نشریه علمی- پژوهشی علوم و فنون نقشه برداری، دوره پنجم، شماره ۲، آبان ماه ۱۳۹۴

مدل‌سازی تغییرات کاربری اراضی چندگانه با استفاده از رگرسیون انطباقی چند متغیره اسپلاین و سیستم اطلاعات مکانی

محمد احمدلو [۱]، محمودرضا دلاور*[۲]

[۱]دانشجوی کارشناسی ارشد سیستم‌های اطلاعات مکانی- دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی پردیس دانشکده‌های فنی- دانشگاه تهران
m_ahmadlou@ut.ac.ir

[۲] دانشیار دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی -پردیس دانشکده‌های فنی-دانشگاه تهران
(قطب علمی مهندسی نقشه‌برداری در مقابله با سوانح طبیعی)
mdelavar@ut.ac.ir

(تاریخ دریافت فروردین ۱۳۹۳، تاریخ تصویب خرداد ۱۳۹۴)

چکیده

لزوم بررسی تغییرات کاربری اراضی، با توجه به تاثیرات قابل ملاحظه‌ای که این پدیده بر روی زندگی انسان و همچنین محیط زیست می‌گذارد، امری لازم و ضروری به نظر می‌رسد. مدلسازان تغییرات کاربری اراضی عمدتا به مدلسازی تغییرات کلاسهای کاربری اراضی یگانه (مانند شهری) پرداخته اند و کمتر مدلسازی تغییرات کلاسهای کاربری اراضی چندگانه (مانند کشاورزی و شهری) مورد توجه قرار گرفته است. در این تحقیق، به شبیه سازی تغییرات کاربری‌های اراضی شهری و کشاورزی با استفاده از یکی از فنون داده کاوی مرسوم، یعنی رگرسیون انطباقی چند متغیره اسپلاین، در تلفیق با سیستم اطلاعات مکانی و به کمک تصاویر ماهواره‌ای در بخشی از شمال ایران، شامل شهرهای ساری و قائم شهر و در یک بازه زمانی ۲۲ ساله از سال ۱۹۹۲ تا سال ۲۰۱۴ پرداخته شده است. بعد از انجام شبیه سازی، یک تحلیل حساسیت بر روی پارامترهای موثر در تغییر کاربری اراضی انجام شد. نتایج مربوط به تحلیل حساسیت نشان داد که پارامترهای فاصله از مناطق کشاورزی، فاصله از مناطق شهری و ارتفاع به ترتیب بیشترین تاثیر را در روند تغییر کاربری اراضی در بازه زمانی ذکر شده دارند. برای ارزیابی کارآیی مدل، از شاخص عامل نسبی و شاخص عامل کلی استفاده شد. مساحت زیر منحنی شاخص عامل نسبی برای کاربری های شهری و کشاورزی به ترتیب برابر با ۶۵٪ و ۶۱/۰۱٪ بدست آمد. همچنین برای نشان دادن مقادیر موجود در داخل جداول احتمال، در منحنی شاخص عامل کلی مربوط به کاربری اراضی کشاورزی و شهری، حدود آستانه به ترتیب ۰/۶۷ و ۰/۴۰ برچسب گذاری شد که این حدود آستانه نشان دهنده احتمال تغییر کاربری اراضی هستند. طبق نتایج بدست آمده، درصد پیکسل‌هایی که در واقعیت دچار تغییر کاربری اراضی شده اند و توسط مدل نیز به عنوان پیکسل‌های تغییر یافته شناخته شده‌اند برای حدود آستانه فوق در نمودار شاخص عامل کلی، به ترتیب برابر با ۳۶/۸٪ و ۶۷/۰۶٪ از کل پیکسل‌های متعلق به دو کاربری کشاورزی و شهری بدست آمدند.

واژگان کلیدی: تغییرات کاربری اراضی، رگرسیون انطباقی چند متغیره اسپلاین، شاخص عامل نسبی، شاخص عامل کلی

---

* نویسنده رابط

۱- مقدمه

در چند دهه اخیر، مدلسازی تغییرات کاربری اراضی با توجه به اهمیت آن، مورد توجه محققان زیادی در سراسر جهان قرار گرفته است. عوامل زیادی در تغییرات کاربری اراضی دخالت دارند که در یک تقسیم‌بندی کلی می‌توان آنها را به عوامل جمعیتی (مانند پدیده مهاجرت از روستاها به شهرها و رشد جمعیت)، عوامل اقتصادی، عوامل بیوفیزیکی (مانند ارتفاع و جنس خاک)، عوامل اجتماعی (مانند سیاست‌ها) و عوامل فرهنگی تقسیم کرد [۱]. عوامل مذکور به عنوان عوامل محرک برای تغییرات کاربری اراضی محسوب می‌شوند. اگر یک شهر به عنوان یک موجود زنده تصور شود، عوامل مذکور باعث تغییر رفتار این موجود می‌شوند و این عامل نیز به نوبه خود باعث تغییر روال زیست‌بوم می‌شود.

با تغییرات کاربری اراضی در یک منطقه ممکن است مسائل و مشکلاتی در چرخه آب و هوای منطقه، امنیت غذایی (با توجه به تخریب زمین‌های کشاورزی اطراف شهرها)، چرخه آب منطقه و زیست‌بوم طبیعی منطقه بوجود آید [۲]. با توجه به اثراتی که تغییرات کاربری در زندگی انسان و محیط زیست دارد لزوم بررسی این پدیده امری ضروری به نظر می رسد. برای بررسی روند تغییرات کاربری اراضی در یک منطقه نیاز به داشتن اطلاعاتی در مورد عوامل محرک تغییرات کاربری اراضی می‌باشد [۳]. این عوامل در طیف وسیعی از مقیاس های مکانی و زمانی و در یک روش غیر خطی عمل می‌کنند [۴] و این عامل باعث پیچیدگی فرآیند تغییر کاربری اراضی می‌شود [۵]. بنابراین به ابزارهای دقیق و پیشرفته برای مدلسازی این تغییرات نیاز می‌باشد. در زمینه مدلسازی تغییرات کلاس-های کاربری اراضی چندگانه (مثل تغییرات مربوط به کلاس‌های کاربری اراضی شهری و کشاورزی) تحقیقات محدودی انجام شده است و اغلب محققان به مدلسازی تغییرات کلاس کاربری اراضی یگانه (مثل تغییرات کلاس کاربری اراضی شهری) پرداخته‌اند [۲]. اما آنچه که در دنیای واقعی اتفاق می افتد مربوط به تغییرات کلاس‌های کاربری اراضی چندگانه است. مثلا در عین حال که کاربری‌های کشاورزی تبدیل به کاربری شهری می‌شوند ممکن است کاربری‌های مربوط به جنگل هم تبدیل به کاربری کشاورزی شده باشند. در این راستا استفاده از فنون داده کاوی برای کشف الگوهای پنهان در محرکان تغییرات کاربری اراضی می‌تواند جهت مدلسازی تغییرات کاربری اراضی مورد استفاده قرار گیرند.

اهدافی که در این مطالعه دنبال می‌گردد عبارتنداز: ۱) شبیه سازی روند تغییرات کاربری اراضی در دوره زمانی ۲۲ ساله، بین سالهای ۱۹۹۲ تا سال ۲۰۱۴ در بخشی از شمال ایران، ۲) انجام تحلیل حساسیت برای تعیین مهمترین فاکتورهای موثر در تغییرات کاربری اراضی چندگانه و نهایتا ۳) استفاده از شاخص عامل نسبی[۱] و شاخص عامل کلی[۲] که شاخص دوم توسط پونتیوس و سی در سال ۲۰۱۴ [۶] برای ارزیابی کارآیی مدلها ارائه شده و جهت مدلسازی تغییرات کاربری اراضی برای تغییرات کاربری اراضی چندگانه کمتر استفاده شده است.

۲- پیشینه تحقیق

اغلب مدلسازی‌هایی که در زمینه تغییرات کلاس‌های کاربری اراضی در سطح جهان انجام گرفته مربوط به تغییرات کلاس کاربری اراضی یگانه (غیر شهری به شهری و بالعکس) بوده و تحقیقات کمی به مدلسازی تغییرات کلاس‌های کاربری اراضی چندگانه پرداخته اند [۲]. به عنوان مثال در زمینه مدلسازی تغییرات کلاس‌های کاربری اراضی یگانه، می توان به مطرح شدن ایده کلی مدلسازی فیزیکی توسعه شهرها توسط پیجانویسکی و همکاران در سال ۱۹۹۷ [۷] و مدلسازی این پدیده توسط پیجانویسکی و همکاران در سال ۲۰۰۲ [۸] با استفاده از شبکه عصبی مصنوعی اشاره کرد. به همین ترتیب مدلسازی های زیادی تا به امروز در زمینه تغییرات کلاس‌های کاربری اراضی یگانه انجام شده است. در ایران نیز مدلسازی تغییرات کلاس‌های کاربری اراضی یگانه (توسعه شهری) انجام شده است [۹،۱۰،۱۱،۱۲].

اما در زمینه مدلسازی تغییرات کلاس‌های کاربری اراضی چندگانه همانگونه که اشاره شد تحقیقات محدودی انجام شده است. لی و یه در سال ۲۰۰۲ [۱۳] از تلفیقی از شبکه عصبی مصنوعی و اتوماتای سلولی برای مدلسازی تغییرات کلاس‌های کاربری اراضی چندگانه شامل گندم زارها، باغات، مکانهای توسعه، مناطق بشر ساخت، جنگل و

---

۱ Receiver Operating Characteristics
۲ Total Operating Characteristics

آب استفاده کردند. وربرگ و همکاران در سال ۲۰۰۲ [۱۴] از CLUE-S[1] برای مدلسازی تغییرات کلاس‌های کاربری اراضی چندگانه شامل جنگل، مرتع، مزارع نارگیل و مزارع برنج استفاده کردند. بالستروز و همکاران در سال ۲۰۱۲ [۱۵] از تلفیقی از درخت تصمیم گیری J48 و اتوماتای سلولی برای مدلسازی تغییرات کلاس‌های کاربری اراضی چندگانه شامل شهری، کشاورزی، جنگل، زمینهای خشک، تالاب و آب استفاده کردند. رالها و همکاران در سال ۲۰۱۳ [۱۶] از مدل عامل مبنا برای مدلسازی تغییرات کاربری های کشاورزی، شهری و جنگل استفاده کردند. طیبی و پیجانویسکی در سال ۲۰۱۴ [۲] سه فن داده کاوی شامل شبکه عصبی مصنوعی، درخت تصمیم گیری و رگرسیون و رگرسیون انطباقی چندمتغیره اسپلاین را برای مدلسازی تغییرات کلاس‌های کاربری اراضی شهری، کشاورزی و جنگل مورد استفاده قرار دادند.

