



ارزیابی عملکرد تبدیل آنالیز مؤلفه‌های اصلی، مستقل و کسر حداقل نویز در بهبود صحت استخراج اطلاعات از داده‌های ماهواره سنتینل-۲

صیاد اصغری سراسکانرود، حسن حسنی مقدم، حسین فکرت

دریافت: ۱۲ شهریور ۱۳۹۹ / پذیرش: ۱۲ آبان ۱۳۹۹

دسترسی اینترنتی: ۳۰ تیر ۱۴۰۰

چکیده

پیشینه و هدف استفاده از انواع تبدیلات جهت بهبود صحت استخراج اطلاعات از تصاویر ماهواره‌ای به‌طور فزاینده‌ای رو به افزایش است. در این بین انتخاب تبدیل بهینه اهمیت بالایی داشته و نتایج خروجی را تحت تأثیر قرار خواهد داد. با توجه به ماهیت همبسته تصاویر سنجش‌ازدور، استفاده از انواع تبدیلات جهت بهبود صحت استخراج اطلاعات از این تصاویر امری ضروری است. با توجه به مطالعات انجام‌شده هدف این تحقیق بررسی روش‌های مختلف تبدیلات تصویر در بهبود روند طبقه‌بندی تصاویر ماهواره‌ای و افزایش میزان صحت نقشه‌های کاربری اراضی است. با در نظر گرفتن این نکته که منطقه مورد مطالعه و به‌صورت کلی مناطق شمالی کشور با شرایط خاص درهم تنیدگی کاربری‌ها روبه

رو هستند، لذا به‌کارگیری انواع روش‌های تبدیلات و همچنین روش ترکیبی پیشنهادی در این پژوهش باعث افزایش صحت و دقت اطلاعات خروجی و درنهایت امکان تفکیک و بررسی‌های دقیق‌تر کاربری‌ها و شناسایی عوامل تغییر آن‌ها را جهت برنامه‌ریزی‌های آینده فراهم می‌کند. در این پژوهش به‌منظور ارزیابی عملکرد تبدیلات آنالیز مؤلفه‌های اصلی، آنالیز مؤلفه‌های مستقل و کسر حداقل نویز از تصویر ماهواره سنتینل-۲ شهرستان رضوانشهر استفاده گردید.

مواد و روش‌ها در این پژوهش به‌منظور ارزیابی عملکرد روش‌های آنالیز مؤلفه‌های اصلی، آنالیز مؤلفه‌های مستقل و روش کسر حداقل نویز از تصاویر ماهواره سنتینل-۲ شهرستان رضوانشهر استفاده گردید. از الگوریتم گرام اشمیت (Gram-Schmit) به‌منظور ادغام این داده‌ها با یکدیگر و رسیدن به قدرت تفکیک ۱۰ متر استفاده شد. بعد از اعمال پیش‌پردازش‌های لازم و ادغام تصاویر با یکدیگر هر سه تبدیل روی تصویر اعمال و همچنین ترکیبی از مؤلفه‌های این سه روش تولید گردید. در ادامه نتایج حاصل از تبدیلات با استفاده از الگوریتم حداکثر احتمال در ۸ کلاس کاربری طبقه‌بندی گردید. با استفاده از ضریب شفیلد و محاسبات آماری بین مؤلفه‌های به‌دست‌آمده، ترکیب مؤلفه‌های اول آنالیز مؤلفه‌های اصلی، اول کسر حداقل نویز و مؤلفه دوم آنالیز مؤلفه‌های مستقل، به‌عنوان ترکیب بهینه انتخاب گردید. شناخت کلی از منطقه و مطابق آن تفسیر

صیاد اصغری سراسکانرود^(✉) ۱، حسن حسنی مقدم ۲، حسین فکرت ۳

۱. دانشیار گروه جغرافیای طبیعی، دانشکده ادبیات و علوم انسانی، دانشگاه محقق اردبیلی، اردبیل، ایران

۲. کارشناسی ارشد سنجش‌ازدور و سیستم اطلاعات جغرافیایی، دانشگاه خوارزمی، تهران، ایران

۳. دانشجوی کارشناسی ارشد سنجش‌ازدور و GIS، دانشکده ادبیات و علوم انسانی، دانشگاه محقق اردبیلی، اردبیل، ایران

پست الکترونیکی مسئول مکاتبات: s.asghari@uma.ac.ir

<http://dorl.net/dor/20.1001.1.26767082.1400.12.2.4.7>

مؤلفه‌های مستقل و روش کسر حداقل نویز و انتخاب مؤلفه‌های اول الگوریتم‌های آنالیز مؤلفه‌های اصلی و کسر حداقل نویز و مؤلفه دوم آنالیز مؤلفه‌های مستقل، دقت کلی به ۹۲ درصد و ضریب کاپا ۰/۹۴ افزایش یافت.

نتیجه‌گیری در این تحقیق بعد از ارزیابی عملکرد تبدیل آنالیز مؤلفه‌های اصلی، آنالیز مؤلفه‌های مستقل و روش کسر حداقل نویز، یک روش ترکیب بهینه از مؤلفه‌های این تبدیلات پیشنهاد گردید. نتایج تحقیق نشان داد طبقه‌بندی تصویر اصلی بدون اعمال تبدیلات و با همان نمونه‌های تعلیمی دقت کلی و ضریب کاپای پایینی داشته است. بنابراین لزوم اعمال تبدیلات جهت کاهش همبستگی باندها و مجزا سازی اجزای تصاویر احساس می‌شود. همان‌طور که نتایج تحقیق نشان داد طبقه‌بندی تصویر اصلی بدون اعمال تبدیلات و با همان نمونه‌های تعلیمی دقت کلی و ضریب کاپای پایینی داشته است. نتایج تحقیق بیانگر عملکرد نزدیک این روش‌ها به یکدیگر است که نشان از وجود هر دو نوع توزیع گوسی و غیر گوسی متغیرها دارد. تبدیل کسر حداقل نویز، میزان نویز داده‌ها را به حداقل رسانده و در خروجی نتایج بهتری از دو تبدیل آنالیز مؤلفه‌های مستقل و آنالیز مؤلفه‌های اصلی داشته است. با توجه به اینکه این تبدیلات به‌تنهایی قادر به استخراج تمامی اجزای تصویر نیستند، لذا ترکیبی از مؤلفه‌های این تبدیلات بر اساس ضریب شفیلد جهت مفروض بودن توزیع گوسی و غیر گوسی متغیرها با حداقل نویز ممکن انتخاب گردید.

واژه‌های کلیدی: سنتینل-۲، ضریب شفیلد، تبدیلات، حداکثر احتمال، رضوانشهر

بصری خروجی‌ها و همچنین برداشت ۱۲۰ نقطه زمینی توسط GPS مبنای ارزیابی صحت نقشه‌های خروجی بوده است.

نتایج و بحث بعد از اعمال پیش‌پردازش‌های موردنیاز و ادغام تصاویر با یکدیگر، هرکدام از این الگوریتم‌ها بر روی تصویر اعمال گردید و خروجی هرکدام با استفاده از الگوریتم حداکثر احتمال و در ۸ کلاس کاربری طبقه‌بندی گردیدند. نتایج نقشه‌های خروجی نشان داد که تبدیل آنالیز مؤلفه‌های اصلی با توجه به اینکه برای متغیرها توزیع گوسی در نظر می‌گیرد و سعی بر غیر همبسته کردن مؤلفه‌های استخراج‌شده دارد، در نمونه‌هایی با توزیع غیر گوسی دارای ضعف بوده و عملکرد پایینی از خود نشان می‌دهد. الگوریتم کسر حداقل نویز مشابه الگوریتم آنالیز مؤلفه‌های اصلی عمل می‌کند با این تفاوت که نویزها را بهتر مؤلفه‌بندی می‌کند. این الگوریتم در جداسازی کلاس‌ها خطای کمتری داشته و همین عامل باعث عملکرد بهتر و دقت بالاتر نسبت به دو تبدیل دیگر شده است. در الگوریتم آنالیز مؤلفه‌های مستقل باندهای همبسته تصویر منطقه مورد مطالعه به مؤلفه‌های مستقل تبدیل‌شده و اطلاعات جدیدی از منطقه استخراج کرده است. تفسیر بصری دقت بالای نتیجه طبقه‌بندی را نشان می‌دهد و برای کمی کردن دقت تصویر طبقه‌بندی‌شده از ماتریس خطا (کانفیوژن) استفاده شده است.

