بررسی دقت روشهای طبقه‌بندی نظارت شده در تهیه نقشه کاربری و پوشش اراضی [1]

مسعود مسعودی، سید کاظم علوی پناه و مسعود نجابت [2]

چکیده

در حال حاضر داده‌های ماهواره‌ای یکی از منابع مهم در تولید نقشه‌های کاربری و پوشش اراضی محسوب می‌شوند. چندین روش طبقه‌بندی رقومی برای تهیه نقشه‌های مذکور پیشنهاد شده و هم اکنون مورد استفاده قرار می‌گیرند. بنابراین ارزیابی و مقایسه روشهای مختلف به منظور تعیین روش مناسب و اصولی در شرایط مختلف کشور از اهمیت زیادی برخوردار است. در این پژوهش چند روش طبقه‌بندی شامل K-نزدیکترین همسایه [3]، حداکثر احتمال [4] و روش جعبه‌ای [5] با استفاده از داده‌های رقومی TM در منطقه کوار استان فارس واقع در جنوب شیراز مورد ارزیابی قرار گرفت. همچنین از GIS [6] برای تصحیح بعضی از پیکسلهای غلط استفاده گردید. بررسی روشهای ذکر شده بر اساس دقت کلی [7]، ضریب کاپا [8] با استفاده از تجزیه و تحلیل

۱- تحقیق فوق با همکاری مرکز تحقیقات مناطق کویری و بیابانی ایران و مرکز تحقیقات منابع طبیعی و امور دام استان فارس انجام شده است.

۲- به ترتیب کارشناس ارشد مؤسسه تحقیقات جنگلها ومراتع، استادیار مرکز تحقیقات مناطق بیابانی و کویری ایران وابسته به دانشگاه تهران، کارشناس ارشد و عضو هیأت علمی مرکز تحقیقات منابع طبیعی و امور دام استان فارس.

3- K-Nearest Neighbour

4- Maximum Likelihood

5- Box Classification

6- Geographic Information System

7- Overall Accuracy

8- Kappa Index

ماتریس خطا[1] انجام شد. طبق نتایج حاصل از این تحقیق مقادیر دقت کلی برای روش حداکثر احتمال برابر ۹۴/۷ ٪، K- نزدیکترین همسایه ۶۱/۷ ٪ و روش جعبه‌ای برابر ۶۰/۴ ٪ می‌باشد که برتری روش طبقه‌بندی حداکثر احتمال نسبت به سایر روشها را نشان می‌دهد. نتیجه نهایی این تحقیق بر تأثیر نوع روش طبقه‌بندی در دقت نقشه‌ها دلالت دارد.

واژه‌های کلیدی: کاربری و پوشش اراضی، داده‌های TM، روش حداکثر احتمال، روش جعبه‌ای، روش K-نزدیکترین همسایه، سیستم اطلاعات جغرافیایی.

مقدمه

اتخاذ تدابیر و سیاستهای صحیح در جهت استفاده بهینه از منابع طبیعی و توسعه پایدار در برنامه‌ریزیهای منطقه‌ای نیازمند اطلاعات کافی وجامع از منابع زیست محیطی است. نقشه‌های کاربری و پوشش اراضی بخش عمده‌ای از اطلاعات مورد نیاز مدیران و برنامه‌ریزان را در امر برنامه ریزی به منظور استفاده بهینه و مطلوب از منابع طبیعی و محیط زیست در جهت توسعه پایدار تامین می‌نماید. نقشه‌های کاربری و پوشش اراضی مدیران و برنامه ریزان را در اتخاذ تدابیر صحیح و تصمیم گیریهای اصولی یاری نموده و نقش بسیار مهمی در برنامه‌ریزی و توسعه منطقه‌ای ایفا می‌کنند.

روشهای مختلفی برای تولید نقشه‌های کاربری و پوشش اراضی وجود دارد. در این رابطه می‌توان به روشهای سنتی با برداشتهای زمینی و همچنین روشهای پیشرفته‌تر با استفاده از اطلاعات ماهواره‌ای اشاره کرد. امروزه قابلیت جایگزینی روشهای پردازش تصویر با روشهای سنتی که به‌طور عام وقت‌گیر و در بعضی موارد کم دقت هستند از پرسشهای اساسی بسیاری از تحقیقات است. به منظور پاسخ به این پرسشها کاربرد

1- Confusion Matrix

تصاویر ماهواره‌ای از جنبه‌های گوناگون مورد آزمایش قرار می‌گیرد که مطالعه حاضر از جمله این تلاشهاست.