همچنین مطالعات محدودی نیز وجود دارند که فنون مختلف داده کاوی را در حیطه مدلسازی تغییرات کلاس کاربری اراضی با هم مقایسه کرده‌اند. در زمینه مدلسازی تغییرات کلاس‌های کاربری اراضی یگانه طیبی و همکاران در سال ۲۰۱۴ [۱۷] سه فن درخت طبقه بندی و رگرسیون[2] ، رگرسیون انطباقی چند متغیره اسپلاین[3] و همچنین شبکه عصبی مصنوعی را برای مدلسازی تغییرات کلاس‌های کاربری اراضی یگانه (یعنی توسعه شهری، توسعه کشاورزی و توسعه جنگل) در سه ناحیه از مناطق مختلف جهان با هم مقایسه کردند. آنها نتیجه گرفتند که شبکه عصبی مصنوعی و رگرسیون انطباقی چند متغیره اسپلاین دقت بهتری نسبت به درخت طبقه بندی و رگرسیون دارند ولی دقت‌ها خیلی نزدیک به هم هستند و هر سه مدل از دقت قابل قبولی برای مدلسازی تغییرات کلاس‌های کاربری اراضی برخوردارند. در زمینه مدلسازی تغییرات کلاس‌های کاربری اراضی چندگانه، طیبی و پیجانویسکی در سال ۲۰۱۴ [۲] باز سه فن شبکه عصبی مصنوعی، درخت کلاسه بندی و رگرسیون و رگرسیون انطباقی چند متغیره اسپلاین را برای مدلسازی تغییرات کلاس‌های کاربری اراضی چندگانه (کلاسهای کاربری اراضی شهری، کشاورزی و جنگل) با هم مورد مقایسه قرار دادند. آنها به این نتیجه رسیدند که شبکه عصبی مصنوعی دارای دقت بهتری نسبت به دو مدل دیگر است ولی این بهبود دقت چندان هم قابل توجه نیست و هر سه مدل می‌توانند برای انجام مدلسازی تغییرات کاربری اراضی مورد استفاده قرار گیرند.

۳- محدوده و قلمرو تحقیق

ناحیه مورد مطالعه در این تحقیق بخشی از اراضی استان مازندران شامل شهرهای ساری (با طول جغرافیایی ۵۳ درجه و ۱ دقیقه و ۴۸ ثانیه شرقی و عرض جغرافیایی ۳۶ درجه و ۱۹ دقیقه و ۴۸ ثانیه شمالی) و قائم شهر (با طول جغرافیایی ۵۲ درجه و ۵۱ دقیقه و ۳۶ ثانیه شرقی و عرض جغرافیایی ۳۶ درجه و ۲۷ دقیقه و ۳۶ ثانیه شمالی) که مربوط به زون ۳۹ شمالی سیستم تصویر UTM هستند، می‌باشد. جمعیت دو شهر ساری و قائم شهر در مجموع بر اساس آمارهای رسمی کشور در سال ۱۳۸۵ برابر با ۴۵۶۲۱۲ نفر بود که در سال ۱۳۹۱ این رقم به ۵۰۳۲۶۷ نفر رسیده بود که این رشد جمعیت می‌تواند ناشی از افزایش مهاجرت از روستاها و همچنین شهرهای دیگر به این شهرها باشد.

۴- روش تحقیق

با توجه به پیچیدگی فرآیند مدلسازی تغییرات کاربری اراضی، استفاده کردن از ابزارهای غیر خطی همراه با سیستم اطلاعات مکانی می تواند مدلسازان تغییرات کاربری اراضی را در به تصویر کشیدن این فرآیند یاری کند. در واقع، فرآیند رایج در مدلسازی تغییرات کاربری اراضی به این صورت می‌باشد که بعد از انجام تصحیحات لازم روی تصاویر ماهواره‌ای، تمامی لایه‌های داده ای موثر در مدلسازی با استفاده از تحلیل های مکانی موجود در سیستم اطلاعات مکانی آماده می‌شوند. به عنوان مثال متغیرهای ورودی مدل‌ها، از قبیل فاصله از مناطق شهری، فاصله از مناطق کشاورزی، فاصله از جاده ها، فاصله از رودخانه، شیب و جهت با استفاده از تحلیل‌های مکانی در محیط سیستم اطلاعات مکانی آماده می‌شوند. همچنین ارزیابی دقت مدل حاصل شده نیز در محیط این سیستم انجام می‌شود.

---

۱ Conversion of Land Use and its Effects
۲ Classification And Regression Tree
۳ Multivariate Adaptive Regression Splines

رگرسیون انطباقی چند متغیره اسپلاین یک روش غیرپارامتریک محلی است که نسبت به سایر مدل‌ها دارای مزایای زیر می‌باشد: ۱) این مدل در مقایسه با مدلهای دیگر استفاده شده در زمینه مدلسازی تغییرات کاربری اراضی مانند رگرسیون منطقی، مدلی انعطاف پذیر می‌باشد زیرا ساختار مدل قبل از مدلسازی مشخص نیست و همچنین امکان تنظیم درجه تعامل بین متغیرهای ورودی وجود دارد (به عنوان مثال برای هر ترم مدل رگرسیون می توان از تعامل دو متغیر مستقل ورودی استفاده کرد و به جای برازش خط از سهمی استفاده نمود)، ۲) برخلاف دیگر مدل‌ها، مانند شبکه عصبی مصنوعی که ساختار مدل قابل تفسیر نیست (جعبه سیاه)، ساختار مدل رگرسیون انطباقی چند متغیره اسپلاین قابلیت تفسیری بالایی دارد و ۳) این مدل نیازمند در نظر گرفتن فرض آماری بین متغیرها نیست (به عنوان مثال نرمال بودن داده ها قبل از انجام عمل مدلسازی اهمیتی ندارد). با در نظر گرفتن موارد فوق در این مطالعه از یک فن داده کاوی غیر پارامتریک محلی، یعنی رگرسیون انطباقی چند متغیره اسپلاین برای شبیه‌سازی تغییرات کلاس‌های کاربری اراضی چندگانه در بخشی از نواحی شمال ایران شامل شهرهای ساری و قائم شهر (شکل ۱) در بین سالهای ۱۹۹۲ و ۲۰۱۴ استفاده شد. همچنین در عمده مطالعات قبلی ذکر شده، به بحث تحلیل حساسیت توجه نشده است. از اینرو، یک تحلیل حساسیت روی مجموعه‌ای از فاکتورهای تاثیر گذار روی مدلسازی تغییرات کاربری اراضی چندگانه در این تحقیق انجام شده است. فاصله از مناطق شهری، فاصله از مناطق کشاورزی، فاصله از جاده های اصلی، فاصله از رودخانه، ارتفاع، جهت و شیب به عنوان محرکان تغییرات کاربری اراضی درنظر گرفته شدند که بر اساس مطالعات قبلی انجام شده در ایران این فاکتورها بیشترین تاثیر را در روند تغییرات کاربری اراضی دارا می‌باشند [۱۹،۱۸]. برای انجام مدلسازی، داده‌ها با استفاده از روش نمونه برداری تصادفی طبقه‌بندی شده[۱] [۲۰] به دو دسته داده‌های آموزشی و داده‌های اعتبارسنجی تقسیم شدند. مشابه با سایر مدلسازی‌های انجام شده در تغییر کاربردی اراضی [۲] ۶۰٪ از داده‌ها به عنوان داده‌های آموزشی و ۴۰٪ از آنها به عنوان داده‌های اعتبارسنجی در نظر گرفته شد. تصاویر ماهواره‌ای لندست ۷ و ۸ با قدرت تفکیک ۳۰ متر (طبق بررسی‌های میدانی به عمل آمده اکثر تغییرات کلاس‌های کاربری اراضی در این تحقیق بیش از ۹۰۰ متر مربع می‌باشد و بنابراین این قدرت تفکیک مناسب برای مدلسازی تغییرات می‌باشد و از طرفی دسترس بودن آسان داده‌ها نیز در انتخاب این تصاویر نیز تاثیرگذار بود) مربوط به سالهای ۱۹۹۲ و ۲۰۱۴ استفاده شد. این تصاویر از طریق پایگاه داده USGS[۲] بصورت زمین مرجع اخذ شد که تمامی تصحیحات رادیومتریک و هندسی روی این تصاویر اعمال شده بودند. برای کلاسه‌بندی تصاویر از روش احتمال بیشینه شباهت [۲۱] در نرم افزار ENVI 4.7 استفاده شد که دقت کلی[۳] این کلاسه بندی برای تصویر سال ۱۹۹۲ برابر با ۹۵/۵٪ و برای تصویر سال ۲۰۱۴ برابر با ۹۳/۳٪ بدست آمد. همچنین شاخص کاپا[۴] برای دو تصویر به ترتیب برابر با ۹۰/۸٪ و ۹۲/۶٪ حاصل شد. با توجه به اینکه دقت کلاسه‌بندی تصاویر در دقت مدلسازی تغییرات کاربری اراضی اثر دارد و از آنجایی که دقت کلی کلاسه‌بندی برای دو تصویر سال ۱۹۹۲ و ۲۰۱۴ بالای ۹۰٪ می‌باشد، لذا تاثیر کلاسه‌بندی تصاویر بر روی دقت مدلسازی کمتر از ۱۰٪ می‌باشد [۲۲]. کلیه لایه‌های داده که به عنوان ورودی مدل مورد استفاده قرار گرفتند با استفاده از نرم افزار ArcGIS10 با فرمت شیپ فایل آماده شدند.