نتایج حاصل از ارزیابی دقت کلی و ضریب کاپا نشان داد که طبقه‌بندی تصویر اصلی بدون اعمال تبدیلات و با همان نمونه‌های تعلیمی خروجی با دقت کلی ۷۶ درصد و ضریب کاپای ۰/۷۸ بیشترین خطا را داشته است. همچنین نتایج سایر خروجی‌ها به ترتیب برای طبقه‌بندی حاصل از تبدیل آنالیز مؤلفه‌های اصلی، دقت کلی ۸۰ درصد و ضریب کاپا ۰/۸۳، برای طبقه‌بندی حاصل از تبدیل کسر حداقل نویز، دقت کلی ۸۵ درصد و ضریب کاپا ۰/۸۸ و برای طبقه‌بندی حاصل از تبدیل آنالیز مؤلفه‌های مستقل، دقت کلی برابر با ۷۷ درصد و ضریب کاپا معادل ۰/۸۰ به دست آمد. بعد از انتخاب ترکیب بهینه از مؤلفه‌های روش‌های آنالیز مؤلفه‌های اصلی، آنالیز

مقدمه

سنجش از دور به عنوان یکی از مهم ترین و عمده ترین منابع داده های مکانی و موضوعی قلمداد می شود. این فناوری با توجه به وسعت مناطق مورد مطالعه، اطلاعات طیف های متعدد و مشاهده تقریباً ثابت، می تواند یک جایگزین مناسب برای روش های سنتی باشد (۱۹). روش های مختلفی جهت استخراج اطلاعات از تصاویر ماهواره ای وجود دارد، از جمله این روش ها می توان انواع تبدیلات جهت افزایش میزان صحت استخراج اطلاعات از تصاویر، طبقه بندی، ادغام تصاویر و بهبود قدرت تفکیک مکانی و طیفی تصاویر را نام برد (۱۵). داده های سنجش از دور بخصوص تصاویر با وضوح بالا معمولاً دارای حجم زیاد، اطلاعات فراوان و در بعضی موارد اطلاعات اضافی هستند و این اطلاعات اضافی باعث افزایش زمان پردازش و کاهش دقت فرایند طبقه بندی می شود (۲۷). از جمله روش هایی که تأثیر آن روی دقت طبقه بندی اثبات شده است، الگوریتم های کاهش ابعاد داده است که با استفاده از آن ها ابعاد ذاتی داده ها مشخص می شود (۴).

روش های متعددی برای کاهش تعداد متغیرهای ورودی و افزایش حجم اطلاعات خروجی از تصاویر پیشنهاد شده است که از آن جمله می توان به آنالیز مؤلفه های اصلی (۶)، آنالیز مؤلفه های مستقل (۳۱) و کسر حداقل نویز (۲۹) اشاره کرد. روش آنالیز مؤلفه های اصلی با حذف اثر هم خطی، چارچوبی منطقی را برای مطالعه اثر متغیرهای مؤثر با همبستگی متقابل فراهم می آورد. هدف اصلی این آنالیز فشرده سازی اطلاعات موجود در متغیرهای اصلی است. آنالیز مؤلفه های اصلی تعداد متغیرها را به متغیرهای جدید غیر همبسته به عنوان مؤلفه های اصلی دربرگیرنده واریانس کل مجموعه داده ها کاهش می دهد (۱۸). روش آنالیز مؤلفه های مستقل یک روش گسترش یافته آنالیز مؤلفه های اصلی است که به منظور جداسازی منابع مستقل از ترکیب خطی آن ها به کار می رود. ایده اصلی این روش این است که داده ها به صورت خطی توسط مجموعه از منابع جداگانه مستقل ترکیب شده اند (۳۰).

اساس الگوریتم آنالیز مؤلفه های مستقل، استقلال میان منابع ترکیب شده و غیر گوسی بودن تابع توزیع احتمالی منابع مستقل هست (۲۰). یکی دیگر از روش ها که جهت بهبود استخراج اطلاعات از تصاویر ماهواره ای به کار برده می شود تبدیل کسر حداقل نویز است. این الگوریتم از داده ها یک تبدیل خطی ارائه می دهد به طوری که مقادیر نویز داده ها در تبدیل اولیه کمتر و رفته رفته مقدار نویز افزایش می یابد. مهم ترین ویژگی این الگوریتم برآورد دقیق ماتریس کوواریانس است که یک گام اساسی در فرآیند کاهش ابعاد داده ها و بهبود فرآیند پردازش داده ها محسوب می شود (۸).

لذا این تبدیل به صورت گسترده در پردازش های سنجش از دور به عنوان یک روش کاهش ابعاد داده مورد استفاده قرار می گیرد. مطالعات متعددی هم راستا با موضوع این پژوهش صورت گرفته و با توجه شرایط منطقه، نوع داده و روش های مورد استفاده، نتایج متفاوتی در پی داشته است. از جمله می توان به تحقیقات ذیل اشاره کرد؛ ویدواتی (۲۷)، در تحقیقی به کاربرد آنالیز مؤلفه های اصلی در تصویربرداری دیجیتال و ارتباطات پزشکی پرداخته است. در این تحقیق ابتدا به تشریح الگوریتم آنالیز مؤلفه های اصلی پرداخته و در ادامه نقش آن در پردازش تصاویر پزشکی و فشرده سازی تصویر بررسی شده است. نتایج تحقیق وی نشان دهنده کارایی و قدرت بالای این الگوریتم در بهبود پردازش تصاویر پزشکی است. لو و همکاران (۱۸)، تحقیقی با عنوان مقایسه آنالیز مؤلفه اصلی و تبدیل کسر حداقل نویز به عنوان یک روش پیش پردازش جهت حذف نویز از تصاویر فرا طیفی انجام داده اند. در این تحقیق آن ها به مقایسه این دو روش پرداخته و نقاط ضعف و قوت هر الگوریتم را مورد بررسی قرار داده اند. نتایج کار آن ها بیانگر برتری الگوریتم تبدیل کسر حداقل نویز در بهبود نسبت سیگنال به نویز در تصاویر فرا طیفی بوده در حالی که روش آنالیز مؤلفه اصلی در حذف نویزهای با توزیع گوسی عملکرد بهتری داشته است.

برزگری و همکاران (۵) با استفاده از روش های مختلف بارزسازی تصاویر ماهواره ای اقدام به جداسازی واحدهای

به صورت کلی مناطق شمالی کشور با شرایط خاص در هم تنیدگی کاربری ها روبه رو هستند، لذا به کارگیری انواع روش های تبدیلات و همچنین روش ترکیبی پیشنهادی در این پژوهش باعث افزایش صحت و دقت اطلاعات خروجی و در نهایت امکان تفکیک و بررسی های دقیق تر کاربری ها و شناسایی عوامل تغییر آن ها را جهت برنامه ریزی های آینده فراهم می کند.

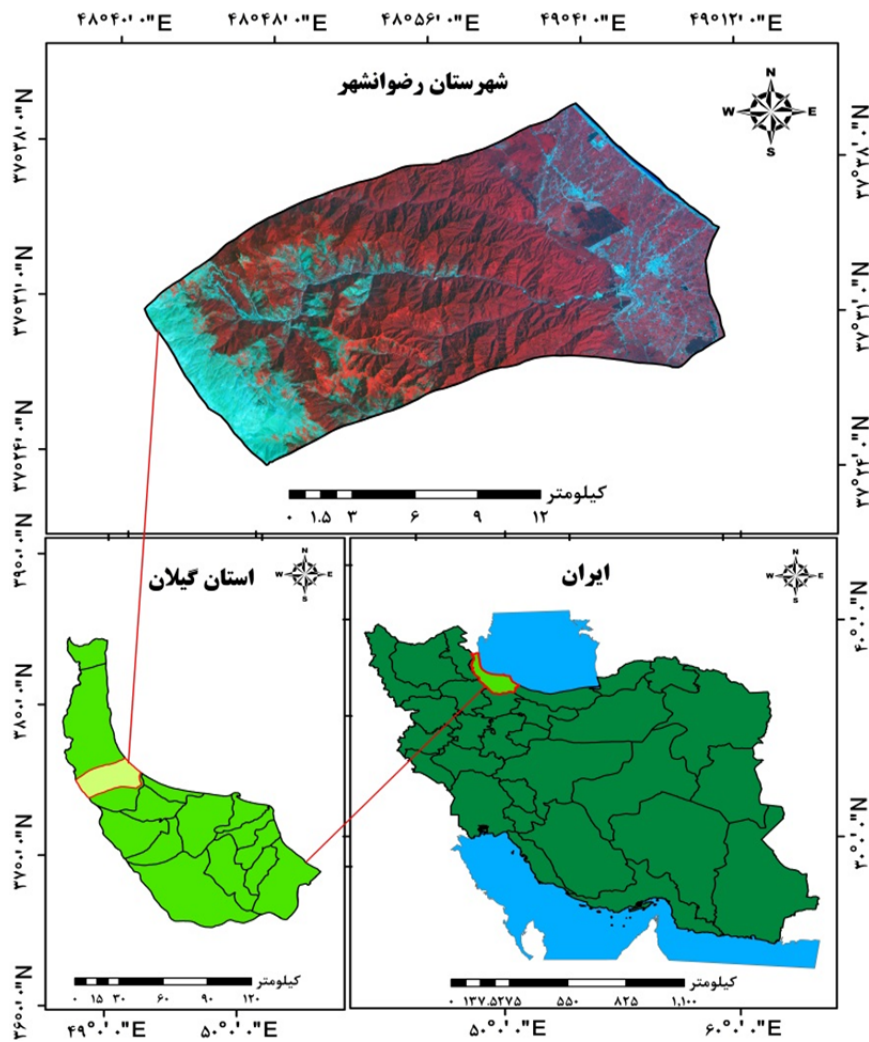
روش تحقیق

منطقه مورد مطالعه

شهرستان رضوانشهر در شمال غرب استان گیلان واقع گردیده و از لحاظ موقعیت جغرافیایی بین عرض های $25^{\circ} 22'$ تا $37^{\circ} 34' 40''$ شمالی و طول جغرافیایی $48^{\circ} 40' 21''$ تا $49^{\circ} 13' 07''$ شرقی قرار گرفته است (شکل ۱). این شهرستان از نظر جهات جغرافیایی از طرف شمال به دریای خزر و شهرستان تالش، از سمت شرق به دریای خزر، شهرستان های بندر انزلی و صومعه سرا و از سوی جنوب و جنوب شرقی به شهرستان ماسال و از طرف غرب به شهرستان خلخال در استان اردبیل محدود می گردد.