تهیه نقشه‌های کاربری و پوشش اراضی از طریق طبقه‌بندی اطلاعات ماهواره‌ای به دو صورت بصری (بدون استفاده از رایانه) و رقومی (استفاده از رایانه) انجام می‌شود. طبقه‌بندی رقومی خود به دو دسته کلی طبقه‌بندی نظارت شده[1] و نظارت نشده[2] تقسیم می‌شود. در طبقه‌بندی نظارت شده با استفاده از اطلاعات نمونه‌های آموزشی و الگوریتم‌های طبقه‌بندی مانند حداکثر احتمال، اقدام به طبقه‌بندی داده‌های طیفی می‌نمایند. در طبقه‌بندی نظارت نشده بدون استفاده از اطلاعات نمونه‌های آموزشی، اطلاعات طیفی را به چند طبقه تقسیم نموده و بعد طبقات اطلاعاتی مربوط به هر طبقه را مشخص می‌کنند.

نتایج حاصل از اجرای روشهای مختلف طبقه‌بندی می‌توانند در شرایط مختلف، متفاوت باشند. با توجه به شرایط خاص هر منطقه اجرای یکی از روشها می‌تواند نسبت به روش دیگر برتر باشد. در این مورد ارزیابی و مقایسه روشهای مختلف به منظور تعیین مناسبترین روش در شرایط مختلف و با داده‌های ماهواره ای مختلف از اهمیت زیادی برخوردار هستند. هدف این پژوهش بررسی روشهای طبقه‌بندی با نظارت شامل حداکثر احتمال، K- نزدیکترین همسایه و روش جعبه ای در تعیین کاربری اراضی با استفاده از داده‌های TM می‌باشد.

در زمینه تعیین الگوریتم‌های مناسب برای تعیین کاربری اراضی مطالعات متعددی توسط متخصصان سنجش از دور در کشورهای مختلف بعمل آمده است. از مطالعات انجام شده در این مورد می‌توان به تحقیقات بووث[3] و اولدفیلد[4] (۱۹۸۹)، گانگ[5]

1- Supervised Classification

2- UnSupervised Classification

3- Booth

وهـــووارث[۶] (۱۹۹۲)، ردی[۷] و دیگـران (۱۹۹۲)، ویـلسون[۸] (۱۹۹۲) و زیـــووان[۹] و هپینگ[۱۰] (۱۹۹۸) اشاره کرد که هر یک به ترتیب روشهای حداکثر احتمال، فرکانس پوششی، حداکثر احتمال، حداکثر احتمال محدود شده[۱۱] و حداکثر احتمال را در میان روشهای مورد مقایسه خود، بهترین روش معرفی نموده‌اند. از تحقیقات صورت گرفته در داخل کشور می‌توان به تحقیق نصیری (۱۳۷۶) اشاره کرد که به ترتیب روش فرکانس پوششی و حداکثر احتمال را مناسبترین روشها برای تعیین کاربری اراضی با استفاده از داده‌های اسپات[۱۲] برشمرده است.

در این تحقیق تنها سه الگوریتم ذکر شده، مورد بررسی قرار گرفت و از بررسی سایر الگوریتم‌ها مانند فرکانس پوششی به علت اینکه ماهیت اصلی آن همان طبقه‌بندی حداکثر احتمال می‌باشد، پرهیز شد.

موقعیت منطقه مورد مطالعه

این منطقه در جنوب شهر شیراز در طول جغرافیایی َ۳۰ و °۵۲ تا َ۴۵ و °۵۲ شرقی و عرض جغرافیایی °۲۹ تا َ۱۵ و °۲۹ شمالی قرار گرفته است. مساحت محدوده مورد مطالعه ۶۹۵۰۸ هکتار می‌باشد. حداکثر و حداقل ارتفاع آن به ترتیب ۲۹۱۰ و ۱۴۵۰ متر، متوسط بارندگی سالانه منطقه فوق ۴۱۹ میلیمتر و اقلیم منطقه بر اساس روش کوپن از ارتفاع ۱۴۰۰ تا ۲۵۰۰ متر از نوع مدیترانه‌ای برّی و از ارتفاع ۲۵۰۰ متر به بالا از نوع مدیترانه‌ای بحری می‌باشد.

4- Oldfield
5- Gong
6- Howarth
7- Reddy
8- Wilson
9- Xiuwan
10- Heping
11- Penalized Maximum Likelihood
12- SPOT

مواد و روشها

به منظور ارزیابی الگوریتم‌های مناسب در مطالعات کاربری اراضی از تصاویر ماهواره‌ای لندست با هفت باند TM به تاریخ ۲۱ مهر ماه ۱۳۶۹، نقشه‌های توپوگرافی به مقیاس‌های ۱:۵۰۰۰۰ و ۱:۲۵۰۰۰۰، نقشه‌های قابلیت و طبقه‌بندی اراضی استفاده گردید و تحلیل نقشه‌ها به وسیله نرم‌افزارهای ایلویس[1] و آرکینفو[2] انجام شد.