۴-۱- ایده‌های موجود برای انجام طبقه‌بندی چندگانه

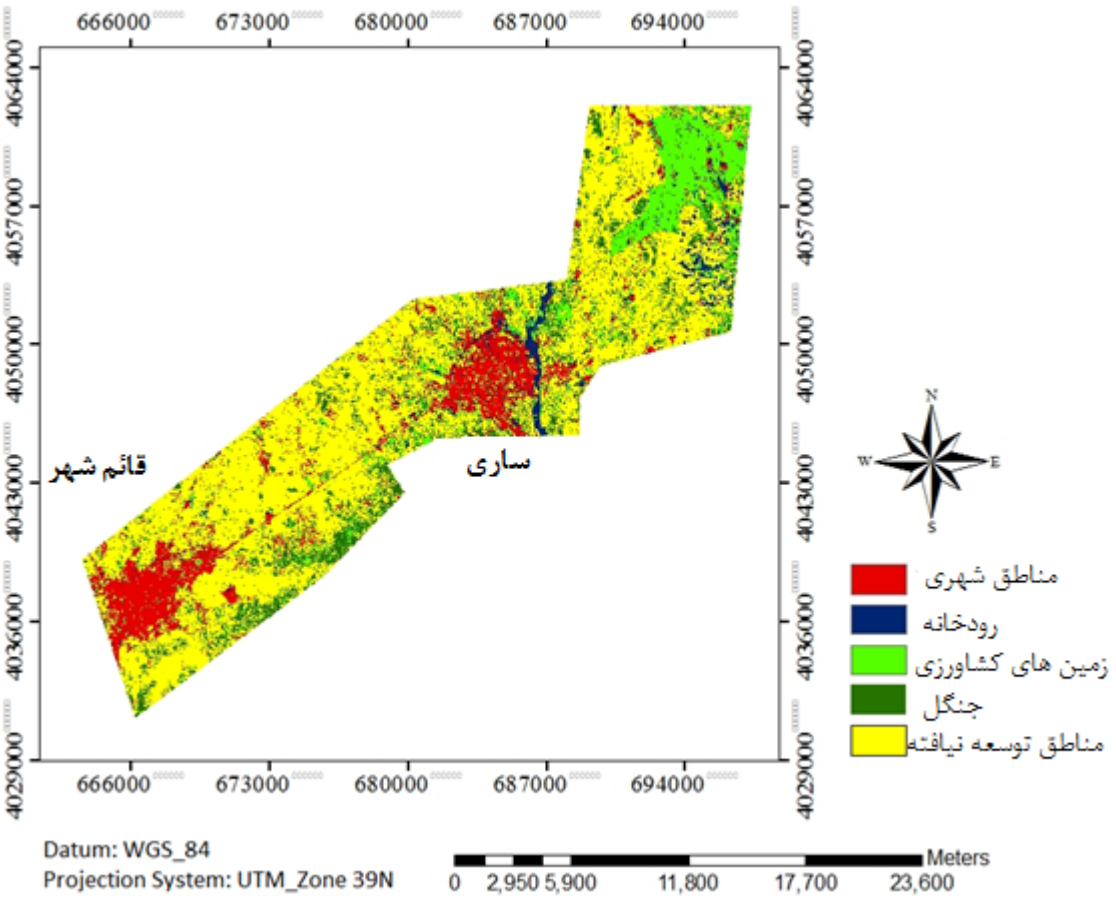
در حالت کلی برای انجام کلاسه‌بندی چندگانه دو ایده مطرح است. ایده اول این است که کلاسه‌بندی چندگانه به کلاسه‌بندی‌های دوگانه زیادی تجزیه شود و بعد هر کدام از این کلاسه‌بندی‌ها با استفاده از کلاسه‌بندی کننده‌های دوگانه انجام شود. ایده دوم این است که کلاسه‌بندی دوگانه به کلاسه‌بندی چندگانه تعمیم داده شود [۲]. برای ایده اول روش‌های مختلفی وجود دارد که در اینجا به صورت خلاصه تشریح شده‌اند (خوانندگان می‌توانند برای اطلاعات بیشتر به مرجع [۲۳] مراجعه کنند).

۱ Stratified random sampling

۲ United States Geological Survey

۳ Overall Accuracy: بیانگر یک شاخص کیفی کلان در کلاسه‌بندی است که بصورت یک مقدار متوسط در کل تصویر کلاسه‌بندی شده ارائه می‌شود و لزوما نشان دهنده دقت در تمامی پیکسل‌ها یا کلاس‌های اطلاعاتی نیست.

۴ Kappa Index



شکل ۱- منطقه مورد مطالعه در سال ۱۹۹۲

۱) One-Vs-All: در این روش که ساده ترین روش تجزیه کلاسه بندی چندگانه به کلاسه‌بندی دوگانه می‌باشد، در هر مرحله از مدلسازی یک کلاس با n-1 کلاس دیگر مقایسه می شود(n نشان دهنده تعداد کلاسها می‌باشد). بدین ترتیب نیاز به n کلاسه‌بندی کننده باینری، وجود دارد. در عمل با توجه به این که یک کلاس تعداد اعضای بسیار کمی نسبت به n-1 کلاس دیگر دارد، فرآیند یادگیری در این روش بسیار سخت می‌باشد [۲۴].

۲) All-Vs-All: در این روش به مقایسه دو به دوی تمامی کلاسها در کلاسه‌بندی چندگانه پرداخته می‌شود. بدین ترتیب به $\binom{n}{2}$ کلاسه‌بندی دوگانه نیاز می‌باشد. از ضعف‌های این روش می‌توان به تعداد زیاد طبقه‌بندی کننده‌های باینری اشاره کرد [۲۳].

۳) روش کدهای خروجی تصحیح کننده خطا(ECOC[۱]): این روش به عنوان یک ایده کلی برای تجزیه کلاس‌های چندگانه به کلاس‌بندی دوگانه مطرح شد [۲۵]. الگوریتم به این روش عمل می‌کند که ابتدا به هر کدام از کلاسها یک کد منحصر به فرد که طول آن باید بزرگتر یا مساوی با m باشد(برای مطالعه روشهای مختلف کدگذاری کلاسها به مرجع [۲۵] رجوع شود) اختصاص داده می‌شود. وقتی کدها به تمامی کلاسها اختصاص داده شد، ماتریسی حاصل می‌شود که آنرا ماتریس کد می‌نامند. ابعاد این ماتریس n*m می‌باشد که n نشان دهنده تعداد کلاسها و m نشان دهنده طول کد اختصاص داده شده به هر کدام از کلاسها می‌باشد. بعد از این مرحله در هر یک از ستونهای ماتریس کد، m طبقه‌بندی کننده‌ی دوگانه ایجاد شده که کلاسها را به مسائل دوگانه تبدیل می‌نمایند. بعد از این مرحله طبقه‌بندی کننده‌ها شروع به یادگیری می‌کنند. بعد از یادگیری اگر یک کد به طول m به طبقه‌بندی کننده‌ها تخصیص داده شود هر کدام یک خروجی ارائه می‌دهند که در آن، کلاسی به عنوان کلاس کد ورودی در نظر گرفته می‌شود که کمترین فاصله را با هر کدام از کدهای مربوط به کلاسها داشته باشد.

۱ Error-Correcting Output-Coding

برای ایده دوم (تعمیم کلاسه‌بندی دوگانه به کلاسه-بندی چندگانه) می‌توان از طرح کدگذاری داده کاوی استفاده نمود [۲] برای انجام این کار دو روش زیر وجود دارد:

۱-کدگذاری One-per-Class: در این حالت می‌توان از n خروجی باینری به جای یک خروجی باینری استفاده کرد که هر کدام از خروجی‌های دوگانه برای جداسازی داده‌های مربوط به یک کلاس به کار گرفته می‌شود. به عنوان مثال در شبکه عصبی مصنوعی برای مدلسازی تغییرات کاربری چندگانه می‌توان به جای داشتن یک نقطه گرهی[۱] خروجی دوگانه که نشان دهنده تغییر یا عدم تغییر کلاس کاربری داده شده می‌باشد، از تعداد n نقطه گرهی خروجی دوگانه متناسب با n کلاس کاربری استفاده کرد.

۲-کدگذاری توزیع یافته: در این نوع کدگذاری برای هر کدام از کلاسها یک کد منحصر به فرد از 1 تا n اختصاص داده می‌شود. بعد از فرآیند یادگیری، در مرحله آزمایش با استفاده از مدل مورد نظر، کد محاسبه شده برای یک نمونه آزمایشی با هر کدام از کدهای مربوط به کلاسها مقایسه شده و کدی به عنوان کلاس نمونه انتخاب می‌شود که کمترین فاصله را از آن داشته باشد.

در این مطالعه از کدگذاری توزیع یافته از روش‌های مربوط به ایده دوم برای کدگذاری کلاسهای کاربری اراضی استفاده گردید. برای هر کدام از کلاسهای کاربری شهری و کاربری کشاورزی به ترتیب کدهای ۱ و ۲ اختصاص داده شد. در بخش‌های زیر به تشریح مدل رگرسیون انطباقی چندمتغیره اسپلاین و شاخص‌های اهمیت متغیرها و همچنین شاخص‌های ارزیابی مدلها پرداخته شده است.

۴-۲- رگرسیون انطباقی چندمتغیره اسپلاین

در یک تقسیم‌بندی کلی می‌توان فنون داده کاوی را به دو دسته کلی مدلهای غیر پارامتریک محلی و مدلهای پارامتریک کلی تقسیم کرد [۱۷]. ترم پارامتریک به این معنی می‌باشد که ساختار مدل قبل از اینکه عمل مدلسازی انجام شود مشخص و معلوم هست، مثال‌های مدل فوق شبکه عصبی مصنوعی و رگرسیون منطقی هستند. در مقابل اگر ساختار مدل قبل از مدلسازی معلوم نباشد مدلهای غیر پارامتریک مطرح می‌شود که از این دسته از مدلها می‌توان به رگرسیون انطباقی چند متغیره اسپلاین اشاره کرد. همچنین اگر مدل استفاده شده از تمام داده های مربوطه به صورت یکجا برای مدلسازی استفاده کند مدل را کلی و اگر داده ها را به زیر مجموعه هایی تقسیم کرده و بعد عمل مدلسازی را برای هر کدام از این زیر مجموعه ها انجام دهد مدل را محلی می‌نامند [۱۷].

