

ارزیابی قابلیت رگرسیون وزنی جغرافیایی در بهبود پیش‌بینی رشد اراضی شهری با استفاده از سلول‌های خودکار

بابک میرباقری^{۱*}، عباس علیمحمدی^۲

- ۱- دانشجوی دکتری سیستم‌های اطلاعات مکانی، دانشگاه صنعتی خواجه نصیرالدین طوسی، دانشکده مهندسی ژئودزی و ژئوماتیک
۲- دانشیار دانشگاه صنعتی خواجه نصیرالدین طوسی، دانشکده مهندسی ژئودزی و ژئوماتیک، قطب علمی فناوری اطلاعات مکانی

تاریخ دریافت مقاله ۱۳۹۵/۰۳/۲۵ تاریخ پذیرش مقاله: ۱۳۹۵/۱۱/۱۳

چکیده

رگرسیون وزنی جغرافیایی منطقی (GWLR) نسخه محلی مدل رگرسیون منطقی (LR) است که در هر مکان رابطه متفاوتی را بین متغیرهای وابسته و مستقل برآورد می‌کند. در تحقیق حاضر از مدل GWLR در توسعه قوانین تبدیل سلول‌های خودکار (CA) استفاده و کارایی آن در پیش‌بینی رشد شهری در مقابل مدل CA مبتنی بر رگرسیون منطقی (Logistic-CA) مورد ارزیابی قرار گرفت. همچنین پارامتری تحت عنوان ضریب رشد حاشیه ای برای تعیین توازن بهینه میان فرایندهای رشد حاشیه ای و رشد خود انگیخته که از جمله فرایندهای مهم در رشد اراضی شهری محسوب می‌شوند، تعریف گردید. جهت ارزیابی دقت مدل نیز از آماره کاپای فازی در کنار ضریب معمول کاپا استفاده شد. مدل پیشنهادی در تحقیق حاضر جهت پیش‌بینی توسعه بخشی از اراضی شهری واقع در جنوب غرب منطقه کلانشهری تهران برای دوره زمانی ۱۳۹۱-۱۳۸۳ اجرا گردید. نتایج این مطالعه نشان داد استفاده از مدل GWLR در تعریف قوانین تبدیل CA منجر به افزایش قابل توجه دقت پیش‌بینی رشد شهری در مقایسه با مدل Logistic-CA می‌گردد. دقت پیش‌بینی مدل توسعه داده شده در تحقیق حاضر بر حسب ضریب کاپا برابر ۰/۵۴ می‌باشد که افزایش دقتی معادل ۰/۲۴ را در مقایسه با مدل Logistic-CA نشان می‌دهد. همچنین دقت مدل پیشنهادی بر اساس شاخص کاپای فازی با در نظر گرفتن فواصل ۵۰ و ۱۰۰ متری برای تابع نمایی افت فاصله، به ۰/۶۹ و ۰/۷۶ می‌رسد.

واژه‌های کلیدی: سلول‌های خودکار، رگرسیون وزنی جغرافیایی، رگرسیون منطقی

* نویسنده مکاتبه کننده: تهران، تقاطع میرداماد و ولیعصر (عج)، دانشکده مهندسی ژئودزی و ژئوماتیک، دانشگاه صنعتی خواجه نصیرالدین طوسی

۱- مقدمه

گسترش حاشیه نشینی و رشد بی قاعده شهرها موجب به وجود آمدن مسائل و مشکلات گوناگون نظیر تبدیل اراضی کشاورزی و چشم اندازهای طبیعی و کاهش فضای سبز، تراکم بیش از اندازه جمعیت، آلوده ساختن محیط زیست، استفاده بیش از حد از ظرفیت زمین، ایجاد اختلال در حمل و نقل، به وجود آمدن مشکلات اقتصادی و اجتماعی نظیر گسترش فقر در حاشیه‌ها، ظهور مشاغل کاذب، کمبود خدمات و تسهیلات و کاهش سرانه های شهری؛ و در نهایت کاهش کیفیت زندگی در میان ساکنین این مناطق می‌شود.

پیش بینی تغییرات کاربری به دلیل فراهم نمودن امکان هدایت صحیح رشد شهری و یا کاهش مساحت تغییرات آن، در برنامه ریزی های محیطی اهمیت بالایی یافته است [۱] و در این میان شبیه سازی سناریوهای رشد شهری و ارزیابی محیطی آنها می‌تواند در انتخاب آلترناتیوهای برنامه ریزی در راستای توسعه پایدار مفید واقع شود [۲].

سلول های خودکار^۱ (CA) از جمله معروفترین مدل های شبیه سازی مکانی هستند. CA به عنوان یک مدل پایین به بالا می‌تواند الگوهای پیچیده ای را با تکرار قوانین تعریف شده در سطح پایین شبیه سازی نمایند [۳]. در نتیجه با استفاده از CA می توان پدیده های پیچیده مکانی-زمانی را بررسی و از آن برای آزمون آلترناتیوهای مختلف برنامه ریزی استفاده نمود [۴]. از زمانی که تابلر در سال ۱۹۷۹ برای اولین بار امکان به کارگیری CA در جغرافیا را مطرح نمود [۵] تا کنون این مدل در موارد متعددی جهت شبیه سازی تغییرات کاربری و به خصوص رشد شهری مورد استفاده قرار گرفته است. مدل های CA ارائه شده توسط وایت و انگلن: [۶، ۷ و ۸]، مدل مشهور SLEUTH [۹]، مدل متداول مبتنی بر رگرسیون

منطقی^۲ (LR) [۱۰]، مدل مبتنی بر شبکه های عصبی [۱۱]، مدل های کرنل پایه [۱۲ و ۱۳]، مدل CA فازی [۱۴] و مدل های CA مبتنی بر الگوریتمهای تکاملی [۱۵ و ۱۶] مثالهایی از پیشرفتهای در زمینه شبیه سازی رشد شهری می باشند.

CA را می توان ترکیبی از فضای سلولی با چهار مولفه: سلولهای دو بعدی تشکیل دهنده فضا، وضعیت سلول در زمان مشخص، همسایگی و قوانین تبدیل^۳ دانست [۱۷] که فرایندی متغیر با زمان را شبیه سازی می کند. ضمن اینکه با توجه به محیط و موضوع مورد مطالعه برخی فرضیات در نظر گرفته شده برای CA کلاسیک نظیر منظم بودن بازه های زمانی و شکل هندسی سلولها، تعریف یکسان همسایگی در منطقه مطالعاتی، بسته بودن سیستم و ثابت بودن قوانین تبدیل در زمان و مکان، قابل تغییر است [۱۸]. از طرفی رابطه میان رشد شهری و پارامترهای تاثیر گذار بر آن لزوما یک رابطه ثابت و صادق در کل منطقه مورد مطالعه نیست [۱۹] زیرا که عملا قوانین حاکم بر تغییرات کاربری در مکانهای مختلف متفاوت است و این ناهمگنی مکانی می تواند تاثیر مستقیم بر کارکرد مدل های پیش بینی داشته باشد [۲۰]. ضمن اینکه عملکرد محلی فاکتورهای تاثیر گذار بر فرایند توسعه، منجر به ظهور الگوهای متفاوت از رشد شهری می شود [۲۱]. بنابراین از آنجاییکه قلب CA قوانین تبدیل آن است [۱۲، ۱۴، ۲۲ و ۲۳] و نقش اصلی را در نتایج به دست آمده از آن بازی می کند، تغییر نگرش در تعریف قوانین تبدیل از حالت ثابت به متغیر^۴ یا محلی [۲۴] می تواند نقش تعیین کننده ای در بهبود دقت پیش بینی داشته باشد.

تعریف قوانین تبدیل به صورت متغیر در تحقیقات معدودی مورد توجه قرار گرفته است. چنگ و مسر (۲۰۰۴) با تخصیص وزنه های متفاوت به عوامل

² Logistic Regression

³ Transition Rules

⁴ Non-stationary

¹ Cellular Automata

کاربری اراضی و توسعه شهری را به خود اختصاص داده است [۱۰، ۳۰، ۳۱، ۳۲، ۳۳، ۳۴، ۳۵، ۳۶ و ۳۷]. همچنین این مدل در تحقیقات متعددی به عنوان مبنای مقایسه مدل‌های جدید پیشنهادی نیز مورد استفاده قرار گرفته است: [۱۲، ۱۳، ۱۵، ۱۶، ۲۶، ۲۸ و ۳۸]. در این مدل رابطه میان تغییرات کاربری و عوامل تاثیر گذار بر آن در کل منطقه مورد مطالعه ثابت فرض شده و در نتیجه تفاوت‌های محلی تاثیر عوامل نادیده انگاشته می‌شود [۲۸ و ۳۹]. مدل فراگیر LR علی‌رغم تفسیر پذیری بالا، قادر به مدل سازی روابط پیچیده میان متغیرهای مستقل و وابسته نیست [۴۰].

رگرسیون وزنی جغرافیایی^۳ (GWR) مدل محلی آماری است که بر خلاف مدل‌های فراگیر در هر مکان رابطه متفاوتی را بین متغیرهای مستقل و وابسته برآزش می‌کند [۴۱]. این مدل از زمان معرفی در سال ۱۹۹۶ تاکنون در مطالعات متعددی مورد استفاده قرار گرفته است [۴۲، ۴۳، ۴۴، ۴۵ و ۴۶]. از مهمترین مزایای استفاده از GWR علاوه بر غیرثابت بودن رابطه میان متغیرها، نادیده گرفتن مرزهای سیاسی و قراردادی در برآورد رابطه رگرسیونی است. در صورتی که متغیر وابسته از نوع دو حالت باشد، آنگاه این مدل تحت عنوان رگرسیون وزنی جغرافیایی منطقی^۴ (GWLR) نامیده می‌شود [۴۷، ۴۸ و ۴۹]. تاکنون از GWLR برای مدل سازی استاتیک احتمال توسعه اراضی شهری و بررسی تغییرات مکانی عوامل موثر بر آن نیز استفاده شده است [۱۹ و ۵۰]. با توجه به قابلیت مدل GWLR در مدلسازی محلی احتمال توسعه شهری، از این مدل در تعریف قوانین تبدیل CA استفاده شده و اثر آن در افزایش دقت پیش بینی مورد ارزیابی قرار گرفته است.

در ادامه مقاله پس از معرفی منطقه مورد مطالعه و داده‌های مورد استفاده، مدل GWLR معرفی می‌شود.