این مطالعه شامل مراحل زیر می‌باشد :

۱ - **مرحله پیش پردازش[3] و بارزسازی تصاویر[4]** : در این مرحله ابتدا تصحیحات اتمسفری و هندسی بر روی داده‌های خام انجام و بعد با اعمال روشهای مورد اشاره در زیر بر روی باندها و تصاویر پدیده‌های موجود بر روی تصاویر واضحتر گردیدند. روشهایی که در این مرحله مورد استفاده واقع شدند عبارتند از:

الف - افزایش مغایرت و وضوح تصاویر با استفاده از کشش خطی[5] و هیستوگرام متعادل شده[6] نمودار فراوانی ارزشهای طیفی

ب - استفاده از تصاویر رنگی کاذب[7] (FCC)

ج - استفاده از فیلترهای هموارکننده و وضوح بخش لبه‌ها

۲ - **طبقه‌بندی اطلاعات** : طبقه‌بندی اطلاعات[8] ماهواره‌ای به دو صورت بصری[9]

1- (Integrated Land and Water Information System) ILWIS

2- ARCINFO

3- Pre-Processing

4- Image Enhancement

5- Linear Stretching

6- Equalization Histogram

7- False Color Composite

8- Data Classification

9- Visual

و رقومی [1] صورت گرفت. در تفسیر بصری تصویر رنگی کاذب (FCC) مورد تجزیه و تحلیل قرارگرفت و بدین‌وسیله ویژگی طیفی و رنگ عوارض پدیده‌های مورد نظر مورد شناسایی قرار گرفت. طبقه‌بندی رقومی با استفاده از داده‌های رقومی و روشهای مختلف طبقه‌بندی به وسیله رایانه انجام گرفت. در این تحقیق با توجه به نتایج عامل OIF (Optimum Index Factor) که مبتنی بر همبستگی و واریانس کل باندها می‌باشد و اهمیت شناخته شده باندهای ۳ و ۴ در تشخیص پوشش گیاهی از باندهای ۳، ۴، ۵ و ۷ استفاده گردید. همچنین نمونه‌های آموزشی به عنوان مرجع زمینی برای کنترل طبقات کاربری حاصل از طبقه‌بندی رقومی انتخاب گردید. به منظور انتخاب نمونه‌های آموزشی از تصاویر ایجاد شده با رنگ کاذب (FCC) استفاده گردید و نقاط پس از انتخاب در صحرا کنترل گردیدند.

۱-۲ - روش طبقه‌بندی: پس از اتمام نمونه‌برداری از طبقات کاربری در منطقه مطالعاتی، نوبت به طبقه‌بندی طبقات رسید. در این تحقیق از سه روش طبقه‌بندی نظارت شده حداکثر احتمال، روش جعبه‌ای و k- نزدیکترین همسایه برای طبقه‌بندی طبقات استفاده شد. هر یک از سه روش فوق دارای مقادیر و عوامل گوناگون در اجرای طبقه‌بندی می‌باشند. از هر روش، بر اساس مؤلفه‌های آن سه نوع طبقه‌بندی انجام شد. بعد از هر روش نقشه‌ای که دارای بالاترین دقت (براساس عوامل مؤثر بر چگونگی طبقه‌بندی آن) و میزان مساحت تعریف نشده کمتری بود برای بررسی نهایی استفاده شد.

در زیر به شرح انتخاب بهترین مؤلفه برای هر روش پرداخته می‌شود. برای روش

1- Digital

جعبه‌ای تحقیق سه نقشه با مقادیر مختلف ضریب افزایشی[1] برای تعیین انحراف معیار نمونه‌های طبقات تهیه گردید که در جدول شمارهٔ (۱) نتایج آن مشاهده می‌شود.

جدول شمارهٔ ۱ - دقت نقشه‌های طبقه‌بندی شده براساس بررسی سطوح نمونه‌های آموزشی و مقایسه آن با روش جعبه‌ای و مقادیر مختلف ضریب افزایشی برای انحراف معیار

مقادیر مختلف Multiplication Factor	۱/۷۳	۲	۳
دقت کلی به درصد	۸۱/۹	۷۶/۳	۵۲/۵

با توجه به نتیجه بدست آمده، از نقشه حاصل از طبقه‌بندی روش جعبه‌ای با ضریب افزایشی ۱/۷۳ که دارای بالاترین دقت است استفاده گردید. لازم به ذکر است که مقدار انتخابی ۱/۷۳ مقدار پیشنهادی رایانه برای استفاده از این روش و نقشه‌های ایجادی از مقادیر کمتر از آن، دارای میزان مساحت تعریف شده بیشتری بود. بنابراین از بررسی مقادیر کمتر از آن خودداری شد.