زمین شناسی از همدیگر کردند. نتایج پژوهش آن ها نشان داد که روش کسر حداقل نویز نسبت به دو روش آنالیز مؤلفه های اصلی و ترکیب رنگ کاذب (FCC) دقت بیشتری در تفکیک واحدهای زمین شناسی داشته است. علیخاه اصل و همکاران (۲) به ارزیابی روش های مختلف بهبود استخراج اطلاعات از داده های سنجنش ازدور پرداخته اند. آن ها در این تحقیق روش های آنالیز مؤلفه مستقل، تسلدکپ، شاخص نرمال شده تفاضل پوشش گیاهی و طبقه بندی نظارت شده را با یکدیگر مقایسه کردند. نتایج کار آن ها نشان داد که استفاده از الگوریتم آنالیز مؤلفه مستقل موجب افزایش صحت نقشه های خروجی می شود. آل احمدی و آل هامس (۱) به ارزیابی قابلیت الگوریتم های طبقه بندی نظارت نشده و نظارت شده (حداکثر احتمال، حداقل فاصله از میانگین و فاصله ماهالانویی) در تهیه نقشه کاربری اراضی پرداخته اند. نتایج پژوهش آن ها نشان داد که روش حداکثر احتمال از بیشترین دقت نسبت به سایر روش ها برخوردار بوده است.

با توجه به مطالعات انجام شده هدف این تحقیق بررسی روش های مختلف تبدیلات تصویر در بهبود روند طبقه بندی تصاویر ماهواره ای و افزایش میزان صحت نقشه های کاربری اراضی است. با در نظر گرفتن این نکته که منطقه مورد مطالعه و



شکل ۱. موقعیت منطقه مورد مطالعه به صورت ترکیب رنگ کاذب (باندهای ۵، ۴ و ۳)

Fig. 1. Location of the study area area as a false color composite (bands 5, 4 and 3)

سنتینل-۲ به تاریخ ۲۰۱۷/۰۶/۳۰ استفاده گردید (جدول ۱).

روند کلی مراحل انجام پژوهش نیز در شکل ۲ ارائه شده است.

داده‌های مورد استفاده

به منظور ارزیابی عملکرد تبدیل آنالیز مؤلفه‌های اصلی،

آنالیز مؤلفه‌های مستقل و کسر حداقل نویز از تصویر ماهواره

جدول ۱. مشخصات باندهای ماهواره سنتینل-۲ (۹)

Table 1. Specifications of Sentinel-2 satellite bands

شماره باند	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲	۱۳
طول موج (میکرومتر)	۰/۴۳	۰/۴۹	۰/۵۶	۰/۶۵	۰/۷۰	۰/۷۴	۰/۷۸	۰/۸۴	۰/۸۶	۰/۹۴	۱/۳۷	۱/۶۱	۲/۱۹
رزولوشن مکانی (متر)	۶۰	۱۰	۱۰	۱۰	۲۰	۲۰	۲۰	۱۰	۲۰	۶۰	۶۰	۲۰	۲۰



شکل ۲. نمودار روند پژوهش

Fig 2. Flow chart of research methodology

پیش پردازش داده‌ها

تصاویر ماهواره سنجنیل-۲ از ابتدا به لحاظ هندسی تصحیح شده هستند، باین وجود تصویر از لحاظ دقت هندسی به صورت بصری مورد بررسی قرار گرفت. به منظور اعمال تصحیحات اتمسفری از ابزار تصحیح جوی سنجنیل ۲ (Sen2Cor) در محیط نرم افزار سنپ (SNAP) استفاده گردید. این پردازنده منحصراً برای تصاویر ماهواره سنجنیل ۲ سطح A طراحی شده است و با استفاده از یک مدل انتقال تابشی، تصحیح جوی و زمینی را انجام می دهد. روش کار آن بر اساس رویکرد تراکم پوشش گیاهی تیره است و نسبت بین بازتاب پایین جو در طول موج های مختلف در آن ثابت است. این الگوریتم برای مطابقت با پوشش گیاهی تیره به برخی از پیکسل های تصویر نیاز دارد و پس از مشخص شدن وجود چنین پیکسل هایی، الگوریتم به طور خودکار این پیکسل ها را انتخاب می کند و تصویر را تصحیح می کند (۱۰).

با توجه به اینکه باندهای تصاویر ماهواره سنجنیل-۲ دارای قدرت تفکیک مکانی متفاوتی هستند، از الگوریتم گرام

اشمیت (Gram-Schmit) به منظور ادغام این داده ها با یکدیگر و رسیدن به قدرت تفکیک ۱۰ متر استفاده گردید. الگوریتم گرام اشمیت، یک الگوریتم آماری است که با استفاده از کوواریانس باندها، آن ها را غیر همبسته می کند. روش انجام این الگوریتم به این صورت است که ابتدا یک باند پانکروماتیک با استفاده از ترکیب n باند از داده های چند طیفی ورودی شبیه سازی می شود. در ادامه، این تبدیل بر روی باندهای تصویر چند طیفی و باند پانکروماتیک شبیه سازی شده ارائه می گردد. سپس باند پانکروماتیک جایگزین اولین باند گرام اشمیت شده و معکوس تبدیل اعمال می شود (۲۱). در نهایت تصویر ادغام شده به صورت یک ماتریس خروجی برآورد می گردد.

استخراج اطلاعات از تصویر

جهت بهبود صحت استخراج اطلاعات از روش های آنالیز مؤلفه های اصلی، آنالیز مؤلفه های مستقل و کسر حداقل نویز و

استفاده نشده و تأکید بر روی قابلیت استخراج عوارض بوده است. این تبدیل با استفاده از رابطه ۲ انجام شد.

$$X_1 X_2, \dots, X_m \Rightarrow S_1 S_2, \dots, S_d \rightarrow d \leq m \quad [2]$$

در این رابطه؛ X ماتریس داده‌ها و A ماتریس ترکیب ناشناخته و S ماتریس مؤلفه‌های مستقل است. در این تبدیل متغیرهای اندازه‌گیری شده میانگین صفر دارند (۲۵).

روش کسر حداقل نویز (MNF)

معمولاً باندهای موجود در داده‌های فرا طیفی دارای نویزهای متنوعی می‌باشند، لذا یک تبدیل استاندارد مؤلفه‌های اساسی نمی‌تواند مؤلفه‌هایی تولید کند که مقادیر نویز به صورت افزایشی رتبه‌بندی شده باشند و این امر باعث می‌شود که در انتخاب یک حد آستانه مطمئن دچار مشکل شویم. برای به دست آوردن چنین مجموعه‌ای از مؤلفه‌ها که به صورت افزایشی در مقادیر نویز (کاهش در نسبت سیگنال به نویز) رتبه‌بندی شده باشند، تبدیل کسر حداقل نویز توسط گرین و همکاران (۱۲) در سال ۱۹۹۸ معرفی گردید. این الگوریتم از داده‌ها یک تبدیل خطی ارائه می‌دهد و فضای داده‌ها را به دو بخش تقسیم می‌کند: بخش اول که شامل مقادیر زیاد و منسجم هستند و بخش دوم شامل مقادیر غیر منسجم و نویز است. این تبدیل مقدار کمینه جهات متعامد کسر نویز (یا بیشینه نسبت سیگنال به نویز) را پیدا می‌کند (۲۸). الگوریتم کسر حداقل نویز دارای ویژگی‌های کاهش ابعاد داده و تشخیص عوارض است که در این مطالعه ویژگی دوم مدنظر بوده است. جهت مؤلفه‌بندی بهتر نویزها روی تصویر منطقه مورد مطالعه، الگوریتم اعمال و مقادیر نویز داده‌ها در خروجی‌های اولیه کمتر و رفته‌رفته مقدار نویز افزایش یافت و در نهایت مؤلفه اول که کمترین نویز را داشت انتخاب و مورد استفاده قرار گرفت. این الگوریتم با استفاده از رابطه‌های ۳ و ۴ انجام گردید.

$$X = (X_1, \dots, X_n)(x_1, \dots, x_d) \quad [3]$$

همچنین ترکیبی از مؤلفه‌های بهینه این سه روش استفاده گردید که به شرح زیر است؛

روش آنالیز مؤلفه‌های اصلی (PCA)