تاثیر گذار در پروژه‌های توسعه زمین به صورت دانش مبنا در شهر ووهان چین، رویکردی محلی را برای تعریف قوانین تبدیل CA دنبال کرده اند [۲۵]. لی و لیو (۲۰۰۶) با فرض عدم استفاده از یک قانون انتقال ثابت برای شبیه‌سازی فرایندهای پیچیده جغرافیایی از تکنیک استدلال مبتنی بر مورد^۱ (CBR) برای تعریف قوانین تبدیل استفاده نمودند [۲۶]. گریتمن و همکاران (۲۰۰۷) نشان دادند که دقت پیش بینی مدل CA با قوانین همسایگی ناحیه ای در مقایسه با CA با قوانین همسایگی که به صورت فراگیر^۲ برای کل منطقه مطالعاتی تعریف شده است، بالاتر است [۲۷]. لی و همکاران (۲۰۰۸) با تقسیم کل منطقه مطالعاتی به ۶ ناحیه بر اساس مرزهای اداری و کالیبراسیون مدل برای هر یک از ناحیه‌ها با استفاده از الگوریتم ژنتیک، وزن پارامترهای موثر در توسعه را جداگانه استخراج و در نتیجه برای هر منطقه قانون تبدیل جداگانه‌ای تعریف نمودند [۲۱]. لی و همکاران (۲۰۱۵) نیز در تحقیق خود با هدف افزایش دقت شبیه‌سازی در مناطق با دقت پایین تر مدل فراگیر LR، مدلی از ترکیب دو مدل درخت رگرسیون و شبکه عصبی مصنوعی را مورد استفاده قرار دادند [۲۸]. در ایران نیز فروتن (۱۳۹۱) در تحقیق خود سعی نمود با انجام کالیبراسیون مدل به صورت منطقه مبنا بر اساس الگوریتم ژنتیک، خروجی آن را هر چه بیشتر به واقعیت نزدیک نماید. در این تحقیق کل منطقه مورد مطالعه (شهر اصفهان) به مناطقی کوچکتر (۴، ۱۵ و ۵۴ گانه) تقسیم و عمل کالیبراسیون برای هر منطقه به طور جداگانه انجام شد [۲۹].

به علت دو حالت (۱ و ۰) بودن داده‌های توسعه شهری، استفاده از مدل فراگیر LR در قوانین تبدیل CA - که تحت عنوان اختصاری Logistic-CA نیز نامیده می‌شود [۳۰]، بخش عمده ای از مطالعات مربوط به شبیه‌سازی

³ Geographically Weighted Regression

⁴ Geographically Weighted Logistic Regression

¹ Case-Based Reasoning

² Global

سپس نحوه تعریف مولفه‌های CA و بخصوص چگونگی تعریف قوانین تبدیل شرح داده می‌شود. در ادامه نتایج حاصل از شبیه‌سازی مورد بررسی قرار گرفته و نتیجه‌گیری تحقیق ارائه می‌گردد.

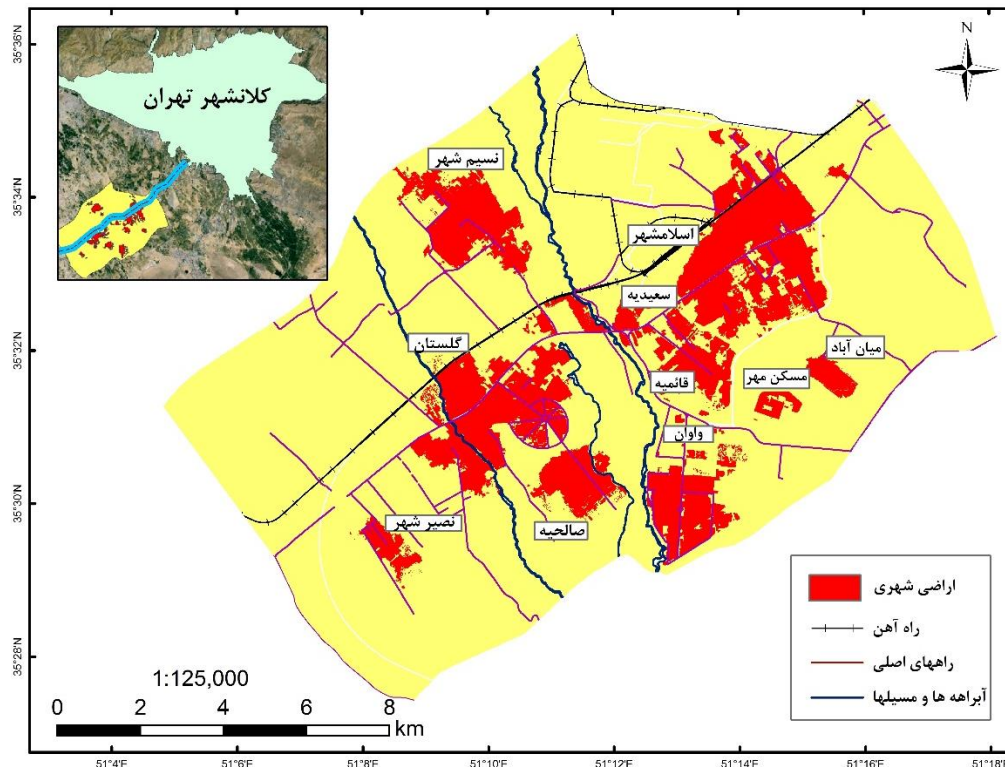
۲- منطقه مطالعاتی

تهران از اواسط دهه بیست شمس به دلیل تمرکز سرمایه‌گذاریهای دولتی و خصوصی، رشد بی سابقه‌ای یافت و در مدتی کوتاه از سایر مناطق کشور پیش افتاد و به مرکز اصلی و قطب بلامنازع اقتصادی کشور تبدیل شد [۵۱]. بطوریکه جمعیت تهران که در سال ۱۳۱۷ فقط دو برابر شهر مشهد دومین شهر ایران بود، در سال ۱۳۴۵ به ۶.۶ برابر رسید [۵۲].

با وجود شکل‌گیری روند مهاجرتی به سوی شهر تهران با توجه به تغییر و تحولات رخ داده، وجود عوامل دافعه شهر تهران نظیر گرانی مسکن و در مقابل وجود عوامل جاذبه نقاط پیرامونی بویژه ارزانی زمین و مسکن موجب شکل‌گیری روند مهاجرتی از شهر تهران به مناطق پیراشهری [۵۳] و در نتیجه توسعه مجموعه شهری تهران گردید. توسعه مراکز جمعیتی مجموعه شهری تهران در قالب حوزه‌های شهری مختلف رخ داده است که یکی از این حوزه‌های فعال شهری (حوزه شهری اسلامشهر-رباط کریم) حول بزرگراه آیت الله سعیدی (جاده ساوه) در جنوب غرب تهران شکل گرفته است [۵۱]. به طور کلی روستاها، شهرها و شهرکهای واقع در این حوزه به علت فاصله نزدیک با این شهر و قابلیت گسترش فیزیکی، طی چند دهه اخیر عرصه‌ای برای گسترش کارکردهای خوابگاهی، کارگاهی و صنعتی بوده است [۵۴] و وجود کارگاههای کوچک و بزرگ در این ناحیه که بدون هیچ ارزیابی اصولی احداث گردیده‌اند، خود باعث جلب مهاجرین بسیار و اسکان جمعیت فراوانی شده است [۵۵]. در این میان اسلامشهر به‌عنوان یکی از مهمترین شهرهای این حوزه، در اثر عوامل ملی و منطقه‌ای، بویژه قرارگیری در حوزه نفوذ

شهر تهران به مهم‌ترین و بزرگ‌ترین کانون جذب مهاجران به تهران تبدیل شد [۵۳]. ضمن اینکه عدم وجود مدیریت یکپارچه شهری منجر به پراکنش ناموزون جمعیت و فعالیت از اسلامشهر به نواحی پیرامونی و بروز و تشدید روند اسکان غیر رسمی در این ناحیه گردیده است [۵۶]. توسعه روستاها و تبدیل آنها به شهر در سالهای اخیر نظیر ایجاد شهر صالحیه (از تجمیع روستاهای امامزاده باقر و صالح آباد)، تبدیل روستای نصیر آباد به شهر در سال ۱۳۸۴ و نیز تاسیس سایت مسکن مهر اسلامشهر با ظرفیت پنج هزار واحد مسکونی بیانگر تداوم روند توسعه شهری در این حوزه است. لازم به ذکر است که افزایش جمعیت این ناحیه و نیاز به ارائه خدمات بهتر به ساکنین آن، منجر به تاسیس شهرستان بهارستان به مرکزیت شهر گلستان در سال ۱۳۸۹ با جدا شدن بخشهای بوستان و گلستان از شهرستان رباط کریم گردید.

لذا با عنایت به تحولات رخ داده در حوزه شهری فوق، محدوده مطالعاتی تحقیق حاضر نیز مناطق عمده‌ای از این حوزه شهری را در بر گرفته است که مشتمل بر بخشهایی از شهرستانهای اسلامشهر، بهارستان و رباط کریم با مساحتی نزدیک به ۱۷۷ کیلومتر مربع می‌باشد. این محدوده شهرهای اسلامشهر، گلستان، نسیم شهر، صالحیه، نصیر شهر و همچنین شهرکهای نظیر شهرک واوان، قائمیه، سعیدیه و میان آباد را شامل می‌شود (شکل (۱)). لازم به ذکر است متوسط رشد سالانه جمعیت نقاط شهری استان تهران در سالهای ۸۵-۹۰ معادل ۱/۴۱ درصد بوده است. این در حالی است که متوسط رشد جمعیت شهرهای اسلامشهر، نصیر شهر، گلستان و نسیم شهر بالاتر از میانگین استان و به ترتیب برابر ۱/۷۱، ۲/۵، ۲/۲۷ و ۳ درصد بوده است و تنها شهر صالحیه با رشدی معادل ۰/۷۷ درصد، رشدی کمتر از متوسط رشد جمعیت استان را تجربه کرده است [۵۷].

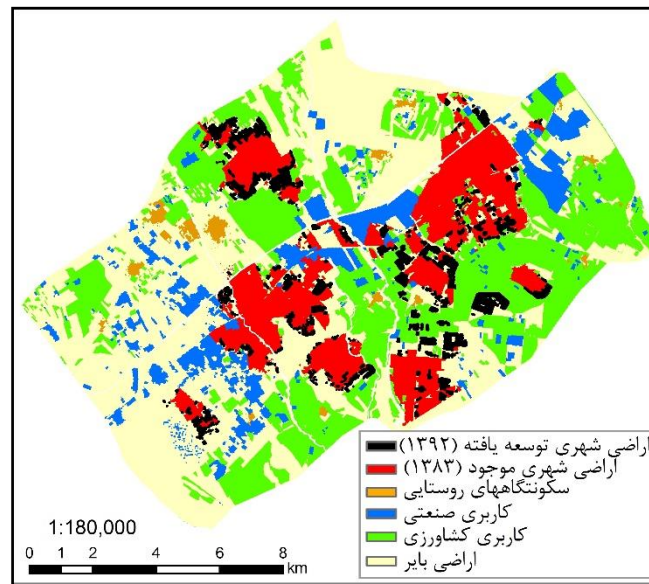


شکل ۱: منطقه مطالعاتی

۳- داده‌ها و متغیرهای مورد استفاده

داده‌های تحقیق شامل کاربری اراضی منطقه مطالعاتی به صورت بصری از تصویر ماهواره‌ای SPOT5 مربوط به سال ۱۳۸۳ و تصویر ماهواره‌ای WorldView-2 مربوط به سال ۱۳۹۱ استخراج شدند. سایر داده‌های مورد نیاز نظیر جاده‌ها و بزرگراه‌ها، رودخانه‌ها و مسیل‌ها و خط آهن نیز از نقشه‌های ۱:۲۵۰۰۰ سازمان نقشه‌برداری تهیه گردیدند. شکل (۲) نقشه کاربری اراضی استخراج شده از منطقه مطالعاتی، مربوط به سال ۸۳ و همچنین اراضی توسعه یافته شهری در دوره زمانی ۸۳-۹۱ را نمایش می‌دهد. با استفاده از منابع داده فوق، متغیرهای فاصله از اراضی شهری، فاصله از مراکز شهرها، کاربری کشاورزی/بایر، متغیر ترکیبی تراکم کاربری صنعتی و فاصله از کاربری صنعتی (که تحت عنوان متغیر تراکم-فاصله از کاربری صنعتی نامیده خواهد شد)، فاصله از شهر تهران، فاصله از روستاها، و

فاصله از راههای ارتباطی به عنوان متغیرهای مستقل (پیشگو) در نظر گرفته شدند که پس از آزمون معناداری و بررسی همبستگی میان آنها، برای ورود به مدل‌های فراگیر و محلی انتخاب می‌گردند. متغیر کاربری کشاورزی/بایر در واقع یک متغیر دو حالتی با ارزش یک برای کاربری بایر و ارزش صفر برای کاربری کشاورزی است. همچنین متغیر تراکم-فاصله از کاربری صنعتی نیز ترکیبی از فاصله از کاربریهای صنعتی و تراکم واحدهای صنعتی است که مقدار آن در فواصل کم و تراکم بالای کاربریهای صنعتی افزایش و با افزایش فاصله و کاهش تراکم، کاهش می‌یابد. افزوده شدن پارامتر تراکم، به فاصله به دلیل فراهم نمودن امکان توصیف محلهایی است که علاوه بر نزدیکی به کاربریهای شهری در کنار واحدهای صنعتی بزرگ و یا مجموعه‌ای از واحدهای صنعتی قرار گرفته‌اند.