در روش طبقه‌بندی k- نزدیکترین همسایه ابتدا تعداد پیکسلهایی را که برای تعیین تعلق یک پیکسل نامشخص به طبقه مورد نظر پذیرفته‌ایم، تحت عنوان تعداد همسایه‌ها[2] تعیین می‌شود. سپس از رایانه خواسته می‌شود که در یک شعاع معین به نام شعاع جستجو[3] به تجسس بپردازد. در صورتی که تعداد پیکسل تعیین شده در آن محدوده از یک کلاس خاص وجود داشته باشد، پیکسل نامشخص به آن طبقه پیوند خواهد خورد.

1- Multiplication Factor

2- Nr. of neighbours

3- Search radius

در روش اخیر نیز با استفاده از مقادیر مختلف تعداد همسایه‌ها برای شعاع جستجو ۵ (با افزایش شعاع جستجو از دقت طبقه‌بندی کاسته می‌شود) سه نقشه تهیه شد که نتایج دقت کلی در جدول شماره ۲ مشاهده می‌شود.

جدول شمارهٔ ۲- دقت نقشه‌های طبقه‌بندی شده از بررسی سطوح نمونه‌های آموزشی با روش k- نزدیکترین همسایه باشعاع جستجو= ۵ و مقادیر مختلف تعداد همسایه‌ها

مقادیر مختلف *nr.of neighbours* *Search radius*=برای۵	۱	۳	۵
دقت کلی به درصد	۸۳/۳	۸۹/۶	۸۶/۷

بنابراین با توجه به جدول بالا برای بررسی نقشه نهایی از روش K- نزدیکترین همسایه از نقشه طبقه‌بندی شده باشعاع جستجو = ۵ و تعداد همسایه =۳ که از دقت بالاتری برخوردار بود استفاده شد.

برای روش طبقه‌بندی حداکثر احتمال نیز سه نقشه با مقادیر مختلف حد آستانه[1] استفاده شد که در جدول شماره (۳) نتایج آن مشاهده می‌شود. افزایش حد آستانه نیز همانند ضریب افزایشی باعث افزایش انحراف معیار نمونه‌های طبقات طیفی و در نتیجه کاهش دقت طبقه‌بندی می‌گردد.

1- Threshold for distance

جدول شماره ۳- دقت نقشه‌های طبقه‌بندی شده از بررسی سطوح نمونه‌های آموزشی با روش حداکثر احتمال و با مقادیر مختلف حد آستانه

مقادیر مختلف *Threshold distance*	۲۰	۲۵	۳۰
دقت کلی به درصد	۹۴/۷	۹۳/۸	۸۹/۱

بنابراین با توجه به جدول فوق از نقشه طبقه‌بندی با روش حداکثر احتمال با حد آستانه ۲۵ که دارای دقت بالایی بوده و میزان پیکسلهای طبقه‌بندی نشده آن خیلی کمتر از نقشه با حد آستانه ۲۰ می‌باشد، استفاده گردید.

۲-۲ - عملیات بعد از طبقه‌بندی [1] : در این مرحله ۲۴ طبقه طیفی حاصل از طبقه‌بندی به ۱۰ طبقه اطلاعاتی فوق تبدیل گردیدند.

۱ - اراضی طبقه‌بندی نشده ۲ - زراعت آبی

۳ - مراتع ۴ - اراضی شخم خورده

۵ - شالیزار ۶ - جنگل

۷ - باغها ۸ - تاکستان

۹ - زراعت دیم ۱۰ - اراضی زراعی برداشت شده

لازم به ذکر است که با توجه به تاریخ تصویر برداری توسط ماهواره، در ۲۱ مهر ماه، بیشتر اراضی زراعی به خصوص دیم برداشت گردیده و اراضی زراعی اعم از دیم یا آبی دارای پوشش گیاهی بسیار کمی بودند. این موضوع باعث سردرگمی برای تفکیک این عرصه از سایر عرصه‌ها مانند مراتع می‌گردید. برای رفع این محدودیت در طبقه‌بندی

1- Post - Classification

تصویر و دست‌یابی به نقشه نهایی و صحیح کاربری اراضی از GIS استفاده گردید. به عنوان مثال برای اینکه دو طبقه اراضی زراعی شخم خورده و زراعی برداشت شده را جزو طبقات زراعی آبی و دیم دربیاوریم، از اطلاعات موجود در زمینه ارتفاع و شیب طبقات منطقه و محیطهای ایلویس و آرکینفو بهره‌گیری شد که شرح آن در پی خواهد آمد. همچنین برای تصحیح مناطق مسکونی و رودخانه‌ها بر روی تصویر طبقه‌بندی شده که با طبقه‌بندی رقومی داده‌های TM توأم با خطاهای زیاد بود از روش مشابه استفاده شد.

