land cover classification, *International Journal of remote sensing*, 23(4), 725-749 .
14. Jafarzadeh, J., & H. Nazmfar., (2019). Investigating the efficiency of satellite image classification methods in evaluating urban land use changes using scale optimization in object-oriented processing (Case: Ardabil city). *Journal of Urban Research and Planning*. 10(36), 117-128 . (In persian)



15. Khazaei, Z.; Mohammad, M.; Anahita, R.; & A. Aliabad, 2019. Comparing the accuracy of different classification methods in preparing land use map (Case study: Yazd city), *Geographical Exploration of Desert Areas*, 7(1), 165-178 . (In persian)
16. Lambin, E. F.; Turner, B. L.; Geist, H. J.; Agbola, S. B.; Angelsen, A.; Bruce, J. W.; ... & C. Folke, 2001. The causes of land-use and land-cover change: moving beyond the myths, *Global environmental change*, 11(4), 261-269 .
17. Maitima, J. M.; Mugatha, S. M.; Reid, R. S.; Gachimbi, L. N.; Majule, A.; Lyaruu, H.; ... & S. Mugisha, 2009. The linkages between land use change, land degradation and biodiversity across East Africa, *African Journal of Environmental Science and Technology*, 3(10).
18. Mariye, M.; Mariyo, M.; Changming, Y.; Teffera, Z. L.; & B. Weldegebrial, 2020. Effects of land use and land cover change on soil erosion potential in Berhe district: A case study of Legedadi watershed, Ethiopia, *International Journal of River Basin Management*, 1-13 .
19. Moradi, H., & V. Rezaei., (2021). Comparison of Land use Type Classification Algorithms in the Land use Mappreparation Inzenouzchai Watershed. *Degradation and Rehabilitation of Natural Land*. 1(2), 80-88.
20. Shafiee, S.; Alikhah Asl, M.; & M. Rezvani, 2016. Preparation of land cover map of Islamshahr city using maximum probability and fuzzy algorithms of 2015 year, *Man and Environment journal*, 15(2), 27-37 . (In persian)
21. Tolessa, T.; Dechassa, C.; Simane, B.; Alamerew, B.; & M. Kidane, 2020. Land use/land cover dynamics in response to various driving forces in Didessa sub-basin, Ethiopia, *GeoJournal*, 85(3), 747-760 .
22. Watson, R. T.; Noble, I. R.; Bolin, B.; Ravindranath, N.; Verardo, D. J.; & D. J. Dokken, 2000. Land use, land-use change and forestry: a special report of the Intergovernmental Panel on Climate Change. Cambridge University Press .
23. Yousef, S.; Tazeh, M.; Mirzaee, S.; Moradi, H.; & S. Tavangar, 2011. Comparison of different classification algorithms in satellite imagery to produce land use maps (Case study: Noor city), *JOURNAL OF RS AND GIS FOR NATURAL RESOURCES*, 2(2), 15-23. (In persian)



# Assessing the Accuracy of Supervised Classification Algorithms for Land Use Map Extraction (Study Area: Taham Watershed)

Ghobad Rostamizad<sup>1</sup>: Assistant Professor, Soil Conservation and Watershed Management Department, Zanjan Agricultural and Natural Resources Research Center, AREEO, Zanjan, Iran

Zahra Khanabaei: PhD in Geomorphology from the University of Tehran - Teacher of Education of Zanjan Province

Mohammad Tahmoreh: Assistant Professor, Soil Conservation and Watershed Management Department, Zanjan Agricultural and Natural Resources Research Center, AREEO, Zanjan, Iran

Article History (Received: 2022/01/20

Accepted: 2022/07/13)



[20.1001.1.22517812.1401.12.4.7.8](https://doi.org/10.22517/12.4.7.8)

## Extended abstract

### 1- Introduction

Remote sensing and Geographic information system have the capability of detecting and/or monitoring the features of the earth's surface using satellite images and have different radiometric, spectral, spatial, and temporal resolutions. These technologies resolutions have several advantages in order to minimize the time and cost of extracting Land cover and Land use (LULC). In addition, the remote sensing provides important and different kinds of remote data sources to extract LULC information. Remote sensing data are widely used and applied to perform classification of LULC in the world. Remote sensing data has the ability of updating information about all of the features located on the earth's surface. The classification of remote sensing imagery is an important method in order to determine the LULC information. Classification approaches are divided into two categories: pixel base and objected oriented base classifiers. The present study's aims are conducting land use classification by making a comparison between different algorithms in Taham watershed in Zanjan province.