شکل ۲: کاربری اراضی منطقه مطالعاتی در سال ۸۳ و اراضی توسعه یافته در بازه زمانی ۸۳-۹۱

۴-۱- رگرسیون وزنی جغرافیایی منطقی

GWR مدلی است مبتنی بر آمار محلی که در آن ضرایب متغیرهای مستقل برای مکان (s_i) ، به صورت جداگانه و بر اساس نمونه‌های همسایه s_i ، برآورد می‌شود (رابطه (۱)):

رابطه (۱)

$$y(\beta(s_i)) = \beta_0 + \sum_k \beta_k(s_i) x_k(s_i) + \varepsilon(s_i)$$

که در آن y متغیر وابسته، $\beta_k(s_i)$ ضریب متغیر مستقل k در s_i ، $x_k(s_i)$ مقدار متغیر مستقل k در s_i و $\varepsilon(s_i)$ مقدار مانده در s_i است.

در رگرسیون محلی نقاط نمونه بسته به میزان نزدیکی آنها به نقطه رگرسیون، با وزنهای متفاوتی در کالیبراسیون مدل شرکت می‌کنند. آنچه در تخصیص مقدار وزن به نقاط نمونه شرکت کننده در مدل نقش اساسی دارد، پهنای باند است.

در صورتیکه متغیر وابسته مانند توسعه و عدم توسعه شهری، به صورت دو حالت باشد، آن‌گاه مدل GWLR به‌عنوان یکی از انواع مدل‌های

کلیه متغیرهای مستقل در بازه صفر تا یک با استفاده از تابع خطی صعودی، استاندارد شده و در کالیبراسیون مدل‌های آماری مورد استفاده قرار گرفتند. لازم به ذکر است به علت واقع شدن منطقه مطالعاتی در دشتهای دامنه ای جنوب رشته کوه البرز، تغییرات ارتفاع در منطقه مطالعاتی ناچیز بوده و لذا متغیرهای شیب و یا ارتفاع در مدل مورد استفاده قرار نگرفتند. متغیر دو حالت اراضی شهری توسعه یافته / نیافته (توسعه یافته = ۱، توسعه نیافته = ۰) نیز به عنوان متغیر وابسته در کالیبراسیون مدل‌های LR و GWLR مورد استفاده قرار گرفت. همچنین اراضی شهری، صنعتی و روستایی موجود در ابتدای دوره شبیه سازی و همچنین حریم خطوط آهن و رودخانه کرج و مسیل‌ها و بزرگراه‌ها، به‌عنوان محدودیت در مدل CA وارد شدند.

۴- روش تحقیق

در ادامه به بیان روش تحقیق اجرا شده در این مقاله پرداخته می‌شود، روشهای رگرسیون و مدل طراحی شده و قوانین تدوین شده در قسمت‌های مختلف این بخش ارائه می‌شوند.

یکنواخت نبودن پراکندگی نمونه های مورد استفاده در فرایند کالیبراسیون، روش انطباقی^۴ در مقابل روش ثابت^۵ برای تعیین پهنای باند استفاده گردید [۴۹]. در روش انطباقی پهنای باند معادل فاصله تا N امین نزدیکترین همسایه است. لذا در محلهایی که تراکم نمونه ها پایین تر است شعاع همسایگی افزایش می یابد تا نقاط بیشتری در فرایند کالیبراسیون شرکت کنند. اما در روش ثابت، پهنای باند بر اساس فاصله و به صورت یکنواخت برای کل منطقه مطالعاتی تعریف می گردد. جهت تعیین میزان پهنای باند بهینه و دستیابی به توازن میان اریبی و واریانس از روش مینیمم سازی شاخص AIC^۶ تصحیح شده (AICc) [۴۹] استفاده شد. شاخص AICc برای مدل GWLR بر اساس رابطه ذیل محاسبه می گردد:

رابطه (۵)

$$AIC_c = -2 \log \left[\sum_{i=1}^n L(y_i) \right] + 2q + \frac{2 \times q(q+1)}{n-q-1}$$

که در آن f مقدار تابع درستنمایی^۷ برای مقادیر y_i در s_i می باشد و q تعداد پارامترهای برآورد شده برای مدل محلی است و n تعداد کل نمونه هاست. با کاهش پهنای باند، نیکویی برازش افزایش یافته و در مقابل مقدار q نیز افزایش می یابد. مینیمم مقدار AICc در واقع نقطه توازن میان نیکویی برازش و تعداد پارامترهای برآورد شده است.

در تعیین پهنای باند بهینه و همچنین کالیبراسیون GWLR از بسته نرم افزاری R، تحت عنوان GWmodel [۵۸] استفاده شد. همچنین برای کالیبراسیون هر یک از مدل های GWLR و LR، به صورت اختیاری از ۵۰۰۰ نمونه تصادفی (به تعداد مساوی نمونه های توسعه یافته و توسعه نیافته) استفاده شد.

GWGLM^۱ مورد استفاده قرار می گیرد (رابطه (۲)):

رابطه (۲)

$$\ln \left(\frac{P(s_i)}{1-P(s_i)} \right) = \beta_0 + \sum_k \beta_k(s_i) x_k(s_i)$$

که در آن $P(s_i)$ احتمال تبدیل متغیر وابسته y در s_i ، به یک است که در مطالعه حاضر برابر احتمال توسعه اراضی شهری است.

تعیین نوع تابع وزن دهی مکانی از مهمترین مواردی است که در کالیبراسیون مدل GWR مد نظر قرار می گیرد. توابع گوسی^۲ و دو مجذوره^۳ دو تابع آلترناتیو مشهور برای تعیین وزن مکانی نقاط می باشند (رابطه های (۳) و (۴)).

$$w_{ij} = \exp \left[-\frac{1}{2} \left(\frac{d_{ij}}{b} \right)^2 \right] \quad \text{رابطه (۳)}$$

$$w_{ij} = \begin{cases} \left[1 - \left(\frac{d_{ij}}{b} \right)^2 \right]^2 & \text{if } d_{ij} < b \\ 0 & \text{Otherwise} \end{cases} \quad \text{رابطه (۴)}$$

که در آنها w_{ij} وزن مکانی نمونه j تا نمونه i (نقطه رگرسیون)، d_{ij} فاصله نمونه های i و j و b مقدار پهنای باند می باشد.

به دلیل اینکه تابع گوسی دارای دنباله طولانی بوده و بر خلاف تابع دو مجذوره دارای آستانه همسایگی معین برای تخصیص وزنها نیست؛ لذا امکان کالیبراسیون مدل در پهنای باند کوچکتر هنگام بکارگیری این تابع وجود دارد. در صورت استفاده از تابع دو مجذوره این امکان وجود دارد که در برخی محلهای تنها نمونه های توسعه و یا تنها نمونه های عدم توسعه در محدوده همسایگی قرار گیرند که عملاً امکان کالیبراسیون مدل را از بین می برد. لذا در این تحقیق از تابع گوسی، برای وزن دهی نقاط نمونه استفاده گردید. همچنین به علت

⁴ Adaptive

⁵ Fixed

⁶ Akaike Information Criterion

⁷ Likelihood Function

¹ Geographically Weighted Generalized Linear Model

² Gaussian

³ bi-square

۲-۴- کالیبراسیون مدل رگرسیون منطقی محلی و فراگیر

در مطالعه حاضر ابتدا مدل فراگیر LR بر روی داده‌ها اجرا و سپس محلهایی که در آنها خطای برآورد شده توسط مدل فراگیر از حد آستانه مشخص شده بیشتر بود، برای کالیبراسیون مدل GWLR انتخاب شدند. مقدار خطا بر اساس رابطه ذیل محاسبه گردید:

$$\varepsilon_i = y_i - \hat{P}_i \quad \text{رابطه (۶)}$$

که در آن ε_i مقدار خطای برآورد شده در نقطه i و \hat{P}_i مقدار متغیر وابسته (یک یا صفر) و مقدار احتمال برآورد شده بر اساس مدل LR است. در تحقیق حاضر حد آستانه ۰.۰۱ برای خطای مجاز در نظر گرفته شد. در نتیجه مدل GWLR در مناطقی که احتمال برآورد صحیح مقادیر آنها توسط مدل LR کمتر از ۹۹ درصد بود، کالیبره گردید. این عمل باعث می‌گردد مناطقی که دارای دقت بالایی در مدل فراگیر می‌باشند، مجدداً در مدل محلی استفاده نشوند و ضمن پرهیز از دوباره کاری، در تعداد نمونه‌ها جهت کالیبراسیون مدل محلی نیز صرفه جویی شود.

انتخاب متغیرهای مستقل در مدل نیز بر مبنای معنی داری ضرایب برآورد شده و همچنین عدم همبستگی بالا میان متغیرهای مستقل که از نشانه‌های آن، تغییر علامت ضرایب برآورد شده برای متغیرهاست [۵۹] انجام شد. در این تحقیق در صورت وجود همبستگی میان متغیرهای پیشگو، متغیری که دارای تاثیر کمتری بر دقت بود، از مدل حذف گردید.

۳-۴- طراحی مدل CA

در این بخش به تعریف شیوه طراحی مدل و پارامترهای موثر آن پرداخته می‌شود.

۱-۳-۴- تعریف قوانین تبدیل

در این تحقیق دو نوع قوانین تبدیل ثابت (فراگیر) و متغیر (محلی) که به ترتیب از کالیبراسیون مدل‌های LR

و GWLR به‌دست آمدند برای شبیه‌سازی توسعه اراضی شهری با استفاده از CA مورد استفاده قرار گرفتند.