لایه‌های اطلاعاتی تهیه شده از روی نقشه ۱:۵۰۰۰۰ منطقه تهیه و شامل لایه‌های محدوده شیت نقشه، خطوط کنتور، رودخانه، مناطق روستایی و منطقه شهری می‌باشد. از لایه‌های مهم تهیه شده، لایه خطوط کنتور می‌باشد که از آن DEM[1] یا مدل رقومی ارتفاعی منطقه بدست آمد. از DEM برای تهیه مدل سه بعدی، نقشه شیب [2] و جهت شیب[3] استفاده گردید. با استفاده از DEM و نقشه شیب، تصحیحات لازم بر روی آن دسته از پیکسلهایی که نادرست طبقه‌بندی شده بودند، انجام گرفت. تصحیحات انجام شده همگی در قسمت Map Calculation با توجه به بازدیدهای صحرایی از منطقه صورت گرفت که به شرح زیر برای نقشه طبقه‌بندی شده با روش حداکثر احتمال عبارتند از:

ML 1: = IF ((ML = 4) and (slope < = 2) OR (ML = 10) and (slope < = 2), 2, ML).

با نوشتن فرمول بالا اراضی شخم خورده (طبقه ۴) و زراعی برداشت شده (طبقه ۱۰) که دارای شیب ۲ و کمتر از ۲ درصد می‌باشند به اراضی زراعی آبی (طبقه ۲) تبدیل

1- Digital Elavation Model

2- Slope map

3- Aspect map

می‌شوند. ML تصویر طبقه‌بندی شده با روش حداکثر احتمال می‌باشد که دقیقاً برابر و مطابق با لایه‌های اطلاعاتی و شیت نقشه ۱:۵۰۰۰۰ منطقه مطالعاتی می‌باشد.

ML 2: = IF ((ML1 = 2) and (slope > = 5) OR (MLI = 2) and (DEM > = 1800) OR(ML 1 = 4) OR (ML1 = 10), 9, ML1).

بدین ترتیب اراضی زراعی آبی با شیب مساوی و بیشتر از ۵ و ارتفاع بیش از ۱۸۰۰ متر و باقیمانده اراضی شخم خورده و زراعی برداشت شده که دارای شیب بیش از ۲ درصد می‌باشند به اراضی با زراعت دیم تبدیل می‌شوند.

ML 3: =IF ((ML2 = 9) and (slope > = 12), 3, ML2).

با نوشتن این فرمول اراضی زراعی دیم (طبقه ۹) با شیب برابر و بیش از ۱۲ درصد به اراضی مرتعی (طبقه ۳) تبدیل می‌گردند.

ML 5: = IF ((ML4 = 5) and (DEM > = 1690) and (DEM < = 1700), 2, ML4).

فرمول اخیر نیز شالیزارهایی (طبقه ۵) راکه در ارتفاع بین ۱۶۹۰ تا ۱۷۰۰ متر می‌باشند به زراعت آبی تبدیل می‌نماید.

پس از تصحیحات لازم، طبقات رودخانه، مناطق روستایی و شهری نیز که در محیط ایلویس ترسیم شده بودند بر روی نقشه تصحیح شده منتقل گردیدند.

۳-۲- ارزیابی تصاویر طبقه‌بندی شده: برای تعیین دقت نقشه تولیدی در این تحقیق پس از بازدید مجدد از منطقه و با استفاده از نمونه‌های شاهد (آزمایشی)، خطای طبقه‌بندی از طریق جدول ماتریس خطا محاسبه گردید. خطای حاصل از طبقه‌بندی تحت عنوان دو نوع خطای کمسیون[1] و امسیون[2] ارزیابی گردید. خطای کمسیون یک

1- Comision

طبقه نشان دهنده میزان مساحت اراضی یک کلاس در یک نقشه می‌باشد که به واقع به آن تعلق نداشته و خطای امسیون یک طبقه نشان دهنده میزان مساحت اراضی یک طبقه می‌باشد که جزو طبقات دیگر بر روی نقشه طبقه‌بندی گردیده است. در جداول مورد نظر اعداد واقع بر روی قطر اصلی نشان دهنده پیکسلهای صحیح از نمونه‌های شاهد برای هر یک از کلاسها می‌باشد و سایر اعداد نشان دهنده پیکسلهای غلطی می‌باشند که وارد طبقه دیگری گردیده‌اند. همچنین برای محاسبه دقت هر طبقه‌بندی میزان دقت کلی و کاپای کلی برآورد گردید. دقت کلی از محاسبه درصد تقسیم مجموع پیکسلهای صحیح بر کل پیکسلها بدست آمد.