در آنالیز مؤلفه‌های اصلی، داده‌ها از فضایی با ابعاد بالا به فضایی با ابعاد پایین تبدیل می‌شود. اساس این روش بر این پایه استوار است که یک تبدیل خطی Φ به گونه‌ای ایجاد می‌شود که بردار ویژگی‌های h - بعدی را به بردارهای ویژگی d - بعدی (که $h < d$)، تبدیل می‌کند و هدف از اعمال این تبدیل دستیابی به کمترین مقدار خطای مربعات است (۱۱). بعد از اعمال روش آنالیز مؤلفه‌های اصلی روی تصویر، مقادیر بهینه که نشان‌دهنده وجود بیشترین اطلاعات در هر باند بود، روی باندهای ۱ و ۲ تراکم گردید و اطلاعات اضافی حذف و تعداد باندها کمتر شدند. درحالی‌که قبل از اعمال الگوریتم، پراکندگی اطلاعات در همه‌ی باندها دیده می‌شد. در مقایسه‌ی باندهای ۱ و ۲ نیز باند ۱ دارای بیشترین اطلاعات ممکن بود که در روند کار مورد استفاده قرار گرفت. این الگوریتم با استفاده از رابطه ۱ انجام گردید.

$$T_D = \arg \min_T \{E[Px - \hat{x} P]^2\} \quad [1]$$

در این رابطه؛ T تبدیل خطی که کمترین خطای مربعات را نشان می‌دهد، X ماتریس کوواریانس داده‌هایی با میانگین صفر و P معرف مؤلفه‌ها است.

روش آنالیز مؤلفه‌های مستقل (ICA)

در الگوریتم آنالیز مؤلفه‌های مستقل فرض بر این است که تعداد m متغیر اندازه‌گیری شده می‌تواند به صورت ترکیب خطی از d متغیر ناشناخته مستقل نمایش داده شود (۱۴). با اعمال این تبدیل، مؤلفه‌های باندهای همبسته تصویر منطقه مورد مطالعه به مؤلفه‌های مستقل تبدیل شده و اطلاعات جدیدی از منطقه استخراج گردید. اگرچه کاهش ابعاد داده نیز از ویژگی‌های این تبدیل است، اما در این مطالعه از این ویژگی

کلاس است (۱۳). جهت اعمال طبقه بندی ابتدا ۸ کلاس اراضی جنگلی، اراضی مرتعی غنی، اراضی کشاورزی، اراضی مرتعی فقیر، اراضی مشجر، مناطق مسکونی، اراضی بایر و مناطق آبی تعیین و نمونه های تعلیمی با توجه به شناخت کلی از منطقه، تفسیر بصری و اعمال تبدیلات از کاربری های مختلف اخذ گردید و در نهایت خروجی حاصل از تبدیلات با الگوریتم حداکثر احتمال که با استفاده از رابطه های ۶ و ۷ طبقه بندی گردید.

$$x \in w_i \text{ if } p(w_i|x) > p(w_j|x) \text{ for all } j \neq i \quad [6]$$

در رابطه ۶؛ x بردار مقادیر طیفی پیکسل مورد نظر، w_i کلاس طیفی i و $P(w_i|x)$ احتمال تعلق پیکسل x به کلاس w_i است. برای محاسبه احتمال $P(w_i|x)$ که یک احتمال ثانویه است از قانون بیز کمک گرفته می شود. بدین منظور ابتدا احتمال عکس $P(x|w_i)$ با فرض اینکه توزیع احتمال کلاس ها به صورت توزیع نرمال چندبعدی است در نظر گرفته می شود و از رابطه ۷ محاسبه گردید.

$$P(x|w) = (2\pi)^{-f/2} |\Sigma_i|^{-1/2} \exp\left\{-\frac{1}{2}(x - m_i)^t \Sigma_i^{-1} (x - m_i)\right\} \quad [7]$$

در رابطه ۷؛ m_i بردار میانگین، Σ_i ماتریس کوواریانس داده های کلاس w_i و f تعداد باندها است. سپس احتمال اولیه درصد کلی حضور کلاس طیفی w_i در تصویر $P(w_i)$ با توجه به اطلاعات کلی موجود در رابطه با منطقه محاسبه می شود. جهت ارزیابی تفکیک پذیری نمونه های تعلیمی از آزمون جفریس ماتوسیتا استفاده شده است (۲۳) که رابطه شاخص آن به شرح زیر است؛

$$J_{ij} = 2\{1 - e^{-a}\} \quad [8]$$

$$a = \frac{1}{8} (m_i m_j)^t \left(\frac{\Sigma_i + \Sigma_j}{2}\right)^{-1} (m_i - m_j) + \frac{1}{2} \ln \left(\frac{|\Sigma_i + \Sigma_j|/2}{|\Sigma_i|^{1/2} |\Sigma_j|^{1/2}}\right) \quad [9]$$

در این رابطه؛ X نشان دهنده ماتریس $n \times d$ داده ها، X_i نشان دهنده i امین اسپکتروم به صورت یک بردار $d \times 1$ و X_j بردار n مشاهده در j امین باند است. لذا هر اسپکتروم را می توان به صورت جمع سیگنال و نویز با استفاده از رابطه ۴ به دست آورد.

$$X_i = S_i + \varepsilon_i \quad [4]$$

در این رابطه؛ S_i مقدار کل و ε_i نویز را نشان می دهد. به منظور استفاده توأمان از قابلیت روش های آنالیز مؤلفه های اصلی، آنالیز مؤلفه های مستقل و کسر حداقل نویز استفاده ترکیبی از مؤلفه های این روش ها پیشنهاد گردید. در این روش ابتدا باندهای ۱۰ متری ماهواره سنتینل-۲ با یکدیگر ترکیب شده و یک باند ۱۰ متری از مجموع آن ها تولید می شود که حاوی اطلاعات همه این باندها است. سپس این باند با سایر باندها ادغام می شود. هر یک از روش های تبدیل بالا بر روی تصویر اعمال شده و سپس مؤلفه های برتر هر تبدیل که دارای واریانس بیشتر و نویز کمتری هستند با یکدیگر ترکیب می شوند. با استفاده از رابطه ۵ (۲۳) و محاسبات آماری بین مؤلفه های به دست آمده، ترکیب مؤلفه های اول آنالیز مؤلفه های اصلی، دوم آنالیز مؤلفه های مستقل و اول روش کسر حداقل نویز، به عنوان ترکیب بهینه انتخاب گردید.

$$\text{Sheffield}_{\text{index}} = |\text{cov}_{p*p}| \quad [5]$$

در این رابطه؛ P بیانگر ماتریس کوواریانس در ابعاد P است.

الگوریتم حداکثر احتمال

این الگوریتم به صورت یک تابع غیرخطی عمل می کند و در آن میانگین ماتریس بردار و کوواریانس پیکسل های ناشناخته در هر کلاس محاسبه می شود. سپس احتمال تعلق هر پیکسل به کلاس خود مورد تجزیه و تحلیل قرار می گیرد و در آخر با حداکثر احتمال پیش بینی می شود پیکسل متعلق به کدام

و آب)، طبقه‌بندی گردیدند. انتخاب کلاس‌ها به‌منظور طبقه‌بندی بر اساس معیار اندرسون انجام گرفته است (۳). بر اساس این معیار و با توجه به کاربری‌های موجود در منطقه، دو نوع مرتع انتخاب شد. نوع اول مرتع غنی بود که کاملاً سطح زمین پوشیده از پوشش گیاهی بوده و خاک لخت در آن مشاهده نمی‌شود و دیگری مرتع فقیر بود که با وجود داشتن پوشش گیاهی تنک، در مناطقی در تداخل با خاک لخت قرار داشت. جداسازی اراضی جنگلی و اراضی مشجر بر اساس انبوهی و یکنواختی درختان در مناطق جنگلی بوده است. از سویی دیگر مناطق جنگلی کاملاً یکدست بوده و کاربری‌های با شباهت بالا در آن وجود ندارد. درحالی‌که عمده درختان اراضی مشجر در داخل شهر و در اختلاط با کاربری مسکونی بوده است. کلاس اراضی بایر با توجه به وجود قسمت‌هایی از زمین منطقه که به‌صورت کاملاً لخت و فاقد پوشش گیاهی بوده و مساحت بالایی هم دارد انتخاب گردید. اراضی کشاورزی نیز که عمده کاربری منطقه را شامل می‌شود بیشتر شامل شالیزارهای برنج بوده و با توجه به پهنه وسیع و اهمیت بالای این مناطق به‌عنوان یک کاربری مجزا در نظر گرفته شده و طبقه‌بندی گردید. مناطق مسکونی که عمدتاً با کاربری‌های اراضی مشجر و کشاورزی درهم‌تنیده شده است نیز به‌عنوان کاربری مستقل در نظر گرفته شده است تا این کاربری از کاربری‌های مجاور کاملاً جدا گردد و صحت نتایج خروجی افزایش یابد. همچنین کلاس آب که عمدتاً به‌صورت رودخانه و آب‌بند در منطقه موجود است، به‌عنوان یک کاربری در نظر گرفته شد تا مورد طبقه‌بندی قرار گیرد. درنهایت ترکیبی بهینه از مؤلفه‌های این تبدیل‌ها نیز انتخاب و طبقه‌بندی گردید تا مورد ارزیابی قرار گیرد. شکل‌های ۳ تا ۷ نتایج حاصل از طبقه‌بندی بر روی داده‌ها را نشان می‌دهد.