شکل کلی قوانین تبدیل طراحی شده در تحقیق حاضر جهت به‌کارگیری در مدل CA مطابق رابطه (۵) می‌باشد. بر طبق این رابطه احتمال تغییر کاربری سلول c_{ij} در زمان t ، $P_{CA}^t(c_{ij})$ برابر است با:

رابطه (۷)

$$P_{CA}^t(c_{ij}) = P(c_{ij}) \times \Omega^t(c_{ij}) \times U \times Const.$$

که در آن $P(c_{ij})$ ، احتمال تغییر سلول c_{ij} به شهر است. در صورتی که سلول c_{ij} در مناطق با خطای بیش از آستانه ۰.۰۱ در مدل LR قرار داشته باشد آنگاه $P(c_{ij})$ برابر احتمال حاصل از کالیبراسیون مدل GWLR و در غیر این صورت $P(c_{ij})$ برابر احتمال حاصل از کالیبراسیون مدل LR خواهد بود. در نتیجه در این حالت قوانین تبدیل عملاً حاصل ترکیب دو مدل فراگیر و محلی خواهند بود. در صورت استفاده از قوانین تبدیل ثابت - که همان مدل Logistic-CA می‌باشد - مقدار $P(c_{ij})$ ، تنها بر اساس کالیبراسیون مدل LR محاسبه می‌گردد. $\Omega^t(c_{ij})$ ضریب همسایگی در زمان t است که به‌صورت دینامیک مطابق رابطه (۸) محاسبه می‌شود:

$$\Omega^t(c_{ij}) = \frac{\sum_{m=1}^n S^t(c_{ij}^m)}{n} \quad \text{رابطه (۸)}$$

که در آن c_{ij}^m ، سلول m موجود در همسایگی سلول c_{ij} ، n تعداد سلول‌های موجود در همسایگی سلول c_{ij} است که بسته به اندازه پنجره همسایگی متغیر خواهد بود و $S^t(c_{ij}^m)$ وضعیت این سلول در زمان t می‌باشد. مقدار $S^t(c_{ij}^m)$ نیز بر اساس رابطه (۹) محاسبه می‌گردد:

رابطه (۹)

$$S^t(c_{ij}^m) = \begin{cases} 1 & \text{if } c_{ij}^m \text{ is urban} \\ 0 & \text{if } c_{ij}^m \text{ is not urban} \end{cases}$$

پیاده‌سازی مدل CA در نرم افزار MATLAB، نسخه 2013b انجام شد.

۴-۳-۲- تعریف ضریب رشد حاشیه ای

به‌طور کلی فرایند رشد شهری را می‌توان به سه نوع حاشیه‌ای^۱، خود‌انگیخته^۲ و پرکننده^۳ تقسیم‌بندی نمود [۶۰]. در فرایند رشد حاشیه‌ای توسعه در مرز مناطق شهری موجود روی می‌دهد. رشد خود‌انگیخته فرایندی است که در آن قطعات شهری بدون ارتباط با قطعات شهری موجود توسعه می‌یابند و در فرایند رشد پرکننده که در واقع نوع خاص فرایند رشد حاشیه‌ای است، توسعه در مناطقی که توسط اراضی شهری موجود محصور شده است، رخ می‌دهد. فرایندهای رشد حاشیه‌ای و رشد خود‌انگیخته در تحقیقات مختلف مرتبط با CA مورد توجه ویژه قرار گرفته‌اند. ذکر این نکته ضروری است که اساساً هدف از مطرح شدن مدل CA - با توجه به حضور پارامتر همسایگی در آن - شبیه‌سازی فرایندهای مبتنی بر رشد حاشیه‌ای بوده است. در منطقه مورد مطالعه تحقیق حاضر نیز گرچه غلبه اصلی با رشد حاشیه‌ای است، اما در مواردی نیز رشد خود‌انگیخته مشاهده می‌گردد. نمونه بارز آن طرح مسکن مهر است در شمال شهرک واوان و جنوب شهرک قائمیه، بدون اتصال با قطعات شهری پیش‌تر توسعه یافته، اجرا شده است (شکل ۱). لذا در تحقیق حاضر جهت ایجاد توازن بهینه میان این دو نوع توسعه، پارامتری تحت عنوان ضریب رشد حاشیه‌ای تعریف گردید.

همان‌طور که در رابطه (۷) مشاهده می‌شود؛ احتمال تغییر سلول c_{ij} در زمان t ، $(P_{CA}^t(c_{ij}))$ ، از حاصلضرب ضریب همسایگی، $(\Omega^t(c_{ij}))$ ، در احتمال تغییر سلول به شهر، به‌دست می‌آید. در نتیجه

U نیز جزء تصادفی قانون تبدیل است که بر اساس رابطه (۱۰) محاسبه می‌شود [۶]:

$$U = 1 + (-\ln(md))^{\alpha} \quad \text{رابطه (۱۰)}$$

که در آن rnd یک عدد تصادفی بین صفر و یک است و α پارامتری است که هر چه مقدار آن بالاتر باشد، اختلاف بین مقادیر بالا و پایین متغیر تصادفی بیشتر بوده و باعث تولید الگوی پراکنده می‌گردد [۳۴] و در نهایت $Const.$ نیز محدودیتهای در نظر گرفته شده برای توسعه اراضی شهری و شامل حریم رودخانه‌ها و مسیله‌ها، حریم خط آهن و همچنین اراضی شهری و روستایی و مراکز صنعتی موجود پیش از شروع شبیه‌سازی است. در صورت قرارگیری سلول c_{ij} در این مناطق، مقدار $Const.$ برابر صفر و در غیر این صورت برابر یک می‌باشد.

در این تحقیق بازه‌های زمانی شبیه‌سازی به‌صورت ماهانه در نظر گرفته شد. با توجه به اینکه شبیه‌سازی در دوره زمانی ۹ ساله (۱۳۹۱-۱۳۸۳) انجام می‌گیرد، در نتیجه کل تکرارهای شبیه‌سازی برابر ۱۰۸ تکرار خواهد بود. از آنجاییکه کل سلولهای توسعه یافته در این دوره زمانی ۴۰۲۵۰ عدد می‌باشد، لذا در هر تکرار معادل ۳۷۲ سلول شبیه‌سازی خواهد شد. همچنین مدل CA با پنجره‌های همسایگی مختلف (۳×۳) ، (۵×۵) ، ...، (۱۵×۱۵) آزمون شد. با توجه به اینکه دقت در حالت استفاده از مدل GWLR در قوانین تبدیل با اندازه پنجره همسایگی بیشتر از ۳×۳ ، کمتر از یک درصد افزایش نشان داد، لذا همسایگی ۳×۳ برای شبیه‌سازی رشد اراضی شهری در نظر گرفته شد. همچنین مشابه این وضعیت برای همسایگی ۵×۵ در شبیه‌سازی مدل Logistic-CA رخ داد. در نتیجه در این حالت از پنجره همسایگی ۵×۵ در شبیه‌سازی استفاده شد. همچنین لازم به ذکر است اندازه سلول در نظر گرفته شده برای اجرای مدل CA، با توجه به داده‌های مورد استفاده و حجم داده‌ها (تعداد کل سلولها)، در حد مقیاس ۱:۲۵۰۰۰ و برابر ۱۰ متر تعیین گردید. در این تحقیق

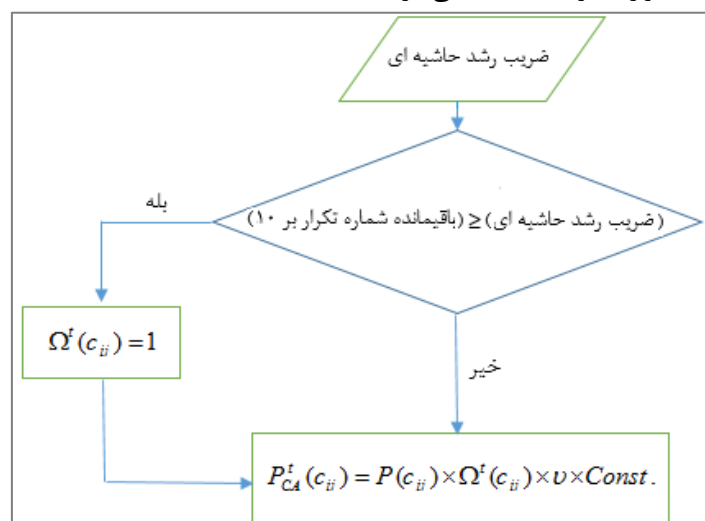
¹ Edge expansion

² Spontaneous

³ Infilling

محاسبه احتمال تغییر سلولها مشارکت داشته است. در نتیجه چنانچه مقدار این ضریب برابر صفر باشد، ضریب همسایگی برای تمامی سلولها، در تمامی تکرارها یکسان در نظر گرفته می شود $(\Omega^i(c_{ij}) = 1)$ و چنانچه ضریب رشد حاشیه ای برابر ۱۰ باشد، در کلیه تکرارها ضریب همسایگی در محاسبه احتمال توسعه دخیل خواهد بود. مقادیر بین صفر و یک نیز امکان شبیه سازی رشد حاشیه ای و خود انگیزه را به صورت توامان و با شدت های متفاوت فراهم می سازند. الگوریتم اعمال ضریب رشد حاشیه ای در فرایند شبیه سازی در شکل (۳) نمایش داده شده است.

احتمال تغییر سلولی که دارای همسایه شهری نیست، به علت صفر بودن ضریب همسایگی (رابطه (۸)) برابر صفر گردیده و در نتیجه از چرخه انتخاب به علت وجود سایر سلولها با مقادیر همسایگی غیر صفر خارج می گردد. این مساله باعث می گردد تا هسته های اولیه رشد خود سازمانده در فرایند شبیه سازی تشکیل نگردد. در این تحقیق این امکان فراهم گشته تا در برخی تکرارهای شبیه سازی، ضریب همسایگی در محاسبه احتمال تبدیل سلول، موثر نباشد. در واقع ضریب رشد حاشیه ای که بین اعداد صفر تا ده تعریف می شود، نشان می دهد در هر ده تکرار شبیه سازی، در چند تکرار ضریب همسایگی در



شکل ۳: الگوریتم اعمال ضریب رشد حاشیه ای در فرایند شبیه سازی

مقادیر مختلف ضریب رشد حاشیه ای (صفر تا ده) اجرا و مقدار بهینه آن تعیین گردید.

۴-۴-۴- ارزیابی دقت

در مطالعه حاضر، جهت مقایسه توسعه پیش بینی شده توسط مدل CA پیشنهادی با توسعه واقعی از ضریب کاپا [۶۱] استفاده شد. مزیت ضریب کاپا محاسبه دقت با در نظر گرفتن حالتی است که در آن سلولها به صورت کاملاً تصادفی شبیه سازی شوند. گرچه در بسیاری از تحقیقات مشابه از این ضریب برای سنجش میزان انطباق دو نقشه پیش بینی و واقعیت

بر این اساس مبنای اعمال ضریب همسایگی در هر تکرار، باقیمانده شماره تکرار بر عدد ۱۰ می باشد که چنانچه بزرگتر از ضریب رشد حاشیه ای باشد، ضریب همسایگی در این تکرار برای تمام سلولها، برابر یک در نظر گرفته می شود. در نتیجه معیار انتخاب تمام سلولها در این تکرار تنها حاصل ضرب $P(c_{ij})$ در جزء تصادفی و پارامتر محدودیت خواهد بود. چنانچه مقدار ضریب رشد حاشیه ای برابر ۱۰ باشد؛ در هیچ تکراری شرط برقرار نبوده و $\Omega^i(c_{ij})$ برابر مقدار واقعی خواهد بود.