نتایج و بحث

ماتریس خطای طبقه‌بندی حاصل از K- نزدیکترین همسایه در جدول شماره (۴) نشان می‌دهد که طبقات مختلف کاربری در این روش تداخل زیادی را با یکدیگر نشان می‌دهند. بیشترین تداخل در میان طبقه مرتع با طبقات زراعت دیم، زراعت آبی و جنگل و بین طبقه تاکستان با زراعت آبی و مرتع و بین دو طبقه زراعت دیم و جنگل و همچنین بین طبقه زراعت آبی با شالیزار و باغها مشاهده می‌شود. بدین شکل که بخش بیشتر طبقه مرتع (۷۱/۹ درصد) به عنوان طبقه جنگل طبقه‌بندی گردیده است. لیکن مساحتهایی از طبقات زراعت دیم و زراعت آبی و جنگل نیز جزو طبقه مرتع طبقه‌بندی گردیده‌اند. همچنین میزان زیادی از طبقه تاکستان جزو طبقه زراعت آبی و به میزان کمتری جزو طبقه مرتع طبقه‌بندی گردیده است. همچنین قسمتی از اراضی دیم به عنوان طبقه جنگل و قسمتهایی از شالیزارها و باغها جزو طبقه زراعت آبی طبقه‌بندی شده‌اند.

2- Emision

جدول شماره ۴ـ ماتریس خطای نتایج طبقه‌بندی روش K- نزدیکترین همسایه

طبقات طبقه‌بندی شده (با علامت اختصاری)	زراعت آبی = I، مرتع =Ra، شالیزار = P، جنگل = F، باغها = O، تاکستان = G، زراعت دیم = D علامت اختصاری طبقات رفرنس (اعداد ذکر شده بر حسب پیکسل می‌باشند.)							جمع	خطای کمسیون (٪)
	I	Ra	P	F	O	G	D		
زراعت آبی (I)	۲۶۴	۶	۲۸	۰	۳۱	۲۱	۰	۳۵۰	۲۴/۶
مرتع (Ra)	۳۴	۱۴۰	۲	۱۴۲	۰	۸	۲۷	۳۵۳	۶۰/۳
شالیزار (P)	۰	۰	۱۵۱	۰	۱۹	۴	۰	۱۷۴	۱۳/۲
جنگل (F)	۲۷	۴۳۹	۸	۵۶۵	۱	۱۱	۴۹	۱۱۰۰	۴۸/۶
باغها (O)	۰	۰	۰	۰	۸۲	۴	۰	۸۶	۴/۷
تاکستان (G)	۰	۰	۰	۰	۰	۳۸	۰	۳۸	۰/۰
زراعت دیم (D)	۰	۲۶	۱	۰	۰	۰	۱۸۹	۲۱۶	۱۲/۵
جمع	۳۲۵	۶۱۱	۱۹۰	۷۰۷	۱۳۳	۸۶	۲۶۵	۲۳۱۷	
خطای امسیون (٪)	۱۸/۸	۷۷/۱	۲۰/۵	۲۰/۱	۳۸/۳	۵۵/۸	۲۸/۷		

دقت کلی (به درصد) = ۶۱/۷ کاپای کلی (به درصد) = ۵۰/۵

جدول شماره (۵) ماتریس خطای بدست آمده از روش جعبه‌ای را نشان می‌دهد. بیشترین تداخل در این طبقه‌بندی بین طبقه مرتع با زراعت دیم، زراعت آبی و جنگل و بین طبقه زراعت آبی با تاکستان، باغها و شالیزار و نیز بین طبقات باغها و شالیزار وجود دارد.

تداخل عمده میان طبقات مذکور بدین شکل است که مقادیری از مساحت هر یک از طبقات زراعت آبی (۱۹٪)، جنگل (۶۵٪) و زراعت دیم(۲۷٪) جزو طبقه مرتع طبقه‌بندی شده‌اند. همچنین در حدود ۱۶ درصد از طبقه تاکستان به طبقه زراعت آبی تعلق گرفته است و نزدیک به ۲۷ درصد از شالیزارهای منطقه جزو اراضی زراعی آبی شده‌اند. از سوی دیگر نزدیک به ۴۰ درصد از باغهای موجود در منطقه جزو طبقه زراعت آبی طبقه‌بندی شده‌اند. همچنین در حدود ۳۷ درصد از شالیزارهای طبقه‌بندی شده بر روی نقشه جزو اراضی باغها می‌باشند.

جدول شماره (۶) ماتریس خطای طبقه‌بندی برای روش حداکثر احتمال را نشان می‌دهد. در این روش دقت طبقه‌بندی بیش از دو روش دیگر و همچنین میزان خطای کمسیون و امسیون طبقات کاربری کمتر از دو روش پیشین است. با مشاهده جدول دیده می‌شود که بیشترین تداخل در این طبقه‌بندی بین طبقه مرتع با جنگل و زراعت دیم، طبقه تاکستان با زراعت آبی و همچنین بین طبقه زراعت آبی با شالیزار و باغها مشاهده می‌شود.