در این رابطه‌ها؛ m_j و m_i به ترتیب بردارهای میانگین برای کلاس‌های j و i بوده و $\sum i$ و $\sum j$ ماتریس‌های کوواریانس برای کلاس i و j هستند.

صحت سنجی

جهت ارزیابی صحت نقشه‌های خروجی از پارامترهای آماری دقت کلی و ضریب کاپا استفاده شد که با استفاده از رابطه‌های ۱۰ و ۱۱ محاسبه گردید (۲۴). برای ارزیابی دقت نقشه‌های طبقه‌بندی شده نیز از ۱۲۰ نقطه زمینی که توسط GPS از کلاس‌های مختلف اخذ گردید، استفاده شد.

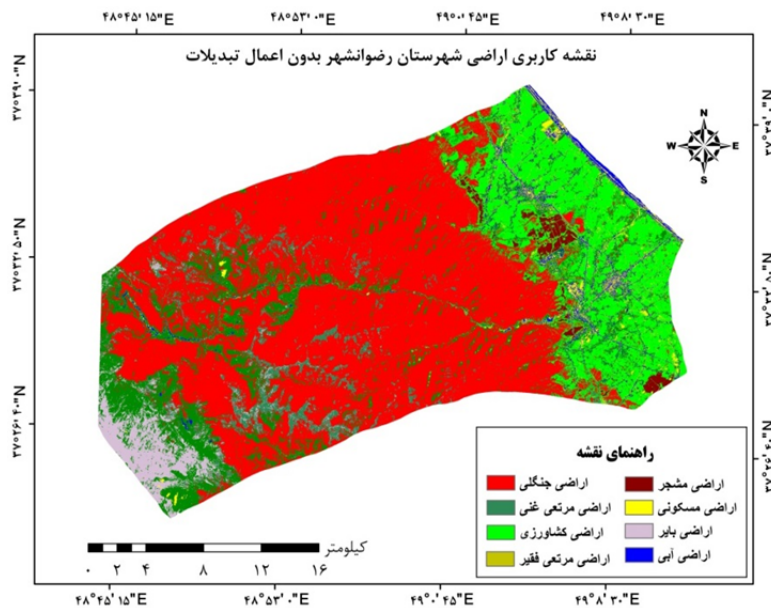
$$OA = \frac{1}{N} \sum P_{ii} \quad [10]$$

$$Kappa = \frac{N \sum_{i=1}^k x_{ii} - \sum_{i=1}^k x_i \times x_i}{N^2 - \sum_{i=1}^k (x_i \times x_i)} \quad [11]$$

در رابطه ۱۰؛ OA دقت کلی، N تعداد پیکسل‌های آزمایشی و P_{ii} عناصر قطر اصلی ماتریس خطا و در رابطه ۱۱؛ N تعداد کل پیکسل‌های تصویر، X_{ii} عناصر قطر اصلی ماتریس، X_i عناصر سطر i ام و $X+i$ عناصر ستون i ام را نشان می‌دهد.

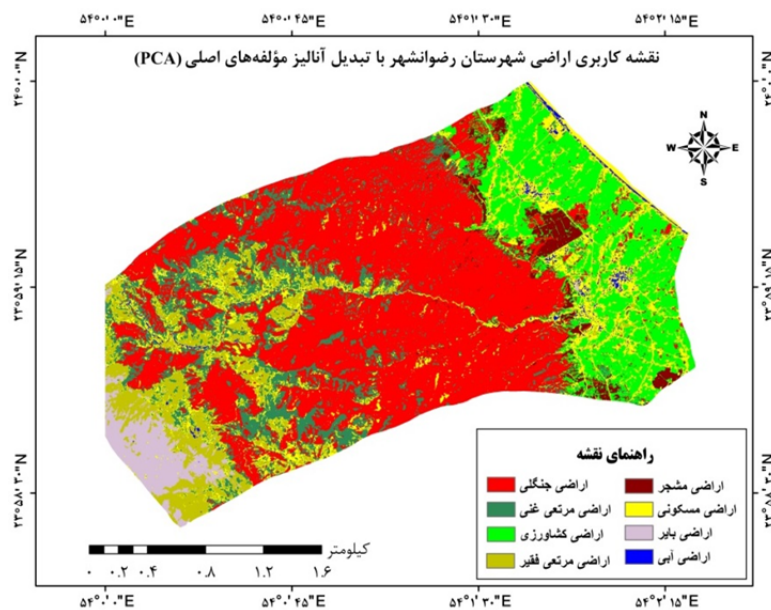
نتایج

به‌منظور ارزیابی عملکرد تبدیل آنالیز مؤلفه‌های اصلی، آنالیز مؤلفه‌های مستقل و کسر حداقل نویز از تصویر ماهواره سنتینل-۲ شهرستان رضوانشهر استفاده گردید. بعد از اعمال پیش‌پردازش‌های لازم و ادغام تصاویر با یکدیگر، هرکدام از این الگوریتم‌ها بر روی تصویر اعمال گردید و خروجی هرکدام با استفاده از الگوریتم حداکثر احتمال و در ۸ کلاس کاربری (اراضی جنگلی، اراضی مرتعی غنی، اراضی کشاورزی، اراضی مرتعی فقیر، اراضی مشجر، اراضی مسکونی، اراضی بایر



شکل ۳. نتایج حاصل از طبقه بندی تصویر اصلی بدون اعمال تبدیلات با ۸ کلاس کاربری

Fig. 3. Results of original image classification without transformation with 8 classes

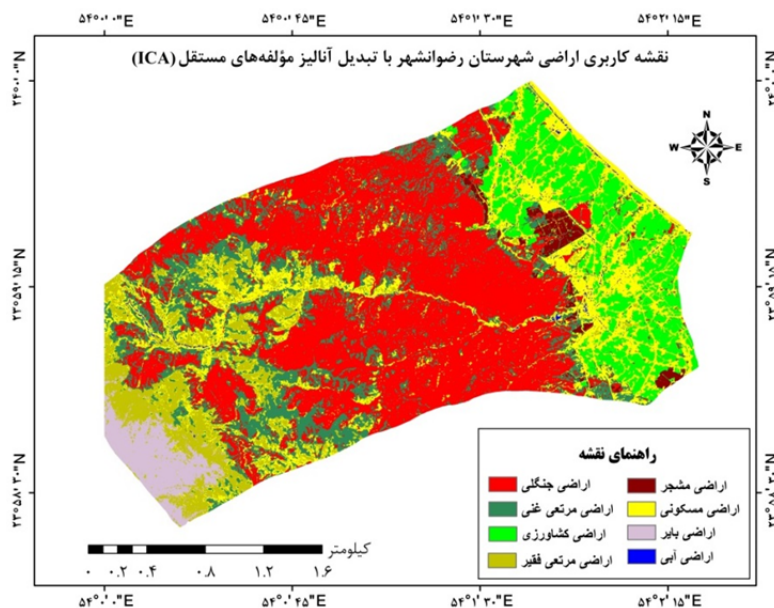


شکل ۴. نتایج حاصل از طبقه بندی خروجی تبدیل آنالیز مؤلفه های اصلی با ۸ کلاس کاربری

Fig. 4. Results of PCA transformation output classification with 8 classes

یکدیگر مشکل داشته و تعداد ۸۷ پیکسل بین این دو کاربری با اختلاط روبه رو بوده است. همچنین کلاس جنگل با مرتع فقیر ۱۴ پیکسل و با مرتع غنی ۱۵ پیکسل جابه جایی پیکسلی داشته که این موارد منجر به پایین آمدن صحت خروجی آن شده است (شکل ۵).

تبدیل آنالیز مؤلفه های اصلی با توجه به اینکه برای متغیرها توزیع گوسی در نظر می گیرد و سعی بر غیر همبسته کردن مؤلفه های استخراج شده دارد، در نمونه هایی با توزیع غیر گوسی دارای ضعف بوده و عملکرد پایینی از خود نشان می دهد. این روش در جداسازی دو نوع مرتع غنی و فقیر از

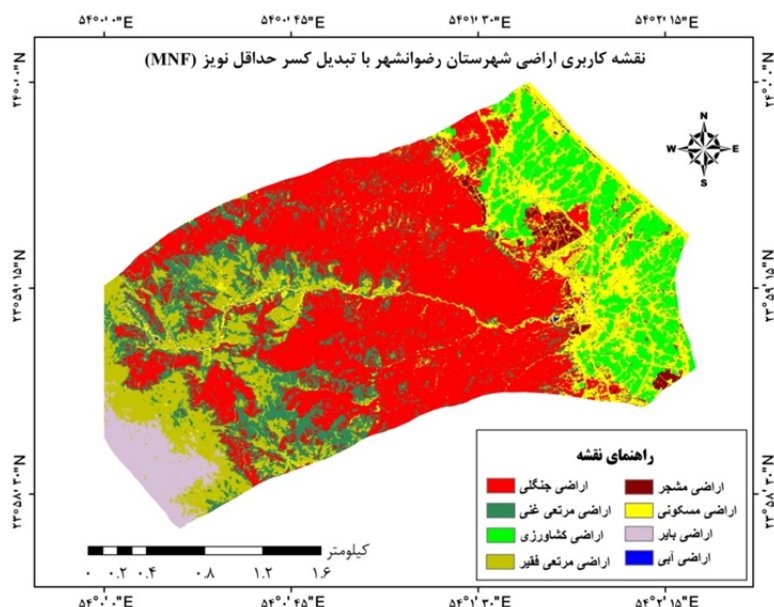


شکل ۵. نتایج حاصل از طبقه‌بندی خروجی تبدیل آنالیز مؤلفه‌های مستقل با ۸ کلاس کاربری

Fig 5. Results of ICA transformation output classification with 8 classes

۵۰ پیکسلی دو کلاس مرتع فقیر و آب داشته است که تا حدی دقت کمتر در تفکیک این دو کلاس را می‌توان به فاصله بیشتر این داده‌ها از توزیع غیرگوسی ارتباط داد (شکل ۶).