در تحقیق حاضر شبیه سازی توسعه اراضی شهری با

تقلیل می یابد که تحت عنوان فاصله نیم کننده^۱ نامیده می شود.

$$f(d) = (0.5)^{d/2} \quad \text{رابطه (۱۲)}$$

میزان توافق دو نقشه A و B در سلول c، $(p_c^{A,B})$ ، که در نهایت مبنای محاسبه کاپای فازی قرار می گیرد، بر اساس رابطه (۱۳) محاسبه می گردد:

$$p_c^{A,B} = \min(\mu_{c,B_c}^A, \mu_{c,A_c}^B) \quad \text{رابطه (۱۳)}$$

که در آن A_c و B_c به ترتیب کلاسهای سلول c در نقشه های A و B هستند. با توجه به روابط فوق، هرچه مقدار d افزایش یابد، میزان تاثیر گذاری اختلافات جزئی میان دو نقشه نیز کاهش یافته و در نتیجه آماره کاپای فازی با فراهم کردن امکان تعریف فواصل مختلف، باعث نادیده گرفتن خطاهای کوچک در مقایسه بین دو نقشه می گردد. این مساله به خصوص برای مدل های پیش بینی مکانی نظیر CA که خروجی آنها اغلب در برنامه ریزی های کلان و استراتژیک شهری و منطقه ای کاربرد دارد و لزوماً نیازمند ارزیابی دقت بر اساس انطباق دقیق سلولها بر یکدیگر نیست، مزیت قابل توجهی به شمار می رود. لازم به ذکر است برای محاسبه کاپای فازی از نرم افزار Map Comparison Kit [۶۴] استفاده گردید.

۵- نتایج و بحث

در ادامه به ارائه نتایج به دست آمده از هر یک از روش ها پرداخته می شود.

۵-۱- کالیبراسیون مدل رگرسیون منطقی فراگیر

نتایج حاصل از کالیبراسیون اولیه و نهایی مدل فراگیر LR بر اساس متغیرهای اشاره شده در بخش ۳، در جدول (۱) ارائه شده است.

موجود استفاده شده است، اما این شاخص تنها میزان انطباق دقیق دو سلول بر یکدیگر را مبنای مقایسه قرار داده و نزدیکی و یا دوری خطای رخ داده با واقعیت موجود را مورد توجه قرار نمی دهد. لذا در این تحقیق جهت اندازه گیری میزان انطباق دو نقشه شبیه سازی شده و واقعیت موجود از آماره کاپای فازی، ارائه شده توسط هگن-زنگر در سال ۲۰۰۹، نیز استفاده شد [۶۲].

در محاسبه این آماره، از مفهوم فازی بودن مکان استفاده می گردد که بر اساس آن، درجه تعلق یک سلول به یک کلاس خاص بستگی به کلاس آن سلول و همسایگان آن دارد که این میزان تاثیر گذاری همسایگان بر اساس تابع افت فاصله تعیین می گردد [۶۲ و ۶۳].

در آماره کاپای فازی میزان تعلق سلول c در نقشه A، به کلاس z در نقشه B بر اساس رابطه (۱۱) محاسبه می گردد:

$$\mu_{c,j}^A = \max_{i=0}^m [M_{i,j} \times f(d_{c,i}^A)] \quad \text{رابطه (۱۱)}$$

که در آن $\mu_{c,j}^A$ درجه تعلق سلول c در نقشه A به کلاس z در نقشه B، $d_{c,i}^A$ نزدیکترین فاصله سلول c در نقشه A به سلول دارای کلاس i، f تابع افت فاصله و m تعداد کلاسها می باشد. $M_{i,j}$ میزان شباهت کلاسهای i و j است که در این تحقیق دو کلاس توسعه و عدم توسعه شباهتی به یکدیگر نداشته و در نتیجه مقدار $M_{i,j}$ در صورتیکه دو کلاس مشابه هم باشند برابر یک و در غیر اینصورت برابر صفر خواهد بود. درجه تعلق سلول c در نقشه B به کلاس i در نقشه A نیز بر اساس رابطه ای مشابه رابطه فوق محاسبه می گردد. همچنین در مطالعه حاضر از تابع افت فاصله نمایی [۶۲] مطابق رابطه (۱۲) استفاده گردید که در آن d فاصله ای است که در آن میزان اثر گذاری سلول همسایه بر سلول مرکزی به ۰/۵

¹ Halving Distance

جدول ۱: نتایج حاصل از کالیبراسیون اولیه و نهایی مدل LR

متغیرهای اولیه	علامت ضریب در رگرسیون چند متغیره	علامت ضریب در رگرسیون تک متغیره	تغییر نسبت درستنمایی در صورت حذف متغیر مستقل	متغیرهای نهایی	ضریب برآورد شده نهایی	تغییر نسبت درستنمایی در صورت حذف متغیر مستقل
فاصله از اراضی شهری	-	-	۱۱۹۷/۸	فاصله از اراضی شهری	*-۱۵/۹۳	۱۳۹۲/۵
کاربری کشاورزی/بایر	+	+	۴۳۹/۲	کاربری کشاورزی/بایر	*۲/۰۱	۳۹۵/۷
تراکم/ فاصله از کاربری صنعتی	-	-	۱۵۴/۰	تراکم/ فاصله از کاربری صنعتی	*-۸/۶۲	۲۸۳/۱
فاصله از جاده‌ها	+	-	۴۱/۹	فاصله از بزرگراه‌ها	*-۳/۳۱	۱۷۳/۱
فاصله از روستاها	+	+	۱۰/۵	فاصله از روستاها	*۱/۶۳	۴۱/۲
فاصله از شهر تهران	-	-	۳۶/۷	فاصله از شهر تهران	*-۱/۴۷	۳۱/۹
فاصله از مراکز شهری	+	-	۱۰۸/۴	مقدار ثابت	*۴/۶۰	-

*معنی‌داری در سطح ۹۹%

جاده‌ها شود، متغیر جدید فاقد همبستگی قابل توجه با متغیر فاصله از اراضی شهری بوده و می‌تواند در مدل LR مورد استفاده قرار گیرد. بر اساس مقادیر نسبت درستنمایی جدید، متغیرهای فاصله از اراضی شهری، کاربری کشاورزی/بایر و تراکم-فاصله از کاربری صنعتی به ترتیب بیشترین تاثیر را در دقت مدل LR دارا می‌باشند.

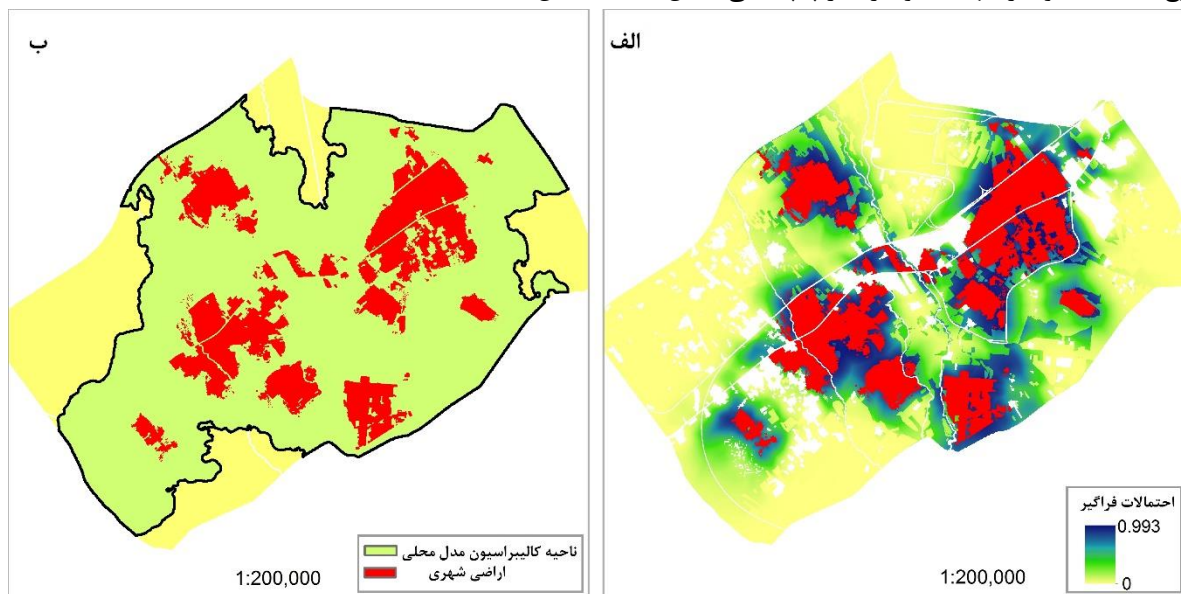
براساس نتایج نهایی به‌دست آمده از مدل LR و با توجه به اینکه متغیرهای مستقل بر اساس تابع خطی صعودی استاندارد شده‌اند، مشخص گردید که احتمال توسعه اراضی شهری با افزایش فاصله از اراضی شهری، فاصله از بزرگراه‌ها و فاصله از شهر تهران کاهش می‌یابد. این در حالیست که احتمال توسعه اراضی شهری با دور شدن از روستاها افزایش می‌یابد. همچنین ضریب مثبت متغیر کاربری کشاورزی/بایر نشان دهنده پتانسیل بالاتر زمینهای بایر در افزایش احتمال توسعه شهری است. ضریب منفی متغیر تراکم-فاصله از کاربری صنعتی نیز نشان می‌دهد پتانسیل توسعه در

در جدول فوق علائم هر یک از ضرایب متغیرهای مستقل در حالت تک متغیره و چند متغیره و تغییر مقدار نسبت درستنمایی در رگرسیون چند متغیره اولیه و نهایی نیز نمایش داده شده است. همانطور که مشاهده می‌شود، علامت دو متغیر مستقل فاصله از جاده‌ها و فاصله از مراکز شهری در حالت رگرسیون چند متغیره و تک متغیره متفاوت بوده و نشان دهنده وجود همبستگی داخلی میان متغیرها است. بررسی بیشتر نشان داد که علت تغییر علامت هر دو متغیر، همبستگی آنها با متغیر فاصله از اراضی شهری است. از آنجایی که تغییر نسبت درستنمایی که به‌عنوان یک شاخص نیکویی برازش در مدل‌های LR مورد استفاده قرار می‌گیرد [۱۹] - در صورت حذف متغیر فاصله از اراضی شهری بیش از سایر متغیرها است؛ لذا این متغیر در مدل باقی ماند و متغیر فاصله از مراکز شهری حذف گردید.

همچنین در صورتی که بزرگراه‌ها از سایر راه‌ها جدا گردد و متغیر فاصله از بزرگراه‌ها جایگزین فاصله از

متغیر فاصله از روستاها نیز در کالیبراسیون مدل محلی مورد استفاده قرار نگرفت. ضمن اینکه دو متغیر فوق بر اساس شاخص نسبت درست‌نمایی نیز از اهمیت کمتری در مدل LR برخوردار هستند. احتمالات حاصل از اجرای مدل LR و ناحیه مشخص شده برای کالیبراسیون مدل GWLR (بر مبنای مقادیر خطای بیش از ۰/۰۱ در مدل LR)، به ترتیب در شکل‌های (۴-الف و ۴-ب) نشان داده شده است.