ماتریس خطای روشهای طبقه‌بندی (جداول شماره ۴، ۵ و ۶) نشان می‌دهد که به ترتیب روشهای حداکثر احتمال، K- نزدیکترین همسایه و شبکه‌ای دارای دقت کلی ۹۴/۷٪، ۶۱/۷٪ و ۶۰/۴٪ می‌باشند. در میان سه روش فوق تنها روش حداکثر احتمال با تفاوت دقت بالا نسبت به دو روش دیگر قابلیت استفاده برای تعیین کاربری اراضی را دارا می‌باشد.

جدول شماره ۵- ماتریس خطای نتایج طبقه‌بندی روش جعبه‌ای

طبقات طبقه‌بندی شده (با علامت اختصاری)	زراعت آبی = I، مرتع=Ra، شالیزار = P، جنگل = F، باغها = O، تاکستان = G، زراعت دیم = D علامت اختصاری طبقات رفرنس (اعداد ذکر شده بر حسب پیکسل می‌باشند.)							جمع	خطای کمسیون (٪)
	I	Ra	P	F	O	G	D		
زراعت آبی (I)	۲۴۳	۲۱	۴۷	۰	۵۵	۱۴	۰	۳۸۰	۳۶/۱
مرتع (Ra)	۵۸	۵۴۹	۱۳	۴۶۰	۰	۱۷	۷۳	۱۱۷۰	۵۳/۱
شالیزار (P)	۰	۰	۱۰۷	۰	۶۶	۰	۳	۱۷۶	۳۹/۲
جنگل (F)	۹	۹	۲	۲۴۷	۰	۲	۱۹	۲۸۸	۱۴/۲
باغها (O)	۰	۰	۰	۰	۱۲	۲	۰	۱۴	۱۴/۳
تاکستان (G)	۰	۰	۳	۰	۰	۵۳	۰	۵۶	۵/۴
زراعت دیم (D)	۰	۳۲	۲	۰	۰	۰	۱۷۳	۲۰۷	۱۶/۴
جمع	۳۱۰	۶۱۱	۱۷۴	۷۰۷	۱۳۳	۸۸	۲۶۸	۲۲۹۱	
خطای امسیون (٪)	۲۱/۶	۱۰/۱	۳۸/۵	۶۵/۱	۹۱/۰	۴۰/۰	۳۵/۴		

دقت کلی (به درصد) = ۶۰/۴ کاپای کلی (به درصد) = ۴۹/۶

جدول شماره ۶- ماتریس خطای نتایج طبقه‌بندی روش حداکثر احتمال

طبقات طبقه‌بندی شده (با علامت اختصاری)	زراعت آبی = I، مرتع =Ra، شالیزار = P، جنگل = F، باغها = O، تاکستان = G، زراعت دیم = D علامت اختصاری طبقات رفرنس (اعداد ذکر شده بر حسب پیکسل می‌باشند.)							جمع	خطای کمسیون (٪)
	I	Ra	P	F	O	G	D		
زراعت آبی (I)	۳۲۷	۰	۱۷	۰	۱۲	۷	۰	۳۶۳	۹/۹
مرتع (Ra)	۵	۵۷۹	۱	۴۱	۰	۵	۰	۶۳۱	۸/۲
شالیزار (P)	۰	۰	۹۱	۰	۲	۰	۰	۱۹۳	۱/۰
جنگل (F)	۰	۲	۰	۶۶۵	۰	۱	۰	۶۶۸	۰/۴
باغها (O)	۰	۰	۱	۰	۱۱۹	۲	۰	۱۲۲	۲/۵
تاکستان (G)	۰	۰	۲	۰	۰	۸۵	۰	۸۷	۲/۳
زراعت دیم (D)	۰	۲۸	۰	۰	۰	۰	۲۶۵	۲۹۳	۹/۶
جمع	۳۳۲	۶۰۹	۲۱۲	۷۰۶	۱۳۳	۱۰۰	۲۶۵	۲۳۵۷	
خطای امسیون (٪)	۱/۵	۴/۹	۹/۹	۵/۸	۱۰/۵	۱۵/۰	۰/۰		

دقت کلی (٪) = ۹۴/۷ کاپای کلی (٪) = ۹۳/۳

$$\text{دقت کلی به درصد} = \frac{\text{مجموع پیکسل‌های صحیح}}{\text{تعداد کل پیکسل‌ها}} \times ۱۰۰$$

$$\text{دقت کلی به درصد} = \left(\frac{۳۲۷+۵۷۹+۱۹۱+۶۶۵+۱۱۹+۸۵+۲۶۵}{۲۳۵۷}\right) \times ۱۰۰ = ٪۹۴/۷$$

نتایج این تحقیق همچنین نشان می‌دهند که چهار باند ۳، ۴، ۵ و ۷ برای طبقه‌بندی و تهیه نقشه کاربری اراضی مناسب می‌باشند و نیازی به استفاده از تمامی باندهای TM نمی‌باشد. همچنین استفاده از GIS به عنوان یک ابزار برای تصحیح کلاسهایی نظیر مناطق مسکونی و رودخانه که ایجاد آنها با طبقه‌بندی رقومی حاوی خطای زیاد می‌باشد، مناسب است.