تفسیر بصری دقت بالای نتیجه طبقه‌بندی را نشان می‌دهد و برای کمی کردن دقت تصویر طبقه‌بندی شده از ماتریس خطا (کانفیوژن) استفاده شده است (۲۲). نتایج نشان از جابه‌جایی

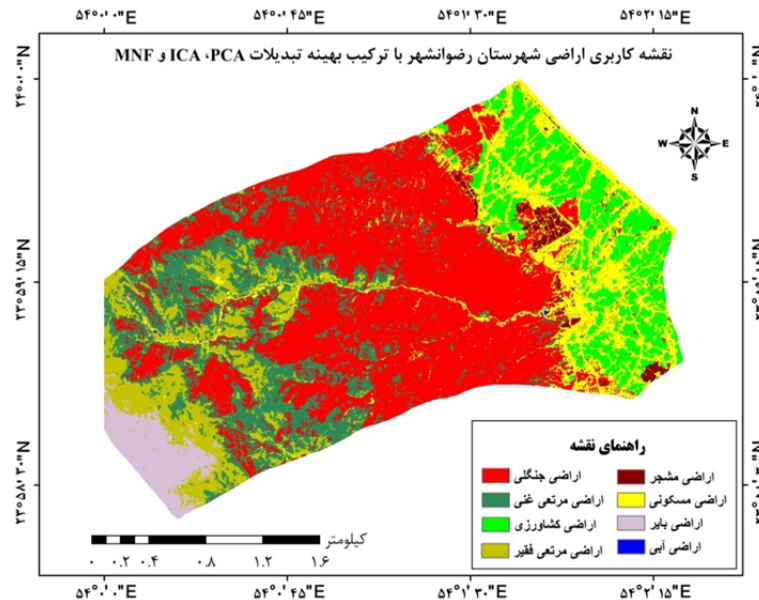


شکل ۶. نتایج حاصل از طبقه‌بندی خروجی تبدیل کسر حداقل نویز با ۸ کلاس کاربری

Fig 6. Results of MNF transformation output classification with 8 classes

کلاس‌ها خطای کمتری داشته و همین عامل باعث عملکرد بهتر و دقت بالاتر نسبت به دو تبدیل آنالیز مؤلفه‌های اصلی و آنالیز مؤلفه‌های مستقل شده است (شکل ۷).

الگوریتم کسر حداقل نویز مشابه الگوریتم آنالیز مؤلفه‌های اصلی عمل می‌کند با این تفاوت که نویزها را بهتر مؤلفه‌بندی می‌کند. مؤلفه‌ها به گونه‌ای مرتب می‌شوند که معمولاً مؤلفه اول کمترین نویز را دارد. این الگوریتم در جداسازی



شکل ۷. نتایج حاصل از طبقه‌بندی ترکیب بهینه مؤلفه‌های آنالیز مؤلفه‌های اصلی، مستقل و کسر حداقل نویز با ۸ کلاس کاربری

Fig 7. Results of classifying the optimal combination of PCA, ICA and MNF components with 8 classes

تفسیر بصری خروجی‌ها و همچنین برداشت ۱۲۰ نقطه زمینی توسط GPS مبنای ارزیابی صحت نقشه‌های خروجی بوده است.

با استفاده از آزمون جفریس ماتوسیتا، تفکیک‌پذیری نمونه‌های تعلیمی موردبررسی قرارگرفته که نتایج آن در جدول ۲ نشان داده شده است. شناخت کلی از منطقه و مطابق آن

جدول ۲. نتایج حاصل از اعمال آزمون جفریس ماتوسیتا بر روی نمونه‌های تعلیمی

Table 2. Results of Jeffris Matusita test on training samples

نمونه‌های تعلیمی	کلاس‌ها	آب	بایر	مسکونی	مشجر	مرتعی فقیر	کشاورزی	مرتعی	جنگلی
۴۲۸	جنگلی	۲	۲	۱/۹۹	۲	۱/۹۹	۱/۹۹	۱/۹۹	-
۱۱۰	مرتعی غنی	۱/۹۹	۱/۹۹	۱/۹۹	۱/۹۹	۱/۹۹	۱/۹۹	-	۱/۹۹
۱۷۳	کشاورزی	۱/۹۹	۱/۹۹	۱/۹۹	۱/۹۹	۱/۹۹	-	۱/۹۹	۱/۹۹
۳۶۳	مرتعی فقیر	۱/۸۶	۱/۸۳	۱/۹۲	۱/۹۹	-	۱/۹۹	۱/۹۹	۱/۹۹
۱۶۷	اراضی مشجر	۲	۲	۱/۹۹	-	۲	۱/۹۹	۱/۹۹	۲
۸۳	مسکونی	۱/۸۴	۱/۹۲	-	۱/۹۹	۱/۹۱	۱/۹۹	۱/۹۹	۱/۹۹
۱۷۸	بایر	۱/۹۳	-	۱/۹۰	۲	۱/۸۰	۱/۹۹	۱/۹۹	۲
۹۵	آب	-	۱/۹۳	۱/۸۷	۲	۱/۸۳	۱/۹۹	۲	۲

میزان بالای ۱/۸۰ دارای جداسازی خوبی بوده و به راحتی از کلاس‌های دیگر جدا می‌شود که نتایج به دست آمده نشانگر جداسازی خوب کلاس‌ها از یکدیگر بوده و صحت نتایج خروجی را کاهش نمی‌دهد.

جدول ۳ دقت کلی و ضریب کاپا مربوط به خروجی‌های تحقیق را نشان می‌دهد. نتایج حاصل از ارزیابی دقت کلی و ضریب کاپا نشان از دقت بالاتر ترکیب بهینه نسبت به تبدیل‌های دیگر دارد. با توجه به جدول ۲ و همچنین معیار رتبه‌بندی میزان همبستگی جفریس ماتوسیتا، نمونه‌هایی با

جدول ۳. دقت کلی و ضریب کاپا خروجی‌های تحقیق

Table 2. Overall accuracy and kappa coefficient of research outputs

تصویر اصلی	کسر حداقل نویز	آنالیز مؤلفه‌های اصلی	آنالیز مؤلفه‌های مستقل	ترکیب تبدیل‌ها	دقت کلی
۷۶	۸۵	۸۰	۷۷	۹۲	دقت کلی
۰/۷۸	۰/۸۸	۰/۸۳	۰/۸۰	۰/۹۴	ضریب کاپا

نتایج بهتری از دو تبدیل آنالیز مؤلفه‌های اصلی و آنالیز مؤلفه‌های مستقل داشته است. این الگوریتم در جداسازی جنگل و مرتع غنی خطای ۲۲ پیکسلی داشته که با توجه به ماهیت این روش و شرایط محیطی منطقه خطای اندکی به حساب می‌آید. بر اساس نتایج صحت سنجی نیز تبدیل کسر حداقل نویز با دقت کلی ۸۵ درصد و ضریب کاپا ۰/۸۸ دقت بالاتر و عملکرد بهتری نسبت به روش‌های آنالیز مؤلفه‌های مستقل و آنالیز مؤلفه‌های اصلی داشته است. تبدیل آنالیز مؤلفه‌های مستقل با فرض غیر گوسی بودن توزیع متغیرها سعی بر مستقل کردن اجزای تصویر از یکدیگر دارد. با توجه به ماهیت این روش، جداسازی بین کاربری شهر و مرتع به دلیل در هم تنیدگی این دو کاربری با خطای ۱۴ پیکسلی روبه‌رو بوده و همچنین در جداسازی دو نوع مرتع خطای ۳۵ پیکسلی در این روش وجود دارد. مقایسه نتایج خروجی‌ها بیانگر عملکرد نزدیک این روش‌ها به یکدیگر است که نشان از وجود هر دو نوع توزیع گوسی و غیر گوسی متغیرها دارد. با توجه به اینکه این تبدیلات به‌تنهایی قادر به استخراج تمامی اجزای تصویر نیستند، لذا ترکیبی از مؤلفه‌های آن‌ها بر اساس ضریب شفیلد جهت مفروض بودن توزیع گوسی و غیر گوسی متغیرها با حداقل نویز ممکن انتخاب گردید. با توجه به معیار کنگولتون و گرین (۷) که ضریب کاپا بالای ۰/۸۵ را مطلوب اعلام کرده‌اند، نتایج تحقیق نشانگر دقت کلی ۹۲ درصد و