### 2- Methodology

In this research, Landsat8 satellite images (25 July) were used to perform preprocessing and processing steps. Geometric, radiometric and atmospheric corrections were used to eliminate the image noise. In the next step, for the existing uses in the region, first, field samples (160 educational points) were collected separately for each use in the region by using field visits and the Global Positioning System (GPS). Educational samples were divided into two categories, 70% of them were used for classification and 30% were used to accuracy of classification methods. After applying the image corrections according to the studied algorithms for each algorithm using EVVI 5.3 software and educational samples, the land use role for the study area was prepared. According to the sampling, five classes in the study area include agricultural land, rangeland, barren land, residential areas and waterbody areas. To evaluate the accuracy of the classified maps, the error matrix was calculated by using ENVI software. For this purpose, in the present study, the coefficients of Overall Accuracy, Kappa Coefficient, Producer Accuracy, User Accuracy, Commission and Omission had been used.

### 3- Results

The aim of this research was to classify and map land-use/land-cover of the study area using remote sensing and Geospatial Information System (GIS) techniques and includes two sections, namely Land use/Landover (LULC) classification and accuracy assessment. In this study, the supervised classification and four algorithms (including SVM, MLC, MDC and MC) were used and then five classes for the study area including agricultural land (3.26),

<sup>1</sup> Corresponding Author: [Rostamizad60@gmail.com](mailto:Rostamizad60@gmail.com)

rangeland land (89.85), residential land (0.15), waterbody (1.19) and barren lands (5.55) were extracted. Also, the results showed the SVM algorithm with overall accuracy (98.24), kappa coefficient (0.969) as the most accurate, MDC algorithm with overall accuracy (94.36) and kappa coefficient (0.9) as the last priority in this study. According to the results of Tables 2 and 3, it can be seen that the percentages related to the five land use classes (residential areas, irrigated areas, agricultural lands, rangelands and bare lands) are mostly above 90%. This indicates that a high percentage of the pixels associated with these uses are properly classified. This study also presents essential source of information whereby planners and decision makers can use to plan sustainable measures for the environment .

#### 4- Discussion & Conclusions

Today, the use of LULC mapping is a vital issue for collecting data in urban planning, comprehensive watershed management, identification of effective factors, and environmental monitoring .The use of remote sensing technology with different satellite images is recommended as the best strategy for preparing LULC maps .By using different systems the limitations and capabilities of these images can be identified .The purpose of this study was to generate LULC thematic maps by comparing different algorithms, using maximum likelihood, minimum distance from mean, Mahalanobis and Support vector machine in Tham watershed in Zanjan province by applying Landsat 8 images. After performing all of the necessary corrections with the steps of preprocessing, processing and analysis of satellite images, four different algorithms were used to classify the satellite images .Support vector machine algorithm with an accuracy equal to 98.24% and kappa coefficient with an accuracy equal to 0.969 has the highest accuracy among other algorithms .Of course, the maximum likelihood algorithm with 97.45% overall accuracy and kappa coefficient equal to 0.961 have the highest accuracy in extracting land use maps and the slightest difference with the support vector machine algorithm, respectively .In other words, it can be said that the two support vector machine algorithms and the maximum likelihood have a high level of accuracy in extracting the land use maps .The two algorithms of Mahalanobis distance and minimum distance from the mean were ranked next.

**Keywords:** Land use, satellite imagery, overall accuracy, remote sensing, kappa coefficient.