مکان‌هایی بالاتر است که دارای تراکم پایین‌تر کاربری‌های صنعتی و نیز فاصله دورتری از آنها باشند. لازم به ذکر است از میان متغیرهای فوق، متغیر فاصله از شهر تهران به دلیل این که اصولاً متغیری فراگیر است و در مقیاس محلی (پهنای باند کوچک) تفسیر معنی‌داری نخواهد داشت؛ در مدل محلی مورد استفاده قرار نگرفت. همچنین به دلیل اینکه توسعه اراضی شهری اکثراً در مناطقی دور از روستاها رخ داده است و نیز جهت جلوگیری از پیچیدگی مدل،

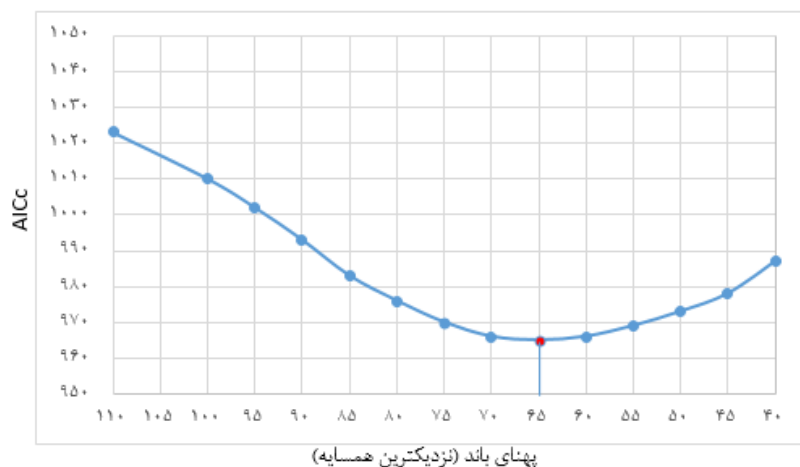


شکل ۴: احتمالات توسعه حاصل از اجرای مدل LR (الف) و ناحیه کالیبراسیون مدل GWLR (ب)

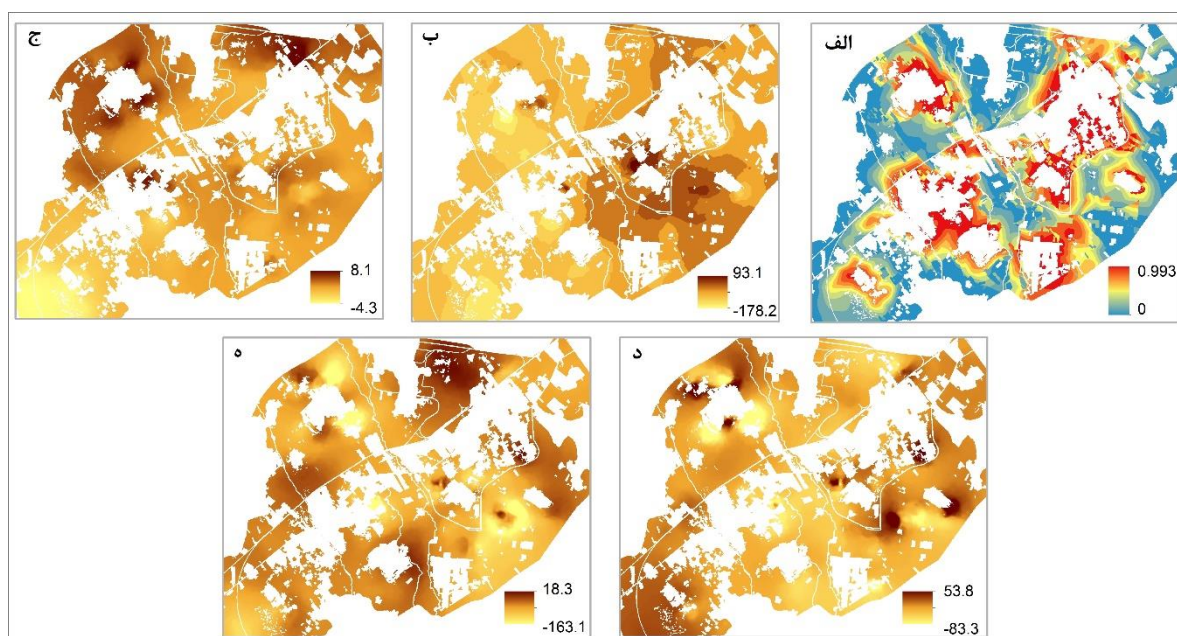
باند بهینه بر حسب تعداد نزدیکترین همسایه برابر ۶۵ می‌باشد. نقشه احتمالات محلی و همچنین ضرایب برآورد شده برای متغیرهای مستقل در مدل GWLR در شکل (۶) نشان داده شده است.

۵-۲- کالیبراسیون مدل GWLR

شکل (۵) تغییرات شاخص AICc را بر حسب پهنای باند برای کالیبراسیون مدل GWLR نمایش می‌دهد. همان‌طور که مشاهده می‌شود، پهنای



شکل ۵: تغییر مقادیر شاخص AICc بر حسب پهنای باند و تعیین ۶۵ همسایه نزدیک بعنوان پهنای باند بهینه

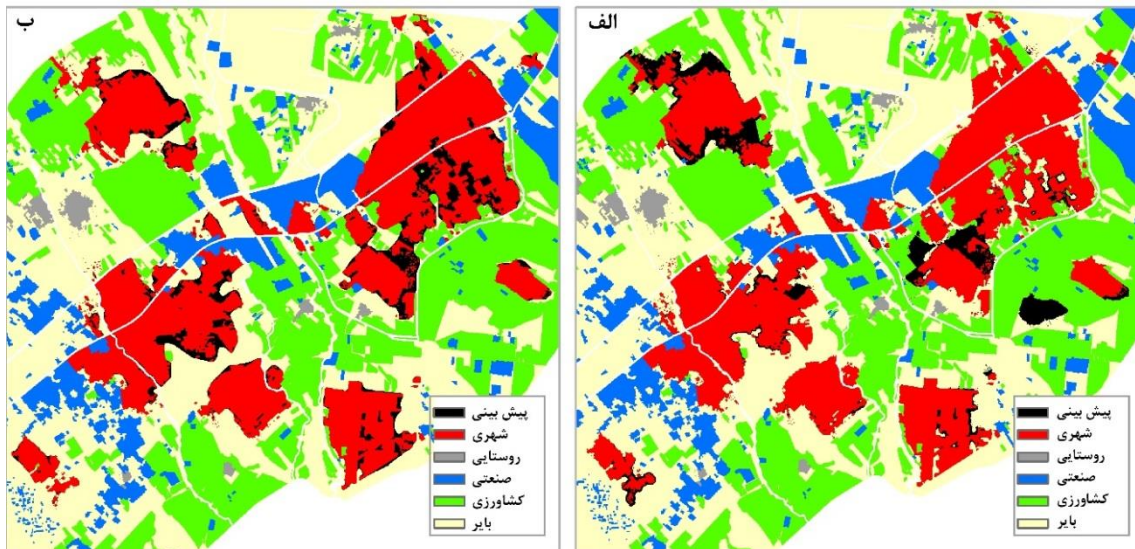


شکل ۶: مقادیر احتمال محلی (الف)، و ضرایب برآورد شده برای متغیرهای فاصله از اراضی شهری (ب)، کاربری کشاورزی/باغ (ج)، فاصله از بزرگراهها (د) و تراکم/فاصله از اراضی صنعتی (ه)

۵-۳- نتایج شبیه‌سازی توسعه شهری

نتایج حاصل از شبیه‌سازی توسعه اراضی شهری بر اساس قوانین تبدیل توسعه داده شده در تحقیق حاضر (استفاده از مدل GWLR در تعریف قوانین تبدیل و ترکیب آن با مدل LR) و همچنین قوانین تبدیل ثابت (Logistic-CA)، در شکل (۷) نمایش داده شده است. دقت پیش بینی رشد شهری توسط این دو مدل برای

سال ۱۳۹۱، بر اساس شاخص کاپا به ترتیب برابر ۰/۵۴ و ۰/۳ می‌باشد، که نشان می‌دهد استفاده از مدل GWLR در شبیه‌سازی رشد اراضی شهری منجر به افزایش قابل توجه دقت (معادل ۰/۲۴) شده است. لازم به ذکر است مقدار بهینه ضریب رشد حاشیه‌ای برای دو مدل فوق به ترتیب برابر ۸ و ۱۰ می‌باشد.



شکل ۷: نتایج شبیه‌سازی با مدل پیشنهادی (الف) و Logistic-CA (ب)

موجود میان توسعه و پارامترهای تاثیر گذار بر آن است موجب افزایش قابل توجه دقت شبیه‌سازی در مقایسه با حالتی که تنها از قوانین تبدیل ثابت در شبیه‌سازی استفاده می‌شود، می‌گردد.

لازم به ذکر است بر اساس ویژگی خاص مدل LR، نسبت نمونه‌های یک و صفر در متغیر وابسته گرچه باعث تغییر در ضرایب متغیرهای پیشگو نمی‌گردد، اما باعث تغییر مقادیر احتمال خروجی به علت تغییر در برآورد عرض از مبدا می‌شود [۶۵]؛ به طوری که بسته به نسبت نمونه‌های یک و صفر، مقادیر پایه احتمالات پیش‌بینی شده نیز افزایش و یا کاهش می‌یابند. در نتیجه، این خاصیت مدل LR در حالت محلی آن باعث می‌گردد تا احتمال توسعه برای قطعات بزرگتر توسعه یافته شهری (با توجه به تعداد بالاتر نمونه‌های مربوط به مناطق توسعه یافته) افزایش و برای قطعه‌های کوچکتر کاهش یابد. در نتیجه می‌توان گفت این پایه احتمالاتی به همراه میزان همبستگی محلی میان توسعه اراضی شهری و متغیرهای مستقل، مقادیر احتمال محلی توسعه شهری را تشکیل داده و در مجموع باعث افزایش دقت پیش‌بینی می‌گردند.

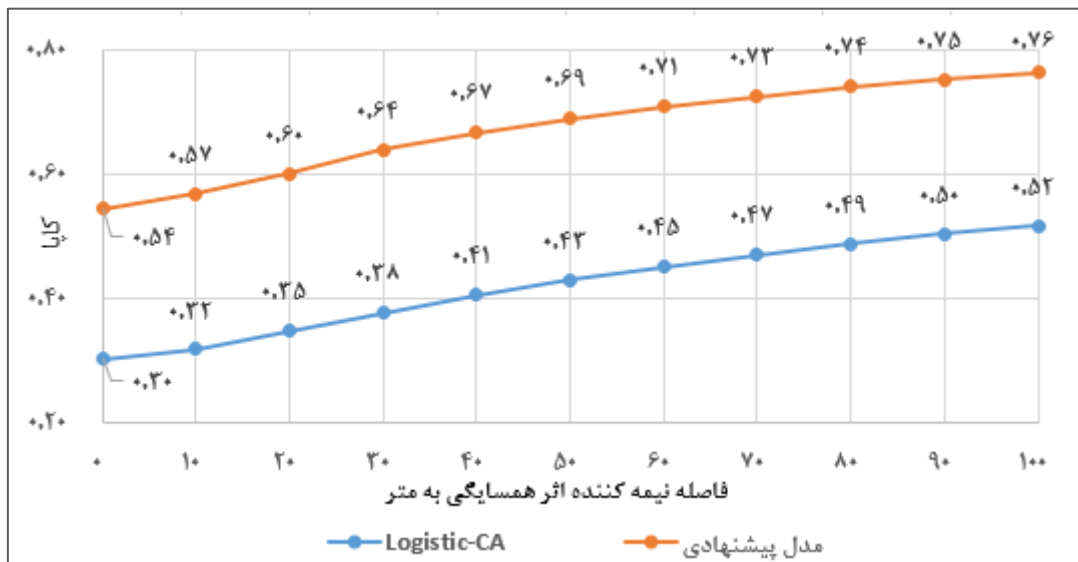
همچنین شکل (۸) دقت شبیه‌سازی دو مدل فوق را بر اساس آماره کاپای فازی برای مقادیر مختلف پارامتر d در تابع افت فاصله نمایش می‌دهد. در حالت $d=0$ مقدار آماره در واقع همان ضریب کاپای مرسوم می‌باشد. همان‌طور که در شکل ذیل مشاهده می‌گردد دقت شبیه‌سازی با ورود مدل GWLR در قوانین تبدیل براساس آماره کاپای فازی نیز از افزایش قابل توجهی برخوردار شده است. همان‌طور که در شکل (۸) مشاهده می‌شود، با محاسبه آماره کاپای فازی بر اساس پارامتر فاصله، دقت افزایش قابل توجهی می‌یابد. به‌عنوان مثال بر اساس مقادیر به‌دست آمده از آماره کاپای فازی می‌توان گفت با فرض تاثیر ۵۰ درصدی کلاس همسایه‌های واقع در فواصل ۲۰، ۵۰ و ۱۰۰ متری سلول مرکزی بر آن، دقت پیش‌بینی مدل پیشنهادی به ترتیب به ۰/۶۰، ۰/۶۹ و ۰/۷۶ می‌رسد. در صورتی که این دقت با اجرای مدل CA با قوانین تبدیل فراگیر به ترتیب معادل ۰/۳۵، ۰/۴۳ و ۰/۵۲ است.