بحث و نتیجه‌گیری

امروزه استفاده از انواع تبدیلات جهت بهبود صحت استخراج اطلاعات از تصاویر ماهواره‌ای رو به افزایش است. در این بین انتخاب نوع تبدیل بستگی به ماهیت تصویر و هدف پردازشی آن دارد و استفاده هر کدام از تبدیلات منجر به نتایج خروجی متفاوتی می‌شوند. در این تحقیق بعد از ارزیابی عملکرد الگوریتم‌های آنالیز مؤلفه‌های اصلی، آنالیز مؤلفه‌های مستقل و روش کسر حداقل نویز یک روش ترکیب بهینه مؤلفه‌های این تبدیلات پیشنهاد گردید. طبق نتایج، طبقه‌بندی تصویر اصلی بدون اعمال تبدیلات و با همان نمونه‌های تعلیمی با دقت کلی ۷۶ درصد و ضریب کاپای ۰/۷۸ پایین‌ترین دقت را نسبت به سایر خروجی‌ها داشته است. با توجه به این موضوع لزوم اعمال تبدیلات جهت کاهش همبستگی باندها، مجزا سازی اجزای تصاویر و به دنبال آن افزایش دقت طبقه‌بندی در این پژوهش نمایان شد. بررسی ماتریس خطای خروجی حاصل از تبدیل آنالیز مؤلفه‌های اصلی نشان از تداخل پیکسلی بالای دو کلاس مرتع غنی و فقیر می‌دهد، همچنین این تبدیل در تفکیک جنگل با مراتع نیز با خطاهایی مواجه بوده است. مقدار دقت کلی برای این تبدیل ۸۰ درصد و ضریب کاپا ۰/۸۳ بوده است. تبدیل کسر حداقل نویز، میزان نویز داده‌ها را به حداقل رسانده و در خروجی

دقت برخوردار بودند. نتایج آماری مطلوب و ارزیابی و مقایسه آن با مشاهدات میدانی از منطقه مورد مطالعه و نقاط کنترل زمینی با استفاده از GPS، حاکی از دقت قابل قبول خروجی‌های این پژوهش دارد که این امکان را فراهم می‌کند تا در مطالعات بعدی با استفاده از تصاویر سال‌های قبل و مقایسه آن با این پژوهش، میزان تغییرات را در کاربری‌های مختلف در طول بازه زمانی مشخص نمایان ساخت. همچنین پیشنهاد می‌شود در تحقیقات بعدی جهت بالا بردن صحت و دقت طبقه‌بندی و نتایج نهایی، از تصاویر ماهواره‌هایی مانند Worldview و Quickbird که توان تفکیک بالاتری دارند بهره برد.

References

1. Al-Ahmadi F, Al-Hames A. 2009. Comparison of four classification methods to extract land use and land cover from raw satellite images for some remote arid areas, Kingdom of Saudi Arabia. *Earth Sciences*, 20(1): 167-119.
 2. Alikhah-Asl M, Elham F, Mohammad N. 2014. Evaluation of different enhancement remote sensing techniques. *International Journal of Agriculture Innovations and Research*, 3(1): 33-37.
 3. Anderson JR. 1971. Land-use classification schemes. *Photogrammetric Engineering*, 37(4): 379-387. <https://trid.trb.org/view/93641>.
 4. Arslan O, Akyürek Ö, Kaya Ş. 2017. A comparative analysis of classification methods for hyperspectral images generated with conventional dimension reduction methods. *Turkish Journal of Electrical Engineering & Computer Sciences*, 25(1): 58-72. doi:<https://doi.org/10.3906/elk-1503-167>.
 5. Barzegari Dehaj V, Zare M, Mokhtari MH, Ekhtesasi MR. 2018. Evaluation of different satellite image enhancement techniques in separating of geological units. *Journal of RS and GIS for Natural Resources*, 9(1): 1-23. http://girs.iaubushehr.ac.ir/article_540405_en.html. (In Persian).
 6. Bellvert J, Jofre-Čekalović C, Pelechá A, Mata M, Nieto H. 2020. Feasibility of using the two-source energy balance model (TSEB) with Sentinel-2 and Sentinel-3 images to analyze the spatio-temporal variability of vine water status in a vineyard. *Remote Sensing*, 12(14): 2299. doi:<https://doi.org/10.3390/rs12142299>.
- ضریب کاپا ۰/۹۴ در ترکیب پیشنهادی است. این ترکیب خطای جداسازی بین دو نوع مرتع را به صفر رسانده و در جداسازی اکثر طبقات از یکدیگر دارای بیشترین دقت بوده و تنها در جداسازی کاربری مرتع غنی از جنگل خطای ۱۶ پیکسلی داشته که منجر به عملکرد بهتر و دقت بسیار بالا شده است. همان‌طور که ذکر گردید روش‌های مورد استفاده در جداسازی کلاس‌های درهم‌تنیده و کلاس‌های با نزدیکی طیفی بالا مشکل دارند که منجر به کاهش دقت خروجی‌های آن‌ها شده است. استفاده ترکیبی از این تبدیلات منجر به کاهش خطای خروجی‌ها و افزایش دقت نقشه‌های نهایی شده است. همسو با موضوع این پژوهش، تحقیقات دیگری نیز صورت گرفته که نتایج متفاوتی به همراه داشته است، از جمله؛ لی و همکاران (۱۶)، جهت تهیه نقشه کاربری اراضی از سه روش آنالیز مؤلفه‌های اصلی، آنالیز مؤلفه‌های مستقل و کسر حداقل نویز استفاده کردند. نتایج پژوهش آن‌ها به صورت کلی نشان داد که هر سه روش کاهش ابعاد داده باعث بهبود دقت طبقه‌بندی شدند و همچنین از بین سه روش، الگوریتم کسر حداقل نویز از دقت بیشتری برخوردار بوده است. البته لازم به ذکر است که شرایط و ویژگی‌های منطقه مورد مطالعه مانند شرایط طبیعی و نحوه توزیع متغیرها از لحاظ گوسی و غیرگوسی بودن، نتایج متفاوتی به دنبال دارد. همچنان که در تحقیق دیگر از دبیری و لانگ (۸) در مقایسه سه الگوریتم، تبدیل آنالیز مؤلفه‌های مستقل نتیجه بهتری نسبت به الگوریتم‌های آنالیز مؤلفه‌های اصلی و کسر حداقل نویز داشته است و یا در پژوهشی دیگر از ارسلان و همکاران (۴) جهت افزایش دقت نقشه طبقه‌بندی از تبدیلات کاهش ابعاد داده استفاده شد که تبدیل آنالیز مؤلفه‌های اصلی از نتیجه بهتر و دقت بالاتری نسبت به روش‌های آنالیز مؤلفه‌های مستقل و کسر حداقل نویز برخوردار بوده است. استریمبو و استریمبو (۲۶) در پژوهشی باهدف ارزیابی استراتژی‌های انتخاب مؤلفه روی تصاویر فرا طیفی، سه تکنیک سنتی کاهش ابعاد داده را مورد مقایسه قرار دادند که به ترتیب تبدیلات کسر حداقل نویز، آنالیز مؤلفه‌های اصلی و آنالیز مؤلفه‌های مستقل از بیشترین