رگرسیون انطباقی چندمتغیره اسپلاین یکی از مدلهای غیر پارامتریک محلی هست که برای اولین بار توسط فردمن در سال ۱۹۹۱ [۲۷] ارائه شد. با توجه به اینکه تغییرات کاربری اراضی در یک روش غیر خطی عمل می‌کند، این روش می‌تواند بدون در نظر گرفتن هیچ فرض آماری بین متغیرهای پیش بینی کاربری اراضی (همان محرکان تغییرات) و متغیر هدف (که همان کلاس کاربری اختصاص داده شده به هر کدام از سلولها است) اقدام به مدلسازی روابط غیر خطی بین این متغیرها نماید [۲۶]. در قیاس با سایر مدل‌های غیر خطی که فقط یک مجموعه از ضرایب را به داده‌ها اعمال می‌کنند این روش با برازش دادن توابع چندجمله‌ای منطقه‌ای جداگانه برای هر یک از زیرمجموعه از داده‌ها اقدام به تشخیص الگوهای پیچیده می‌نماید [۱۷] در واقع این روش داده‌ها را به زیر مجموعه‌هایی تقسیم کرده و متناسب با پیچیدگی داده‌ها در هر ناحیه اقدام به برازش توابعی بنام توابع پایه می‌نماید. این روش می‌تواند با یک فرآیند دو مرحله ای اقدام به مدلسازی تغییرات کاربری اراضی نماید [۲۶]. مرحله اول، مرحله حرکت به سمت جلو می‌باشد. در این مرحله مدل ابتدا فقط با استفاده از یک ترم ثابت شروع می‌شود(که در واقع میانگین مقادیر متغیر هدف یا همان کلاس کاربری اراضی مربوط به پیکسل‌ها است) و بعد به صورت مکرر توابع پایه را برای مدل کردن داده‌ها به ترم ثابت اضافه می‌کند. در مرحله اضافه کردن توابع پایه، مدل توابع پایه‌ای را به معادله اضافه می‌کند که بیشترین کاهش را در مجموع مربعات خطاهای باقی مانده ایجاد کنند. حرکت رو به سمت جلو اغلب باعث ایجاد بیش برازش[۲] به داده‌ها می‌شود. در حقیقت مدلی ایجاد می‌شود که دارای بیشترین برازش به داده‌هایی است که در فرآیند مدلسازی دخیل بوده‌اند اما برای داده‌هایی که در فرآیند

۱ Node

۲ Over fitting

مدلسازی شرکت نکرده باشند برازش خوبی ندارد. برای ساختن مدلی که قابلیت انعطاف بیشتر برای همه داده ها باشد نیاز به مرحله دوم یعنی مرحله حرکت به سمت عقب می‌باشد. در واقع مرحله حرکت به سمت عقب برای هرس کردن مدل و حذف توابع پایه‌ای که کمترین اثر را در فرآیند مدلسازی دارند می‌باشد. برای مشخص کردن زیر مدلهایی که کمترین اثر را در مدلسازی دارند از اعتبارسنجی متقاطع تعمیم‌یافته[۱] استفاده می‌شود.

اعتبارسنجی متقاطع تعمیم‌یافته تعاملی بین برازش و پیچیدگی مدل است به این صورت که در مرحله حرکت به سمت جلو با اضافه شدن توابع پایه، برازش مدل به داده‌ها افزایش پیدا کرده و به تبع آن پیچیدگی مدل نیز بیشتر می‌شود. ولی در مرحله حرکت به سمت عقب، با حذف توابع پایه‌ای که کمترین سهم را در مدلسازی دارند میزان برازش کمتر شده ولی پیچیدگی مدل کاهش پیدا می‌کند. رگرسیون انطباقی چندمتغیره اسپلاین طبق معادله (۱) انجام می‌شود [۲۶]:

$$\hat{Y} = \hat{f}(X) = a_0 + \sum_{m=1}^{M} a_m B_m(X) \qquad (۱)$$

در این معادله a_0 یک مقدار ثابت می‌باشد، M تعداد ترم‌های غیر صفر (یا همان گره‌ها که توابع پایه در آنها تقسیم می‌شوند) می‌باشد، a_m همان ضرایب مربوط به mامین تابع پایه و $B_m(X)$ ، mامین تابع پایه برای مدل است که طبق معادله (۲) محاسبه می‌شود:

$$B_m(X) = \prod_{i=1}^{K_m} [S_{i,m}(X_{v(i,m)} - t_{i,m})]_+^q \qquad (۲)$$

K_m درجه تعامل بین متغیرها در mامین تابع پایه می‌باشد. $S_{i,m} = \pm 1$ و $X_{v(i,m)}$ متغیر vامی می‌باشد که در آن $1 \le v(i,m) \le k$ و k تعداد کل متغیرهای ورودی می‌باشد. $t_{i,m}$ مکان گره، در هر یک از متغیرهای پیش‌بینی متغیر وابسته می‌باشد، q توان توابع پایه می‌باشد، به عنوان مثال برای $q=1$، یک اسپلاین خطی ساده به داده‌ها اعمال می‌شود. اندیس + به معنی اخذ بخش مثبت عبارت داخل براکت می‌باشد به این معنی که اگر عبارت داخل براکت بزرگتر از صفر بود نتیجه خود عبارت داخل براکت می‌شود در غیر این صورت نتیجه برابر با صفر خواهد بود یعنی [۲۶]:

$$[S_{i,m}(X_{v(i,m)} - t_{i,m})]_+^q = \begin{cases} [S_{i,m}(X_{v(i,m)} - t_{i,m})]^q & S_{i,m}(X_{v(i,m)} - t_{i,m}) > 0 \\ 0 & \text{otherwise} \end{cases} \qquad (۳)$$

اعتبارسنجی متقاطع تعمیم‌یافته که بهترین مدل رگرسیون بر اساس آن انتخاب می‌شود نیز طبق رابطه (۴) حاصل می‌شود[۲۶]:

$$\frac{(1/n)\sum_{i=1}^{n}[y_i - \hat{f}(X_i)]^2}{[1-(C(M)/n)]^2} \qquad (۴)$$

که در این معادله y_i مقادیر واقعی کلاس مورد نظر، $\hat{f}(X_i)$ مقدار تخمین زده شده برای مقادیر واقعی کلاس مورد نظر، n تعداد کل مشاهدات و $C(M)$ معیار هزینه-جریمه یک مدل است که شامل M تابع پایه می باشد. به عبارت دیگر $C(M)$ تعداد موثر درجه آزادی است که به موجب آن اعتبارسنجی متقاطع تعمیم یافته یک جریمه را برای اضافه‌کردن متغیرهای ورودی بیشتر به مدل اضافه می‌کند. $C(M)$ از رابطه (۵) محاسبه می‌شود [۱۷]:

$$C(M) = M \times d \qquad (۵)$$

d هزینه برای هر کدام از توابع می‌باشد. در واقع صورت کسر در معادله (۴)، عدم برازش در M تابع برازش به داده‌ها را نشان می‌دهد و مخرج کسر جریمه برای پیچیدگی مدل را نشان می‌دهد[۲۶].

فرآیند مدلسازی در نرم افزار Salford Predictive Modeler [۱۷] انجام شد. بقیه مراحل کار در نرم افزار متلب پیاده شد. با توجه به پیچیدگی منطقه از درجه تعامل ۲ برای انجام مدلسازی استفاده گردید. یعنی در معادله (۲) $q=2$ انتخاب گردید و به ترمهایی از نوع سهمی اجازه وارد شدن در معادله رگرسیون داده شد. ضرائب، متغیرها و گره‌ها در مدل رگرسیون انطباقی چند متغیره اسپلاین تشکیل شده مطابق با جدول (۱) می‌باشد. طبق این جدول معادله رگرسیون کلی دارای ۱۹ ترم می‌باشد که ترم صفر نشان دهنده مقدار ثابت معادله رگرسیون است. تعامل بین متغیرهای ورودی که مقدار دو برای آن انتخاب شده بود را می-توان در توابع پایه ۳، ۴، ۷، ۸، ۹، ۱۰، ۱۱، ۱۴، ۱۵، ۱۶، ۱۷ و ۱۸ مطابق با جدول (۱) مشاهده کرد.

[۱] Generalized Cross-Validation

جدول ۱- ضرایب، متغیرها و گره ها در مدل رگرسیون انطباقی چند متغیره اسپلاین

توابع پایه	ضرایب	متغیرهای تقسیم شونده	علامت متغیرهای تقسیم شونده	علامت متغیر والد	متغیر والد	گره(متر)
۰	۰/۴۸۹۵۷۳					
۱	۰/۰۰۰۷۷۶	فاصله از مناطق کشاورزی	+			۱۳۴/۱۶۴۱
۲	-۰/۰۰۰۶۴۴	فاصله از مناطق کشاورزی	-			۱۳۴/۱۶۴۱
۳	۰/۰۰۰۰۰۱	فاصله از رودخانه	+	+	فاصله از مناطق کشاورزی	۹۵۱۲/۰۳۵۲
۴	۰/۰۰۰۰۰۱	فاصله از رودخانه	-	+	فاصله از مناطق کشاورزی	۹۵۱۲/۰۳۵۲
۵	۰/۰۰۵۵۸۲	ارتفاع	+			-۸/۰۰۰۰
۶	۰/۰۵۲۶۱۰	ارتفاع	-			-۸/۰۰۰۰
۷	-۰/۰۰۰۱۰۰	فاصله از مناطق شهری	+	+	ارتفاع	۶۰/۰۰۰۲
۸	-۰/۰۰۰۲۷۴	فاصله از مناطق شهری	-	+	ارتفاع	۶۰/۰۰۰۲
۹	۰/۰۰۰۰۹۶	فاصله از مناطق شهری	+	+	ارتفاع	۱۵۰/۰۰۰۰
۱۰	۰/۰۰۰۰۱۰	فاصله از مناطق کشاورزی	-	+	ارتفاع	۱۲۰/۰۰۰۰
۱۱	۰/۰۰۰۰۲۹	فاصله از مناطق کشاورزی	+	+	ارتفاع	۱۲۰/۰۰۰۰
۱۲	۰/۰۰۰۰۷۴	فاصله از مناطق شهری	-			۷۲۶/۲۲۳۱
۱۳	۰/۰۰۰۲۴۷	فاصله از مناطق شهری	+			۷۲۶/۲۲۳۱
۱۴	۰/۰۰۰۰۰۱	فاصله از مناطق کشاورزی	-	-	فاصله از مناطق شهری	۳۰/۰۰۰۰
۱۵	-۰/۰۰۰۰۰۲	فاصله از مناطق کشاورزی	+	+	فاصله از مناطق شهری	۳۰/۰۰۰۰
۱۶	۰/۰۰۰۰۰۱	فاصله از رودخانه	-	+	فاصله از مناطق شهری	۱۶۴۱۳/۵۳۷۱
۱۷	۰/۰۰۰۰۰۱	فاصله از رودخانه	+	+	فاصله از مناطق شهری	۱۶۴۱۳/۵۳۷۱
۱۸	۰/۰۰۰۰۰۱	فاصله از رودخانه	-	+	فاصله از مناطق شهری	۱۱۱۳۰/۰۴۰۰