نتایج نشان می‌دهد تعریف قوانین تبدیل به‌صورت محلی با استفاده از مدل GWLR و برآورد محلی ضرایب متغیرهای پیشگو که در واقع بیانگر رابطه محلی

¹ Patch

شده نسبت به یک مدل پایه یکسان (Logistic-CA) می‌باشد؛ عملاً تفاوت‌های موجود در مناطق مطالعاتی، دوره زمانی شبیه‌سازی و عوامل تاثیر گذار بر توسعه اراضی شهری نادیده گرفته شده و امکان مقایسه نتایج آنها با تحقیق حاضر فراهم می‌گردد.

در ادامه، نتیجه به‌دست آمده در تحقیق حاضر با نتایج ۶ تحقیق معتبری که دقت مدل توسعه داده شده در آنها با مدل Logistic-CA مقایسه گردیده‌اند، مورد ارزیابی قرار گرفته است (جدول (۲)). از آنجایی که مبنای مقایسه، میزان افزایش دقت مدل توسعه داده



شکل ۸: نمودار دقت مدل پیشنهادی و مدل Logistic-CA بر اساس آماره کاپای فازی

جدول ۲: مقایسه دقت شبیه‌سازی مدل های توسعه داده شده در ۶ تحقیق معتبر با مدل Logistic-CA

میزان افزایش دقت	دقت مدل Logistic-CA	دقت مدل پیشنهادی	محققین	نام مدل به کار گرفته شده در CA
+0.19	0.34	0.53	لی و لیو (۲۰۰۶) [۲۶]	استدلال مبتنی بر مورد (CBR)
+0.10	0.47	0.57	لیو و همکاران (۲۰۰۸) [۱۳]	کرنل افتراقی فیشر (KFD ^۱)
+0.26	0.44	0.70	یانگ و همکاران (۲۰۰۸) [۱۲]	ماشین بردار پشتیبان (SVM ^۲)
+0.07	0.49	0.56	فنگ و همکاران (۲۰۱۱) [۱۵]	بهینه سازی ازدحام ذرات (PSO ^۳)
+0.07	0.56	0.63	فنگ و لیو (۲۰۱۳) [۱۶]	شبیه‌سازی تبرید (SA ^۴)
+0.02	0.65	0.67	لی و همکاران (۲۰۱۵) [۲۸]	مدل همادی ^۵ (ترکیب درخت رگرسیون و شبکه عصبی)

CA مقایسه شده است، بیشترین افزایش دقت به میزان ۰.۲۶، مربوط به تحقیق یانگ و همکاران (۲۰۰۸) و

همانطور که در جدول فوق مشاهده می‌گردد، در میان تحقیقاتی که مدل پیشنهادی آنها با مدل Logistic-

¹ Kernel Fisher Discriminant

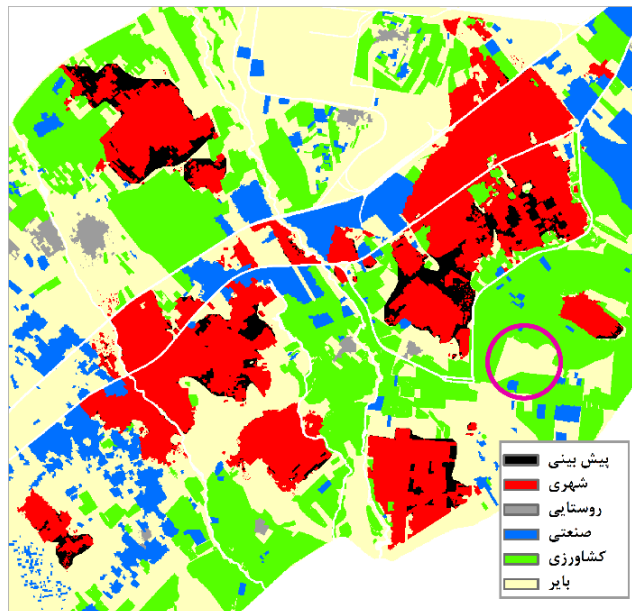
² Support Vector Machines

³ Particle Swarm Optimization

⁴ Simulated Annealing

⁵ Ensemble

شبیه‌سازی را در حالتی که ضریب رشد حاشیه‌ای برابر ۱۰ در نظر گرفته شده است (اعمال ضریب همسایگی در تمامی تکرارها) نشان می‌دهد. همانطور که در شکل مشاهده می‌شود قطعه شهری توسعه یافته مسکن مهر اسلامشهر در این حالت پیش‌بینی نشده است (دایره مشخص شده در شکل (۹)). در صورتیکه با تغییر این ضریب به عدد بهینه ۸، این قطعه توسط مدل قابل پیش‌بینی می‌باشد (شکل (۷-الف)).



شکل ۹: نتیجه حاصل از شبیه‌سازی با لحاظ کردن اثرات همسایگی به طور کامل (ضریب حاشیه ای = ۱۰)

رابطه بین متغیرهای وابسته و مستقل لزوماً در غالب منطقه مطالعاتی پایا نیستند. لذا در این مطالعه از مدل محلی GWLR که بر پایه ناپایداری روابط بین متغیرهای مستقل و وابسته بنا شده است و ترکیب آن با مدل فراگیر LR، جهت تعریف قوانین تبدیل CA استفاده شد.

همچنین در تحقیق حاضر پارامتری نیز تحت عنوان ضریب رشد حاشیه ای تعریف گردید که بر اساس آن سهم رشد خود انگیزه و حاشیه ای در شبیه‌سازی توسعه شهری تعیین و استفاده از آن منجر به پیش‌بینی اراضی شهری که به صورت قطعه‌های

کمترین آن مربوط به تحقیق لی و همکاران (۲۰۱۵) با افزایش دقتی معادل ۰.۰۲ می‌باشد. در این میان تحقیق حاضر با افزایش دقتی معادل ۰.۲۴ نسبت به مدل Logistic-CA در جایگاه مناسبی (رتبه دوم) از لحاظ بهبود دقت پیش‌بینی قرار می‌گیرد.

۵-۴- تاثیر ضریب رشد حاشیه‌ای

نتایج نشان می‌دهد اعمال ضریب رشد حاشیه‌ای تاثیر به‌سزایی در پیش‌بینی قطعات توسعه یافته شهری که به صورت جداگانه توسعه یافته‌اند، دارد. شکل (۹) نتیجه

۶- نتیجه‌گیری

قوانین تبدیل نقش اساسی را در تعیین دقت پیش‌بینی مدل CA ایفا می‌کنند. در صورتی که بتوان قوانین تبدیل را به‌گونه‌ای تعریف نمود که معرف دقیق‌تر روابط میان متغیرهای محیطی باشند، می‌توان انتظار داشت دقت پیش‌بینی نیز متناسب با آن افزایش یابد. به‌علت ماهیت دو حالت توسعه شهری در مطالعات متعددی از مدل فراگیر LR جهت تعریف قوانین تبدیل سلول‌های خودکار شهری استفاده شده است. در مدل فراگیر LR ضرایب برآورد شده برای متغیرهای مستقل در کل منطقه مطالعاتی ثابت فرض می‌شوند، در حالی که

پیش‌بینی شده به صورت قطعی می‌نگرد، درجه تعلق یک سلول به کلاس خاص را بر اساس تابع افت فاصله محاسبه می‌نماید. در نتیجه امکان ارزیابی دقت بر اساس تفرانس های مختلف فاصله در ابعادی فراتر از یک سلول با استفاده از این آماره وجود خواهد داشت. نتایج تحقیق نشان دهنده بهبود دقت پیش‌بینی هنگام به‌کارگیری مدل محلی GWLR در قوانین تبدیل CA در مقایسه با قوانین تبدیل ثابت (مدل Logistic-CA) بر اساس شاخص‌های ضریب کاپا و آماره کاپای فازی می‌باشد. همچنین این افزایش دقت در مقایسه با سایر تحقیقاتی که دقت مدل پیشنهادی خود را نسبت به مدل Logistic-CA سنجیده‌اند، مقدار قابل توجهی می‌باشد.

جداگانه توسعه یافته اند مانند طرح مسکن مهر گردید. لازم به ذکر است گرچه در این تحقیق مقدار بهینه ضریب رشد حاشیه ای بر اساس داده‌های توسعه در دوره زمانی ۸۳-۹۱ به دست آمده است، اما در صورت فقدان داده‌های لازم در مورد رشد خودانگیخته رخ داده در منطقه مورد مطالعه، این ضریب می‌تواند برای پیش‌بینی و مشاهده سناریوهای مختلف رشد خودانگیخته، مستقلاً توسط کاربر نیز تعیین گردد. علاوه بر این در تحقیق حاضر علاوه بر ضریب کاپا که به‌طور معمول برای مقایسه نتایج پیش‌بینی و واقعیت موجود استفاده می‌گردد، از آماره کاپای فازی نیز برای مقایسه دو نقشه استفاده گردید. استفاده از این آماره بر خلاف ضریب کاپای معمول که به موقعیت سلول‌های

مراجع

- [1] R. White, I. Uljee, and G. Engelen, "Integrated modelling of population, employment and land-use change with a multiple activity-based variable grid cellular automaton", *International Journal of Geographical Information Science*, Vol.26, pp. 1251-1280, 2012.
- [2] H. Zhang, Y. Zeng, L. Bian, and X. Yu, "Modelling urban expansion using a multi agent-based model in the city of Changsha", *Journal of Geographical Sciences*, Vol.20, pp. 540-556, 2010.
- [3] X. Li, and A. G-O Yeh, "Modelling sustainable urban development by the integration of constrained cellular automata and GIS", *International Journal of Geographical Information Science*, Vol.14, pp. 131-152, 2000.
- [4] R. M. Itami, "Simulating spatial dynamics: cellular automata theory", *Landscape and Urban Planning*, Vol. 30, pp. 27-47, 1994.
- [5] W. Tobler, "Cellular geography", *Philosophy in geography*, Vol.9, pp. 379-386, 1979.
- [6] R. White, and G. Engelen, "Cellular automata and fractal urban form: a cellular modelling approach to the evolution of urban land-use patterns", *Environment and Planning A*, Vol.25, pp. 1175-1175, 1993.
- [7] R. White, G. Engelen, and I. Uljee, "The use of constrained cellular automata for high-resolution modelling of urban land-use dynamics", *Environment and Planning B*, Vol.24, pp. 323-344, 1997.
- [8] G. Engelen, S. Geertman, P. Smits, and C. Wessels, "Dynamic GIS and strategic physical planning support: a practical application", in: Stillwell S, Geertman S, Openshaw S (Eds.) *Geographical Information and Planning*. Berlin, Springer-Verlag, 1999.
- [9] K. C. Clarke, S. Hoppen, and L. Gaydos, "A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area", *Environment and planning B*, Vol.24, pp.247-261, 1997.
- [10] F. Wu, "Calibration of stochastic cellular automata: the application to rural-urban land conversions", *International Journal of Geographical Information Science*, Vol.16, pp. 795-818, 2002. [65] P. D. Allison, *Logistic regression using SAS®: theory and application*, Second Edition. Cary, NC, SAS Publishing, 2012.