7. Congalton RG, Green K. 2019. Assessing the accuracy of remotely sensed data: principles and practices. CRC press. 139 p.
8. Dabiri Z, Lang S. 2018. Comparison of independent component analysis, principal component analysis, and minimum noise fraction transformation for tree species classification using APEX hyperspectral imagery. *ISPRS International Journal of Geo-Information*, 7(12): 488. doi:<https://doi.org/10.3390/ijgi7120488>.
9. ESA. 2017. (Standard Document), SENTINEL-2 User Handbook, 2.
10. ESA. 2018. SNAP-Sen2Cor, Available online: <http://step.esa.int/main/third-party-plugins-2/sen2cor>.
11. Estornell J, Martí-Gavilá JM, Sebastiá MT, Mengual J. 2013. Principal component analysis applied to remote sensing. *Modelling in Science Education and Learning*, 6: 83-89. doi:<https://doi.org/10.4995/msel.2013.1905>.
12. Green AA, Berman M, Switzer P, Craig MD. 1988. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Transactions on geoscience and remote sensing*, 26(1): 65-74. doi:<https://doi.org/10.1109/36.3001>.
13. Guan H, Liu H, Meng X, Luo C, Bao Y, Ma Y, Yu Z, Zhang X. 2020. A quantitative monitoring method for determining Maize lodging in different growth stages. *Remote Sensing*, 12(19): 3149. doi:<https://doi.org/10.3390/rs12193149>.
14. Hyvärinen A, Oja E. 1997. A fast fixed-point algorithm for independent component analysis. *Neural computation*, 9(7): 1483-1492. doi:<https://doi.org/10.1162/neco.1997.9.7.1483>.
15. Ibarrola-Ulzurrun E, Marcello J, Gonzalo-Martin C. 2017. Assessment of component selection strategies in hyperspectral imagery. *Entropy*, 19(12): 666. doi:<https://doi.org/10.3390/e19120666>.
16. Javan F, Hasani Moghaddam H. 2017. Deforestation detection of Hyrcania forest using satellite imagery and Support Vector Machine (Case study: Rezvanshahr county). *Forest Strategical Approachment Journal*, 2(5): 1-13. <https://www.magiran.com/paper/1706792>. (In Persian).
17. Li X, Chen W, Cheng X, Liao Y, Chen G. 2017. Comparison and integration of feature reduction methods for land cover classification with RapidEye imagery. *Multimedia Tools and Applications*, 76(21): 23041-23057. doi:<https://doi.org/10.1007/s11042-016-4311-4>.
18. Luo G, Chen G, Tian L, Qin K, Qian S-E. 2016. Minimum noise fraction versus principal component analysis as a preprocessing step for hyperspectral imagery denoising. *Canadian Journal of Remote Sensing*, 42(2): 106-116. doi:<https://doi.org/10.1080/07038992.2016.1160772>.
19. Manly BF, Alberto JAN. 2016. *Multivariate statistical methods: a primer*. Chapman and Hall/CRC. 269 p. <https://doi.org/10.1201/9781315382135>.
20. Matkan AA, Nohegar A, Mirbagheri B, Torkchin N. 2014. Assessment relations of land use in heat islands using time series ASTER sensor data (Case study: Bandar Abbas city). *Journal of RS and GIS for Natural Resources*, 5(4): 1-14. http://girs.iaubushehr.ac.ir/m/article_516652.html?lang=en. (In Persian).
21. Nascimento JM, Dias JM. 2005. Vertex component analysis: A fast algorithm to unmix hyperspectral data. *IEEE transactions on Geoscience and Remote Sensing*, 43(4): 898-910. doi:<https://doi.org/10.1109/TGRS.2005.844293>.
22. Pohl C, Van Genderen J. 2016. *Remote sensing image fusion: A practical guide*. Crc Press. 288 p. <https://doi.org/10.1201/9781315370101>.
23. Pu R. 2017. *Hyperspectral remote sensing: fundamentals and practices*. CRC Press. 575 p.
24. Richards JA, Richards J. 2013. *Remote sensing digital image analysis*. Springer, Edition Number 5, XIX, 494 p. <https://doi.org/10.1007/978-3-642-30062-2>.
25. Sheffield C. 1985. Selecting Band Combinations from Multi Spectral Data. *Photogrammetric Engineering and Remote Sensing*, 58(6): 681-687. <https://ci.nii.ac.jp/naid/80002491091>.
26. Strîmbu VF, Strîmbu BM. 2015. A graph-based segmentation algorithm for tree crown extraction using airborne LiDAR data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 104: 30-43. doi:<https://doi.org/10.1016/j.isprsjprs.2015.01.018>.
27. Vidhyavathi R. 2017. Principal component analysis (PCA) in medical image processing using digital imaging and communications in medicine (DICOM) medical images. *International Journal of Pharma and Biol Sciences*, 8(2): 598-606. doi:<http://dx.doi.org/10.22376/ijpbs.2017.8.2.b.598-606>.
28. Wang L, Zhang J, Liu P, Choo K-KR, Huang F. 2017. Spectral-spatial multi-feature-based deep learning for hyperspectral remote sensing image classification. *Soft Computing*, 21(1): 213-221. doi:<https://doi.org/10.1007/s00500-016-2246-3>.

29. Yang M-D, Huang K-H, Tsai H-P. 2020. Integrating MNF and HHT transformations into artificial neural networks for hyperspectral image classification. *Remote Sensing*, 12(14): 2327. doi:<https://doi.org/10.3390/rs12142327>.
30. Zhang X, Chen CH. 2002. New independent component analysis method using higher order statistics with application to remote sensing images. *Optical Engineering*, 41: 1717-1728. doi:<https://doi.org/10.1117/1.1482722>.
31. Zhang Y, Zhang J, Yang W. 2020. Quantifying Information Content in Multispectral Remote-Sensing Images Based on Image Transforms and Geostatistical Modelling. *Remote Sensing*, 12(5): 880. doi:<https://doi.org/10.3390/rs12050880>.



Original
paper

Performance evaluation of principal component analysis, independent component analysis and minimum noise fraction method in increasing the information extracting accuracy of Sentinel-2 satellite data

Sayyad Asghari Saraskanrood, Hasan Hasani Moghaddam, Hossein Fekrat

Received: 2 September 2020 / Accepted: 2 November 2020
Available online 21 July 2021

Abstract

Background and Objective Problem The use of various transformations to improve the accuracy of data extraction from satellite images is increasing sharply. In the meantime, the choice of optimal conversion is very important and will affect the output results. Due to the correlated nature of remote sensing images, the use of various transformations to improve the accuracy of information extraction from these images is essential. According to the studies, the purpose of this study is to investigate different methods of image conversion in improving the process of classification of satellite images and increasing the accuracy of land use maps. Considering that the study area and in general the northern regions of Iran are facing special conditions of entanglement of land uses, so the use of various conversion methods as well as

the combined method proposed in this study increases the accuracy and the accuracy of the output information and finally the possibility of more detailed separation and review of uses and identification of factors changing them for future planning.

Materials and Methods In this study, in order to evaluate the performance of principal component analysis methods, independent component analysis, and minimum noise fraction method, Sentinel-2 satellite images of Rezvanshahr city were used. Gram-Schmit algorithm was used to integrate this data with each other and achieve a resolution of 10 meters. After applying the necessary pre-processing and merging the images together, all three transformations were applied to the image, as well as a combination of the components of these three methods. Then, the results of the transformations were classified into 8 user classes using the maximum likelihood algorithm. Using Sheffield coefficient and statistical calculations between the obtained components, the combination of the first components of principal component analysis, the first component of minimum noise fraction, and the second component of independent component analysis were selected as the optimal combination. General knowledge of the area and accordingly the visual interpretation of the outputs, as well as the perception of 120 ground points by GPS, has been the basis for assessing the accuracy of the output maps.

S. Asghari Saraskanrood^(✉) ¹, H. Hasani Moghaddam ²,
H. Fekrat ³

1. Associate Professor, Department of Natural Geography, Faculty of Humanities, University of Mohaghegh Ardabili, Iran
2. MSc. of Remote Sensing and GIS, Faculty of Humanities, Kharazmi University, Tehran, Iran
3. MSc. Student of Remote Sensing and GIS, Faculty of Humanities, University of Mohaghegh Ardabili, Iran

e-mail: s.asghari@uma.ac.ir

<http://dorl.net/dor/20.1001.1.26767082.1400.12.2.4.7>

Results and Discussion After applying the required preprocessors, each of these algorithms was applied to the image, and the output of each was classified into 8 user classes using the Maximum Likelihood algorithm. The results of output maps showed that the conversion of principal component analysis, considering that it considers Gaussian distribution for variables and tries to decompose the extracted components, is weak in samples with non-Gaussian distribution and shows low performance. The minimum noise fraction algorithm works similarly to the principal component analysis algorithm, except that it classifies the noise better. This algorithm has less error in separating classes and this factor has resulted in better performance and higher accuracy than the other two conversions. In the independent component analysis algorithm, the image correlated bands of the study area have been converted to independent components and new information has been extracted from the area. The visual interpretation shows the high accuracy of the classification result and an error matrix (confusion) is used to quantify the accuracy of the classified image.

The results of the evaluation of overall accuracy and kappa coefficient showed that the classification of the original image without applying transformations and with the same training samples of output with an overall accuracy of 76% and kappa coefficient of 0.78 had the highest error. Also, the results of other outputs for classification resulting from principal component analysis conversion are 80% overall accuracy and kappa coefficient of 0.83, respectively, for classification resulting from minimum noise fraction conversion, total accuracy of 85% and kappa coefficient of 0.88 and for the classification obtained

from the analysis of independent component analysis, the overall accuracy was 77% and the kappa coefficient was 0.80. After selecting the optimal combination of components of principal components analysis methods, independent component analysis and minimum noise fraction method and selecting the first components of principal component analysis algorithms and minimum noise fraction and the second component of total component analysis to 92% independent coefficient and Kappa increased 0.94.

Conclusion In this study, after evaluating the conversion performance of principal component analysis, independent component analysis, and minimum noise fraction method, an optimal combination of components of these transformations was proposed. As the results of the research showed, the classification of the original image without conversions and with the same training samples had low overall accuracy and kappa coefficient. The results show the close performance of these transformations to each other, which indicates the existence of both Gaussian and non-Gaussian distributions of variables. MNF conversion has minimized the amount of data noise and results in better output than ICA and PCA conversion. Since these transformations alone are not able to extract all the components of the image, so a combination of the components of these transformations based on the Sheffield coefficient was chosen to assume the Gaussian and non-Gaussian distributions of the variables with the least possible noise.

Keywords: Sentinel-2, Sheffield, Transformation, Maximum likelihood, Rezvanshahr