مقادیر هر کدام از متغیرها در این جدول در یک مقدار ثابت تقسیم می‌شوند که این مقدار همان گره برای آن متغیر است. به عنوان مثال تابع پایه یک، دارای یک متغیر، یعنی فاصله از مناطق کشاورزی (X_1)، می‌باشد که این متغیر در مقدار ۱۳۴/۱۶۴۱ متر تقسیم می‌شود و علامت خود متغیر مثبت بوده و ضریب این ترم ۰/۰۰۰۷۷۶+ است. تابع پایه دوم نیز دارای یک متغیر (فاصله از مناطق کشاورزی) بوده و در همان مقدار ۱۳۴/۱۶۴۱ متر تقسیم شده ولی علامت متغیر برخلاف ترم یک منفی بوده و دارای ضریب ۰/۰۰۰۰۶۴۴- می‌باشد. در تابع پایه سوم همانطور که قبلا اشاره شد دو متغیر، یعنی فاصله از رودخانه (X_2) و فاصله از مناطق کشاورزی) حضور دارند (درجه تعامل ۲) که فاصله از رودخانه در مقدار ۹۵۱۲/۰۳۵۲ متر تقسیم شده و دارای علامت مثبت می‌باشد و متغیر فاصله از مناطق کشاورزی به عنوان متغیر والد در نظر گرفته شده که دارای علامت مثبت بوده و در متغیر فاصله از رودخانه ضرب می‌شود. ضریب این دو متغیر ۰/۰۰۰۰۰۱+ می‌باشد. طبق توضیحات داده شده معادله رگرسیون انطباقی چند متغیره اسپلاین متشکل از ترم‌های ۰، ۱، ۲ و ۳ مطابق با معادله (۶) می‌باشد:

$$Y = 0.489573 + 0.000776 \times (X1 - 134.1641) - 0.000644 \times (134.1641 - X1) + 0.000001 \times (X2 - 9512.0352) \times X1 \quad (۶)$$

در این معادله متغیر Y همان کلاس‌های کاربری اراضی (متغیر هدف) است که با توجه به کدگذاری توزیع یافته کلاس‌های کاربری اراضی (۰ برای عدم تغییر کلاس کاربری اراضی، ۱ برای کلاس کاربری اراضی شهری و ۲ برای کلاس کاربری اراضی کشاورزی) مقادیری بین ۰ تا ۲ را می‌گیرد که مقادیر نزدیک به ۰ نشان دهنده تمایل سلولها برای عدم تغییر کاربری اراضی و مقادیر نزدیک به ۱ نشان دهنده تمایل سلولها برای تغییر به کلاس کاربری اراضی شهری و مقادیر نزدیک به ۲ نشان دهنده تمایل سلولها برای تغییر به کلاس کاربری اراضی کشاورزی

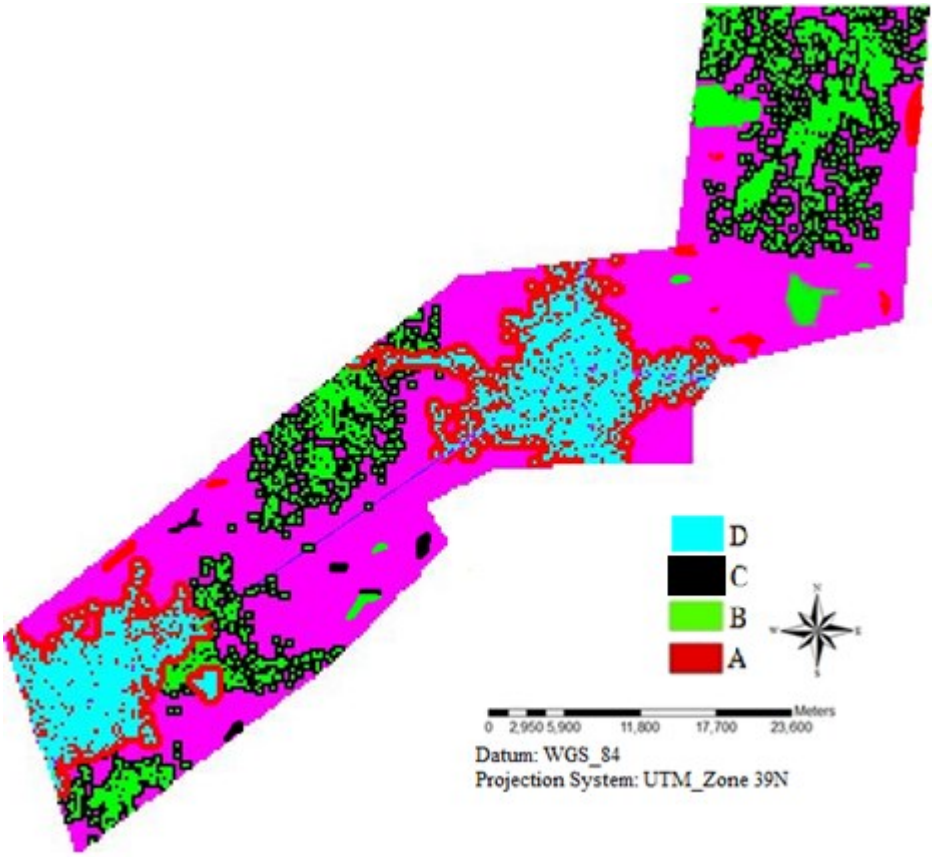
می‌باشند. در سمت راست معادله نیز متغیرهای مستقلی وجود دارند که برای مدلسازی متغیر وابسته Y به کار رفته‌اند. به همین ترتیب می‌توان معادله کلی شامل همه ۱۹ تابع پایه موجود در جدول (۱) را تشکیل داد. هدف اصلی در این مطالعه شبیه‌سازی تغییرات کاربری اراضی چندگانه برای یک دوره ۲۲ ساله و استخراج مهمترین فاکتورهای موثر در این تغییرات بوده است، اما می‌توان مدل بدست آمده برای این بازه زمانی را به متغیرهای مستقل موثر در تغییرات کاربری اراضی در سال ۲۰۱۴ اعمال کرده و نقشه تغییرات کاربری اراضی را برای سال ۲۰۳۶ پیش بینی کرد (شکل (۴)). با این حال ممکن است تاثیر متغیرهای مستقل در نظر گرفته شده در تغییرات کاربری اراضی در دو بازه زمانی یکسان نباشد. همانطور که در شکل (۴) مشاهده می‌شود اکثر سلول‌های توسعه یافته شهری و کشاورزی به ترتیب در نزدیکی مناطق شهری و زمین‌های کشاورزی رخ می‌دهند.

۴-۳- روش شناسایی محرکان کاربری اراضی پر اهمیت در فرآیند مدلسازی

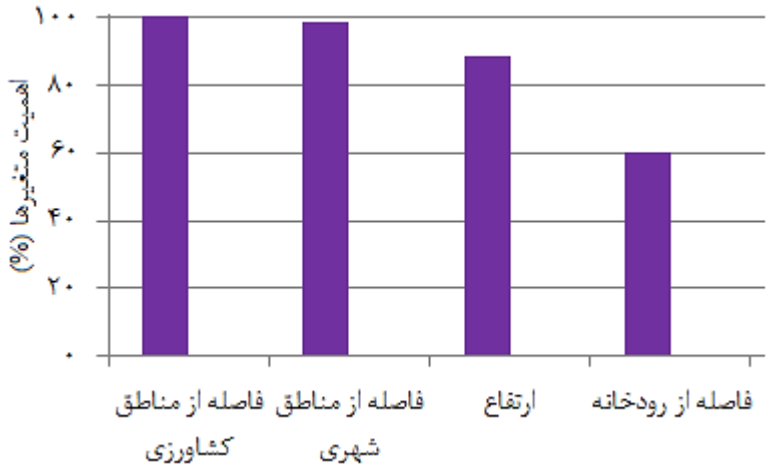
در این مطالعه، برای تشخیص محرکان تغییرات کاربری اراضی پر اهمیت، یک تحلیل حساسیت انجام شد. زمانی که مدل رگرسیون انطباقی چند متغیره اسپلاین برای مجموعه داده‌های آموزشی تشکیل شد با حذف تک تک پارامترهای مربوط به هر کدام از محرکان از معادله مدل تشکیل شده، مقدار تغییری که در برازش مدل (اعتبار سنجی متقاطع تعمیم یافته) رخ می‌دهد محاسبه می‌گردد. محرکی که دارای کمترین اثر در برازش هنگام حذف از مدل باشد به عنوان کم اهمیت‌ترین محرک انتخاب می‌شود. به همین ترتیب محرکی که بعد از حذف شدن بیشترین اثر را در برازش مدل داشته باشد به عنوان پر اهمیت ترین محرک شناخته می‌شود. در شکل (۳) اهمیت متغیرها ارائه شده است. همچنین انحراف معیار هر کدام از محرکان تغییرات کاربری اراضی در شکل (۴) ارائه شده است. هر چه قدر انحراف معیار مربوط به یک محرک بیشتر باشد نشان دهنده قدرت توضیحی و تبیینی بالای آن متغیر در توصیف ارتباط بین ورودی‌ها و خروجی‌ها می‌باشد [۱۷]. همانطور که در نمودار مربوط به شکل (۴) مشاهده می‌شود ترم متشکل از دو فاکتور فاصله از مناطق شهری و ارتفاع با انحراف معیار ۰/۲۷ (معیار همان عدد مربوط به کلاس کاربری متشکل از دو فاکتور فاصله از مناطق شهری و ارتفاع با انحراف معیار ۰/۲۷ (معیار همان عدد مربوط به کلاس کاربری اراضی می‌باشد) دارای بیشترین انحراف معیار و ترم متشکل از دو فاکتور ارتفاع و فاصله از مناطق کشاورزی با انحراف معیار ۰/۰۵، دارای کمترین انحراف معیار می‌باشند و به ترتیب دارای بیشترین و کمترین قدرت توضیحی و تبیینی برای توصیف ارتباط بین خروجی‌ها و ورودی‌ها می‌باشند [۲]. تعیین درجه اهمیت متغیرها و انحراف معیار آنها در نرم افزار Salford Predictive Modeler انجام شده است. در این نرم افزار درجه اهمیت متغیرها بر حسب درصد ارائه شده است.