پیشنهادها

برای تهیه نقشه پوشش و کاربری اراضی مناسبتر است که از داده‌های چند زمانه استفاده گردد و در صورت عدم وجود آن، داده‌های اواخر بهار و یا اوایل تابستان به علت موجود بودن پوشش گیاهی توصیه می‌گردد. همچنین ایجاد اولیه طبقات طیفی برای تشکیل طبقات اطلاعاتی و استفاده از سیستم اطلاعات جغرافیایی و نقشه‌های کمکی ما را در تهیه نقشه‌هایی دقیقتر به خصوص برای افزایش دقت طبقات مرتع، جنگل، زراعت دیم و مسکونی که توأم با اختلاط با سایر طبقات و خطا است، راهنمایی می‌کند.

سپاسگزاری

لازم می‌دانم از همه کسانی که به نحوی در این تحقیق اینجانب را یاری نموده‌اند، به‌خصوص از اعضای محترم جهاد سازندگی فارس، مرکز تحقیقات مناطق کویری و بیابانی و آقای علی اکبر نوروزی کارشناس ارشد مرکز تحقیقات آبخیزداری صمیمانه سپاسگزاری و قدردانی نمایم. همچنین از راهنماییهای آقایان مهدی فرح‌پور، اسماعیل رهبر، مجتبی پاک‌پرور و محمد درویش در تهیه این مقاله تشکر می‌نمایم.

منابع

۱- نصیری، ع. ۱۳۷۶. روش‌های طبقه‌بندی طیفی در تهیه نقشه‌های کاربری و پوشش اراضی. نشریه شماره ۷۶/۲۶ اداره کل آمار و اطلاعات وزارت کشاورزی.

2- Booth, D.J. and R.B. Oldfield. (1989). A Comparison of Algorithms in terms of Speed and Accuracy after the Application of Post Classification Modal Filter. International Journal of Remote Sensing, Vol. 10, 7: 1271-1276.

3- Gong, P. and P.J., Howarth. (1992). Land Use Classification of SPOT HRV Data Using a Cover Freqency Method. International Journal of Remote Sensing, Vol. 13, 8: 1459-1471.

4-Reddy, C.S.S., Cetin, H. and A. Chaturvedi. (1992). An Evaluation of Classification Algorithms for Land Cover Mapping. ASIAN-PACIFIC Remote Sensing Journal, Vol. 5, No. 1.

5- Wilson, J.D. (1992). A Comparison of Procedures for Classifying Remotely Sensed Data Using Simulated Data Sets incorporating Autocorrelation between Spectral Responses International Journal of Remote Sensing, Vol. 13, 2: 365 - 386.

6- Xiuwan, C. and H. Heping. (1998). Evaluation of algorithms for land cover classification from Landsat TM data in Ansan. Republic of Korea, ASIAN-PACIFIC Remote Sensing and GIS Journal, Vol. 10, 2: 61-69.

Accuracy assessment of supervised classification methods for landuse and cover mapping

M. Masoudi[1], *S.K. Alavipanah*[2] *and M. Nejabat*[3]

Abstract

Nowaday, satellite data are considered as one of the most important sources for production of landuse/cover maps. Several digital classification methods have been applied and evaluation and comparison of the methods for classification has a great degree of importance tasks. In this research, some classification methods, e.g. K-Nearest Neighbour, Maximum Likelihood and Box classification were used for evaluating landuse maps in Kavar area which is located in the south of Shiraz. Also, the GIS was applied in this research for correcting the wrong pixels of some classes. Evaluation of the accuracy were performed using over-all accuracy and KAPPA index which are based on analysis of error matrix. According to these results, the over-all accuracy for Maximum Likelihood, K-Nearest Neighbour and Box classification has been reached to 94.7, 61.7 and 60.4 percent respectively. Results of this research show that the capability of Maximum Likelihood is higher than the other methods.

Key words: landuse, TM data, Maximum Likelihood, Box classification, K-Nearest Neighbour, GIS.

1- Junior research scientist of Research Institute of Forests and Rangelands.

2- Assistant Professor of Iran Desert Research Center - University of Tehran.

3- The member of scientific board of Natural Resources Research Center of Fars.