۴-۴- ارزیابی نتایج بدست آمده

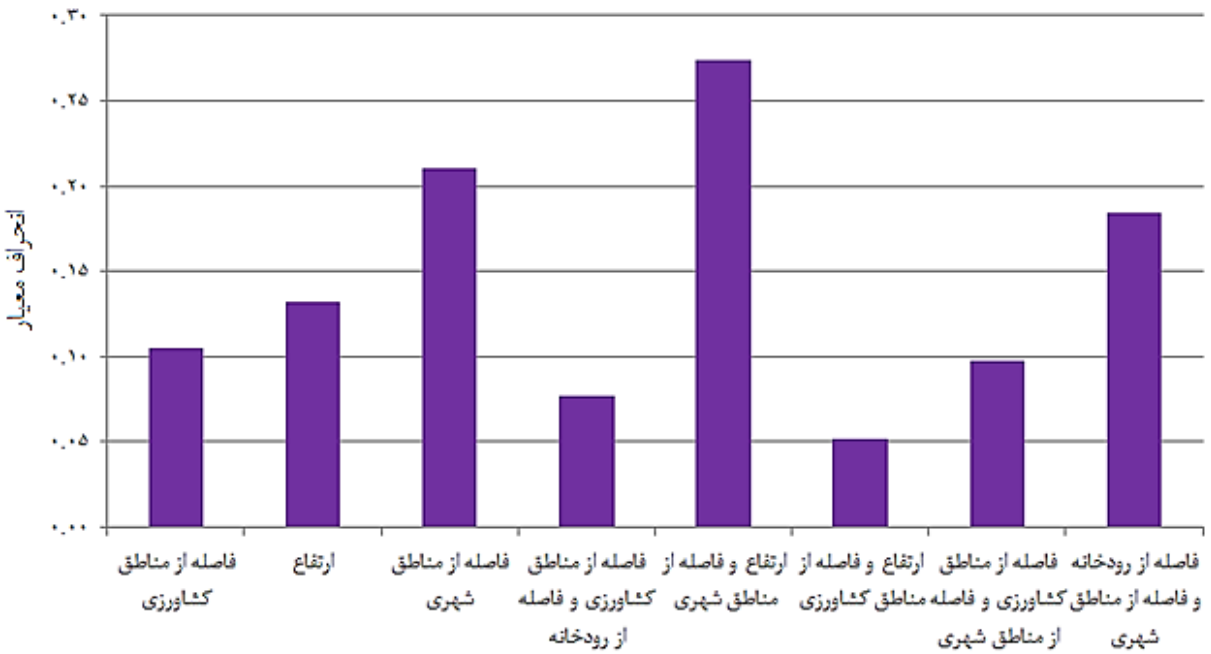
شاخص‌های کالیبراسیون برای ارزیابی دقت و کارآیی مدل به کار رفته، استفاده می‌شوند. خروجی یک مدل شبیه‌سازی تغییرات کاربری اراضی، یک نقشه‌تناسب می‌باشد که احتمال تغییرکاربری را برای پیکسل‌های مختلف با توجه به تغییرات کاربری اراضی نشان می‌دهد. اما آنچه که باید برای ارزیابی کارآیی مدل، با نقشه تغییرات واقعی (نقشه مرجع که از مقایسه تصاویر در دو بازه زمانی متوالی بدست می‌آید و نقش متغیر هدف را در مدلسازی ایفا می‌کند) مقایسه‌شود، نقشه شبیه‌سازی شده می‌باشد که با اعمال حد آستانه به نقشه تناسب حاصل می‌شود. پیداکردن یک حد آستانه مناسب برای نقشه تناسب با توجه به ویژگی‌های منطقه مورد مطالعه و ساختار مدل کار دشواری است [۶]. در این مقاله از شاخص عامل نسبی و شاخص عامل کلی برای ارزیابی کارآیی مدل رگرسیون انطباقی چند متغیره اسپلاین استفاده شده است که هر دوی آنها از حدود آستانه مختلف برای انجام فرآیند ارزیابی مدل استفاده می‌کنند.



شکل ۲- نقشه تغییرات کاربری اراضی پیش‌بینی شده برای سال ۲۰۳۶ با استفاده از مدل بدست آمده برای بازه زمانی ۱۹۹۲ تا ۲۰۱۴. D نشان دهنده مناطق شهری موجود در سال ۲۰۱۴ و A سلول‌های شهری توسعه یافته در سال ۲۰۳۶ می‌باشند. به همین ترتیب B مناطق کشاورزی موجود در سال ۲۰۱۴ و C سلول‌های توسعه یافته کشاورزی در سال ۲۰۳۶ می‌باشند.



شکل ۳- درجه اهمیت محرکان تغییرات کاربری اراضی



شکل ۴- انحراف معیارهای هر کدام از محرکان تغییرات کاربری اراضی

جدول ۳- جدول احتمال استفاده شده برای مقایسه نقشه شبیه-سازی شده و نقشه موجود در واقعیت

نقشه شبیه سازی شده	نقشه موجود در واقعیت		
	تغییر	عدم تغییر	کل
تغییر	A	B	A + B
عدم تغییر	C	D	C + D
کل	A + C = P	B + D = Q	

۴-۴-۱- شاخص عامل نسبی

شاخص عامل نسبی با به کارگیری حدود آستانه مختلف، محدودیت تعریف یک حد آستانه منحصر به فرد را برای مسئله از بین می‌برد [۲۸]. بعد از تعریف حدود آستانه، این مقادیر به نقشه تناسب اعمال شده و مقادیری که در نقشه تناسب بزرگتر از این حد آستانه باشند تبدیل به ۱ (به معنی تغییر کاربری اراضی در پیکسل مورد نظر) و دیگر مقادیر کمتر از حدود آستانه برابر با صفر(به معنی عدم تغییر کاربری اراضی در پیکسل مورد نظر) می‌شوند. نتیجه حاصل با نقشه تغییرات کاربری اراضی در حالت واقعی مقایسه شده و جداول احتمال ۲*۲ مطابق جدول (۳) و برای هر کدام از حدود آستانه محاسبه می‌شود. بعد از ایجاد جداول احتمال، برای رسم منحنی شاخص عامل نسبی، X_t (حساسیت) و Y_t (۱-خصوصیت)را برای حدود آستانه مختلف طبق روابط (۷) و (۸) محاسبه می‌گردد [۲۷]:

$$X_t = \frac{B}{B+D} = \frac{B}{Q} = 1 - \text{خصوصیت} \qquad (۷)$$

$$Y_t = \frac{A}{A+C} = \frac{A}{P} = \text{حساسیت} \qquad (۸)$$

که حساسیت نشان دهنده احتمال طبقه‌بندی صحیح پیکسل‌های تغییر یافته توسط مدل و خصوصیت نشان دهنده احتمال طبقه‌بندی صحیح پیکسل‌های تغییر نیافته است. مساحت زیر منحنی شاخص عامل نسبی نیز طبق رابطه (۹) حاصل می‌شود [۲۹]:

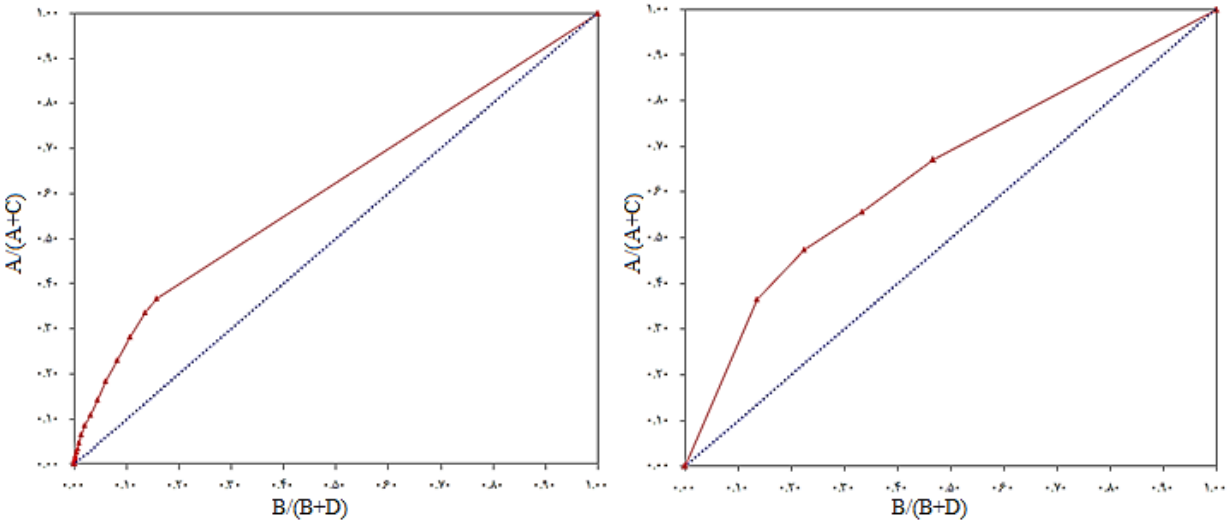
$$\sum_{t=1}^{T-1} [Y_{t+1} + Y_t][X_{t+1} - X_t]/2 \qquad (۹)$$

مساحت زیر منحنی شاخص عامل نسبی برای کاربری اراضی کشاورزی و شهری با استفاده از رابطه (۹) به ترتیب برابر با ۶۱/۰۱٪ و ۶۵٪ بدست آمد. در شکل (۵) نمودار شاخص عامل نسبی برای کاربری‌های کشاورزی و شهری ارائه شده است.

۴-۴-۲- شاخص عامل کلی

با توجه به محدودیت‌هایی که شاخص عامل نسبی در نمایش دادن اعداد جداول احتمال برای هر حد آستانه داشت، پونتیوس و سی در سال ۲۰۱۴ [۶] شاخص عامل کلی را برای غلبه بر این محدودیت به کار بردند. شاخص

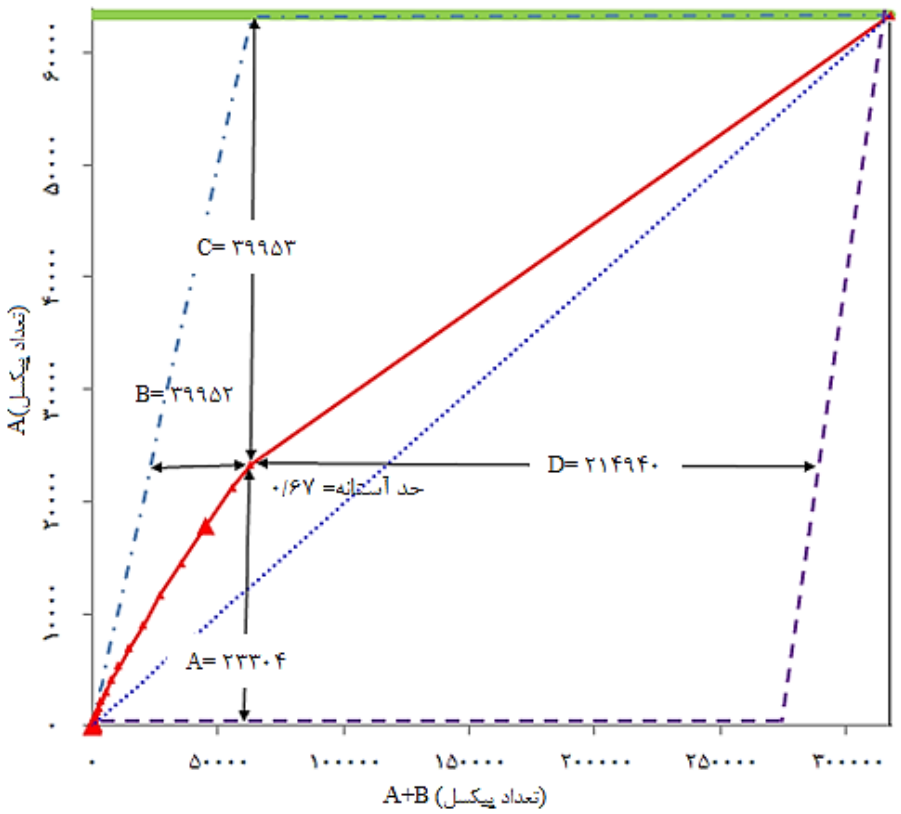
عامل کلی علاوه بر نشان دادن دقت ارائه شده بوسیله شاخص عامل نسبی، چهار عدد موجود در جداول احتمال یعنی A، B، C و D (جدول (۳)) را برای هر کدام از حدود آستانه نشان می‌دهد. محورهای x و y در نمودار شاخص عامل کلی به ترتیب برابر با A + B و A می‌باشند. همانطور که در شکلهای (۶) و (۷) که نمودار شاخص عامل کلی، برای کاربری‌های اراضی کشاورزی و شهری را نمایش می‌دهند مشاهده می‌شود، برای رسم نمودار شاخص عامل کلی، ابتدا یک متوازی‌الاضلاع که اضلاع آن برابر با (۰,۰)، (P,P)، (Q,۰) و (P+Q,P) می‌باشد رسم می‌گردد (مطابق با جدول (۳)، P برابر با A + C و Q برابر با B + D می‌باشد) و سپس نمودار منحنی با استفاده از x و y های محاسبه‌شده از جداول احتمال رسم می‌گردد. برای هر حد آستانه میتوان چهار عدد جدول احتمال را در شکل‌های (۶) و (۷) مشاهده‌کرد. به عنوان مثال برای حد آستانه ۰/۶۷، مقادیر A، B، C و D (همان مقادیر جدول (۳)) در نمودار شاخص عامل کلی برای کاربری اراضی کشاورزی (شکل (۶)) به ترتیب برابر با ۲۳۳۰۴، ۳۹۹۵۲، ۳۹۹۵۳ و ۲۱۴۹۴۰ پیکسل می‌باشد. همچنین برای حد آستانه ۰/۴، مقادیر A ، B، C و D در نمودار شاخص عامل کلی برای کاربری اراضی شهری (شکل (۷)) به ترتیب برابر با ۲۲۱۷۷، ۱۳۲۸۲۰، ۱۰۸۹۱ و ۱۵۲۲۶۱ می‌باشد. همچنین مساحت زیر منحنی شاخص عامل نسبی (همان اعداد ۶۱/۰۱٪ و ۶۵٪ به ترتیب برای کاربری‌های اراضی کشاورزی و شهری) برابر با، نسبت مساحت نمودار شاخص عامل کلی داخل متوازی‌الاضلاع به مساحت کل متوازی‌الاضلاع بدست آمد. نقشه خطا که در واقع حاصل هم پوشانی نقشه پیش‌بینی شده توسط مدل و نقشه تغییرات واقعی اتفاق افتاده می‌باشد، به ترتیب برای کاربری اراضی کشاورزی و شهری در شکل (۸) ارائه شده‌اند. در این نقشه کد A نشان دهنده سلولهایی است که به عنوان سلول تغییر یافته توسط مدل پیش‌بینی شده‌اند و در واقعیت نیز به عنوان سلول تغییر یافته می‌باشند. کد B نشان دهنده سلول‌هایی است که به عنوان سلول تغییر یافته توسط مدل پیش‌بینی شده‌اند ولی در واقعیت این سلول‌ها تغییری نکرده‌اند. کد C نشان دهنده سلول هایی است که توسط مدل به عنوان سلول تغییر نیافته پیش‌بینی شده‌اند. ولی در واقعیت این سلول‌ها دچار تغییر شده‌اند. سرانجام D نشان دهنده سلول‌هایی است که در نقشه پیش‌بینی شده توسط مدل به عنوان سلول تغییر نیافته هستند و در واقعیت نیز تغییری نکرده‌اند.



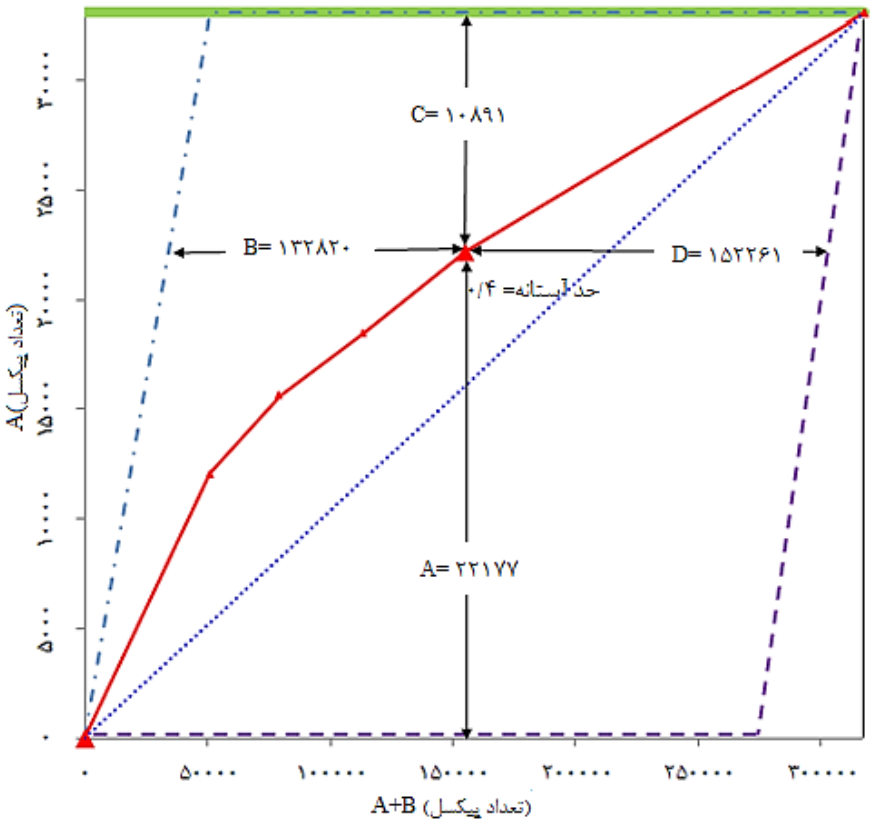
الف: شاخص عامل نسبی برای کاربری اراضی شهری

ب: شاخص عامل نسبی برای کاربری اراضی کشاورزی

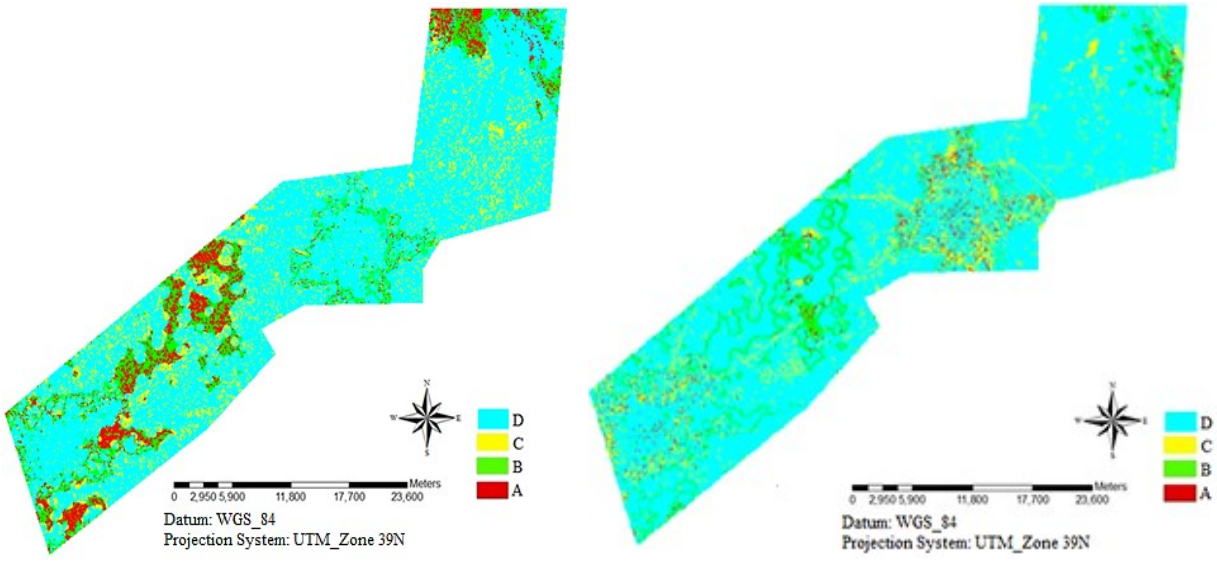
شکل ۵- شاخص عامل نسبی برای کاربری اراضی شهری (الف) و کشلورزی (ب)



شکل ۶- شاخص عامل کلی برای کاربری اراضی کشاورزی



شکل ۷- شاخص عامل کلی برای کاربری اراضی شهری



الف- نقشه خطا برای کاربری اراضی شهری

ب- نقشه خطا شده برای کاربری اراضی کشاورزی

شکل ۸- نقشه خطا برای کاربری های شهری (الف) و کشاورزی (ب). در این نقشه‌ها A نشان دهنده سلولهایی هست که در واقعیت دچار تغییر کاربری شده‌اند و مدل آنها را درست پیش بینی کرده است، B نشان دهنده سلولهایی است که به عنوان سلولهای تغییر پیدا کرده شناسایی شده‌اند ولی در واقعیت تغییری نکرده‌اند، C نشان دهنده سلولهایی هست که به عنوان سلولهای تغییر نیافته شناسایی شده اند ولی در واقعیت تغییر پیدا کرده‌اند و D نشان دهنده سلولهایی هست که به عنوان سلولهای تغییر نیافته شناسایی شده‌اند و در واقعیت هم تغییر نکرده‌اند.

۵- نتیجه گیری

این مطالعه، به مدلسازی تغییرات کاربری اراضی چندگانه با استفاده از تلفیقی از سیستم اطلاعات مکانی، رگرسیون انطباقی چند متغیره اسپلاین و همچنین تصاویر ماهواره‌ای در بخشی از اراضی مربوط به شمال ایران بین سالهای ۱۹۹۲ تا ۲۰۱۴ پرداخته است. در سالهای اخیر، اکثر مدلسازی‌های تغییرات کاربری، به مدلسازی تغییرات کاربری اراضی یگانه (مثل توسعه شهری) مربوط می‌شود. در حالیکه آنچه در دنیای واقعی به وقوع می‌پیوندد تغییرات کاربری اراضی چندگانه می‌باشد. بنابراین ضروری هست که محققان به تحقیق و پژوهش در این زمینه بپردازند. تغییرات کاربری اراضی که در این تحقیق مدلسازی شده‌اند، مربوط به کاربری‌های اراضی شهری و کشاورزی بوده‌اند. فاکتورهایی که به عنوان محرک تغییرات کاربری اراضی در نظر گرفته شده بودند عبارتند از: ارتفاع، جهت، فاصله از مناطق شهری، فاصله از مناطق کشاورزی، فاصله از جاده‌های اصلی، فاصله از رودخانه و نهایتا شیب منطقه. بعد از انجام مدلسازی، تحلیل حساسیت روی این فاکتورها انجام شد. نتایج حاصل از این تحلیل نشان داد که چهار فاکتور فاصله از مناطق کشاورزی، فاصله از مناطق شهری، ارتفاع و فاصله از رودخانه بیشترین نقش را در فرآیند مدلسازی دارند. این نتایج می‌تواند تصمیم گیرندگان محیط زیست را در اتخاذ سیاست‌های مناسب یاری کند. مدل حاصل شده برای بازه زمانی ۱۹۹۲ تا ۲۰۱۴ روی متغیرهای مستقل موجود در سال ۲۰۱۴ اعمال شد و نقشه پیش‌بینی تغییرات کاربری اراضی برای سال ۲۰۳۶ بدست آمد. این نقشه پیش‌بینی می‌تواند مدیران و برنامه‌ریزان شهری را در تخصیص منابع، احداث مراکز خدماتی و توسعه هوشمند کاربری‌های اراضی برای سال‌های آتی یاری کند. همچنین در این مطالعه شاخص عامل کلی که توسط پونتیوس و سی [۶] ارائه شده برای کلاسه‌بندی چندگانه استفاده گردید و از این شاخص برای استخراج اعداد موجود در داخل جداول احتمال استفاده شد. می‌توان گفت که دو شاخص عامل نسبی و شاخص عامل کلی با توجه به پیچیدگی تغییرات اتفاق افتاده در منطقه، جزء شاخص‌های مناسب و دقیق برای ارزیابی مدلسازی تغییرات کاربری می‌باشند.

مراجع

[1] Geist, H., McConnell, W., Lambin, E F., Moran, E., Alves, D. and Rudel, T. (2006). “Causes and Trajectories of Land-Use/Cover Change, Land-Use and Land-Cover Change, Global Change” -The IGBP Series 2006, pp. 41-70.

[2] Tayyebi, A. and Pijanowski, B.C. (2014). “Modeling multiple land use changes using ANN, CART and MARS: comparing tradeoffs in goodness of fit and explanatory power of data mining tools”. International Journal of Applied Earth Observation and Geoinformation, Vol. 28, pp. 102-116.

[3] Verburg, P.H., Soepbaoer, W., Veldkamp, A., Limpiada, R., Espaldon, V. and Mastura, S. (2002). “Modeling the spatial dynamics of regional land use: the CLUE S model.” Environmental Management. Vol. 2, No. 30, pp. 391–405.

[4] Veldkamp, A., and Lambin, E. F. (2001). “Predicting land-use change.” Agriculture, Ecosystems and Environment, VOL. 85, pp. 1–6.

[5] Lambin, E.F. and Geist, H.J. (Eds.). (2006). “Land-use and Land-Cover Change.” Local Processes and Global Impacts, Springer, Berlin.

[6] Pontius Jr., R.G. and Si, K. (2014). “The Total Operating Characteristics to measure diagnostic ability for multiple thresholds.” International Journal of Geographical Information Science, Vol. 28, No. 3, pp. 570-583.

[7] Pijanowski, B.C., D.T. Long, S.H. Gage and W.E. Cooper. (1997). “A Land Transformation Model” Conceptual Elements, Spatial Object Class Hierarchies, GIS Command Syntax and an Application to Michigan's Saginaw Bay Watershed. Land Use Modeling Workshop. Sioux Falls, South Dakota, June 3-5, 1997. Sponsored by NCGIA and USGS.

[9] Tayyebi, A., Delavar, M.R., Saeedi, S. and Amini, J. (2008). “Monitoring the urban expansion by multi-temporal GIS maps. In: Application of Remote Sensing and Imagery (TS 5B), Integrating Generations, FIG Working Week 2008 and FIG/ UNHABITAT Seminar.

[10] Tayyebi, A., Delavar, M.R., Saeedi, S., Amini, J. and Alinia, H. (2008). “Monitoring land use change by multi-temporal Landsat remote sensing imagery.” Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XXXVII.

[11] Foroutan, E., and M. R. Delavar, (2012). “Urban growth modeling using fuzzy logic.” Proc. ASPRS Annual Conference. 2012, Sacramento, California, USA.

[12] Mohammady, S, and M R Delavar (2014). “A spatio-temporal urban growth modeling a case study: Tehran Metropolis.” Journal of Settlements and Spatial Planning, Vol. 5, No. 1, pp. 83-91.

[13] Li, X. and Yeh, A.G.O. (2002).” Neural-network-based cellular automata for simulating multiple land use changes using GIS.” International Journal of Geographical Information Science, Vol. 16, No. 4, pp. 323–343.

[14] Verburg, P.H., Soepbaoer, W., Veldkamp, A., Limpiada, R., Espaldon, V. and Mastura, S. (2002). “Modeling the spatial dynamics of regional land use: the CLUE S model.” Environmental Management, Vol. 30, pp. 391–405.

[15] Ballestores Jr., F., Qiu, Z., Nedorezova, B.N., Nedorezov, L.V., Ferrarini, A., Ramathilaga, A. and Ackah, M. (2012).” An integrated parcel-based land use change model using cellular automata and decision tree.” Proc. International Academy of Ecology and Environmental Sciences, Vol. 2, No. 2, pp. 53–69.

[16] Ralha, C.G., Abreu, C.G., Coelho, C.G., Zaghetto, A., Macchiavello, B. and Machado, R.B. (2013).”A multi-agent model system for land-use change simulation.” Environmental Modeling & Software, Vol. 42, pp. 30–46.

[17] Zha, W. and Chan, W-Y. (2005). “Objective speech quality measurement using statistical data mining.” EURASIP Journal on Applied Signal Processing, Vol. 9. pp. 1410–1424.

[18] Tayyebi, A., Pijanowski, B. C., & Tayyebi, A. H. (2011). An urban growth boundary model using neural networks, GIS and radial parameterization: An application to Tehran, Iran. Landscape and Urban Planning, Vol. 100, No. 1, pp. 35-44.

[19] Tayyebi, A., Pijanowski, B. C., & Pekin, B. (2011). Two rule-based urban growth boundary models applied to the Tehran Metropolitan Area, Iran. *Applied Geography*, Vol. 31, No. 3, pp. 908-918.

[20] De Vries, P. G. (1986). “Stratified random sampling”, Sampling Theory for Forest Inventory, Springer. pp. 31-55.

[21] Tayyebi, A., Pijanowski, B. C., Linderman, M. and Gratton, C. (2014). “Comparing three global parametric and local non-parametric models to simulate land use change in diverse areas of the world.” Environmental Modelling & Software, Vol. 59, pp. 202-221.

[22] Anderson, J. R., Hardy, E.E., Roach, J.T. and Witmer, R. E. (1976). “A land use and land cover classification system for use with remote sensor data.” USDI, Geological Survey Professional Paper No. 964. Washington, DC, pp. 26-36.

[23] Tayyebi, A. H., Tayyebi, A. and Khanna, N. (2014). “Assessing uncertainty dimensions in land-use change models: using swap and multiplicative error models for injecting attribute and positional errors in spatial data.” International Journal of Remote Sensing, Vol. 35, No. 1, pp. 149-170.

[24] Rifkin, R. and Klautau, A. (2004). “Parallel networks that learn to pronounce English text.” Journal of Machine Learning Research, pp. 101–141.

[25] Tsoumakas, G., Katakis, I. and Vlahavas, I. (2010).”Mining multi-label data.” In: Maimon, O., Rokach, L. (Eds.), Data Mining and Knowledge Discovery Handbook. , 2nd ed.Springer, (Chapter 34).pp. 667–685.

[26] Dietterich, T.G. and Bakiri, G. (1995).”Solving multiclass learning problems via error correcting output codes.” Journal of Artificial Intelligence Research, Vol. 2, NO. 1, pp. 263–286.

[27] Friedman, J.H. (1991).”Multivariate Adaptive Regression Splines (with discussion).”Annals of Statistics Vol. 19, No. 1 ,pp. 1-67.

[29] Pijanowski, B.C., Pithadia, S., Shellito, B.A. and Alexandridis, K. (2005). “Calibrating a neural network based urban change model for two metropolitan areas of Upper Midwest of the United States.” International Journal of Geographical Information Science. Vol. 19, pp. 197-215.

[30] Pontius Jr., R.G. and Batchu, K. (2003). “Using the relative operating characteristic to quantify certainty in prediction of location of land cover change in India.” Transactions in GIS, Vol. 4, No. 7, pp. 467-484.