- [11] X. Li, and A. G-O Yeh, "Calibration of cellular automata by using neural networks for the simulation of complex urban systems", *Environment and Planning A*, Vol.33, pp. 1445-1462, 2001.
- [12] Q. Yang, X. Li, and X. Shi, "Cellular automata for simulating land use changes based on support vector machines", *Computers & geosciences*, Vol.34, pp. 592-602, 2008.
- [13] X. Liu, X. Li, X. Shi, S. Wu, and T. Liu, "Simulating complex urban development using kernel-based non-linear cellular automata", *Ecological Modelling*, Vol.211, pp. 169-181, 2008.
- [14] K. Al-Ahmadi, L. See, A. Heppenstall, and J. Hogg, "Calibration of a fuzzy cellular automata model of urban dynamics in Saudi Arabia", *Ecological Complexity*, Vol.6, pp. 80-101, 2009.
- [15] Y. Feng, Y. Liu, X. Tong, M. Liu, and S. Deng, "Modeling dynamic urban growth using cellular automata and particle swarm optimization rules", *Landscape and Urban Planning*, Vol.102, pp. 188-196, 2011.
- [16] Y. Feng, and Y. Liu, "A heuristic cellular automata approach for modelling urban land-use change based on simulated annealing", *International Journal of Geographical Information Science*, Vol.27, pp. 449-466, 2013.
- [17] H. Couclelis, "Cellular worlds: a framework for modeling micro-macro dynamics", *Environment and Planning A*, Vol.17, pp. 585-596, 1985.
- [18] H. Couclelis, "From cellular automata to urban models: new principles for model development and implementation", *Environment and Planning B*, Vol.24, pp. 165-174, 1997.
- [19] J. Luo, and Y. H. Wei, "Modeling spatial variations of urban growth patterns in Chinese cities: The case of Nanjing", *Landscape and Urban Planning*, Vol.91, pp. 51-64, 2009.
- [20] R. I. McDonald, and D. L. Urban, "Spatially varying rules of landscape change: lessons from a case study. *Landscape and Urban Planning*", Vol.74, pp. 7-20, 2006.
- [21] X. Li, Q. Yang, and X. Liu, "Discovering and evaluating urban signatures for simulating compact development using cellular automata", *Landscape and Urban Planning*, Vol.86, pp. 177-186, 2008.
- [22] R. White, and G. Engelen, "High-resolution integrated modelling of the spatial dynamics of urban and regional systems", *Computers, Environment and Urban Systems*, Vol.24, pp. 383-400, 2000.
- [23] H. Wang, S. He, X. Liu, L. Dai, P. Pan, S. Hong, and W. Zhang, "Simulating urban expansion using a cloud-based cellular automata model: A case study of Jiangxia, Wuhan, China", *Landscape and Urban Planning*, Vol.110, pp. 99-112, 2013.
- [24] I. Santé, A. M. García, D. Miranda, and R. Crecente, "Cellular automata models for the simulation of real-world urban processes: A review and analysis", *Landscape and Urban Planning*, Vol.96, pp. 108-122, 2010.
- [25] J. Cheng, and I. Masser, "Understanding spatial and temporal processes of urban growth: cellular automata modelling", *Environment and Planning B*, Vol.31, pp. 167-194, 2002.
- [26] X. Li, and X. Liu, "An extended cellular automaton using case - based reasoning for simulating urban development in a large complex region", *International Journal of Geographical Information Science*, Vol.20, pp. 1109-1136, 2006.
- [27] S. Geertman, M. Hagoort, and H. Ottens, "Spatial - temporal specific neighbourhood rules for cellular automata land - use modelling", *International Journal of Geographical Information Science*, Vol.21, pp. 547-568, 2007.
- [28] X. Li, X. Liu, and P. Gong, "Integrating ensemble-urban cellular automata model with an uncertainty map to improve the performance of a single model",

International Journal of Geographical Information Science, Vol. 29, pp. 762-755, 2015.

- [29] E. Foroutan, "Modeling Urban Growth using the integration of Cellular Automata, Fuzzy and Genetic algorithms ", MSc Thesis, School of Surveying and Geospatial Eng. at The University of Tehran, 2012.
- [30] [30] Y.Chen, X. Li, X. Liu, & B. Ai., "Modeling urban land-use dynamics in a fast developing city using the modified logistic cellular automaton with a patch-based simulation strategy", International Journal of Geographical Information Science, Vol. 28(2), pp. 234-255, 2014.
- [31] D. Z. Sui, and H. Zeng, "Modeling the dynamics of landscape structure in Asia's emerging desakota regions: a case study in Shenzhen", Landscape and Urban Planning, Vol.53, pp. 37-52, 2001.
- [32] B. S. Soares-Filho, G. Coutinho Cerqueira, and C. Lopes Pennachin, "DINAMICA—a stochastic cellular automata model designed to simulate the landscape dynamics in an Amazonian colonization frontier", Ecological Modelling, Vol.154, pp. 217-235, 2002.
- [33] F. Wu, "A parameterised urban cellular model combining spontaneous and self-organising growth", In: Innovation in GIS 7, London, Taylor & Francis, 2000.
- [34] A. M. García, I. Santé, R. Crecente, and D. Miranda, "An analysis of the effect of the stochastic component of urban cellular automata models", Computers, Environment and Urban Systems, Vol.35, pp. 289-296, 2011.
- [35] A. J. Jokar Arsanjani, M. Helbich, W. Kainz, and A. Darvishi Bloorani, "Integration of logistic regression Markov chain and cellular automata models to simulate urban expansion", International Journal of Applied Earth Observation and Geoinformation, Vol.21, pp. 265-275, 2012.
- [36] A. M. García, I. Santé, M. Boullón, and R. Crecente, "Calibration of an urban cellular automaton model by using statistical techniques and a genetic algorithm, Application to a small urban settlement of NW Spain", International Journal of Geographical Information Science, Vol.27, pp. 1593-1611, 2013.
- [37] J. Liao, L. Tang, G. Shao, X. Su, D. Chen, & T. Xu, "Incorporation of extended neighborhood mechanisms and its impact on urban land-use cellular automata simulations", Environmental Modelling & Software, Vol. 75, pp. 163-175, 2016.
- [38] A. M. García, I. Santé, M. Boullón, and R. Crecente, "A comparative analysis of cellular automata models for simulation of small urban areas in Galicia, NW Spain", Computers, Environment and Urban Systems, Vol.36, pp. 291-301, 2012.
- [39] J. Luo, and N. K. Kanala, "Modeling urban growth with geographically weighted multinomial logistic regression", In Proceedings of the Paper presented at the Geoinformatics 2008 and Joint Conference on GIS and Built Environment: The Built Environment and Its Dynamics. Proceedings of SPIE 7144, 2008.
- [40] X. Li, and A. G-O Yeh, "Data mining of cellular automata's transition rules", International Journal of Geographical Information Science, Vol.18, pp. 723-744, 2004.
- [41] C. Brunsdon, A. S. Fotheringham, and M. E. Charlton, "Geographically weighted regression: a method for exploring spatial nonstationarity", Geographical Analysis, Vol.28, pp. 281-298, 1996.
- [42] A. S. Fotheringham, M. E. Charlton, and C. Brunsdon, "Spatial variations in school performance: a local analysis using geographically weighted regression", Geographical and Environmental Modelling, Vol.5, pp. 43-66, 2001.
- [43] C. Brunsdon, J. McClatchey, and D. Unwin, "Spatial variations in the average rainfall–altitude relationship in Great Britain: an approach using geographically weighted regression", International Journal of

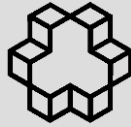
- Climatology, Vol.21, pp. 455-466, 2001.
- [44] M. Cahill, and G. Mulligan "Using geographically weighted regression to explore local crime patterns", *Social Science Computer Review*, Vol.25, pp. 174-193, 2007.
- [45] J. Gao, and S. Li, "Detecting spatially non-stationary and scale-dependent relationships between urban landscape fragmentation and related factors using geographically weighted regression", *Applied Geography*, Vol.31, pp. 292-302, 2011.
- [46] D. Ivajnsič, M. Kaligarič, and I. Žiberna, "Geographically weighted regression of the urban heat island of a small city", *Applied Geography*, Vol.53, pp. 341-353, 2014.
- [47] A. S. Fotheringham, C. Brunsdon, and M. Charlton, *Geographically weighted regression: the analysis of spatially varying relationships*, New York, John Wiley & Sons, 2002.
- [48] M. J. De Smith, M. F. Goodchild, and P. Longley, *Geospatial analysis: a comprehensive guide to principles techniques and software tools*, Leicester, Troubador Publishing Ltd, 2007.
- [49] T. Nakaya, "Geographically weighted generalised linear modelling *Geocomputation: A Practical Primer*. London", Sage Publication, pp. 217-220, 2015.
- [50] H. Shafizadeh-Moghadam, and M. Helbich, "Spatiotemporal variability of urban growth factors: A global and local perspective on the megacity of Mumbai", *International Journal of Applied Earth Observation and Geoinformation*, Vol.35, pp. 187-198, 2015.
- [51] M. Ghamami, "Urban agglomeration of Tehran and around cities", *Center of studies and researches of Iran's architecture and urban planning*, Ministry of Roads and Urban development, 2000 (Persian).
- [52] A. Nazarian, "Urban sprawl of Tehran city and emergence of satellite towns", *Journal of Geographical Researches*, Vol.20, pp. 97-139, 1991 (Persian).
- [53] A. Pourahmad, F. Seifoddini, and Z. Parnoon, "Migration and land use change in Eslamshahr city", *Journal of Geographical Studies of Arid Regions*, Vol. 2(5), pp. 131-150, 2011 (Persian).
- [54] M. Rezvani, "Analysis of city and villages relationships patterns in rural areas of Tehran suburbs", *Journal of Geographical Researches*, Vol.43, pp. 81-94, 2002 (Persian).
- [55] N. Ghasemi, and V. Mohammadi, "Eslamshahr town environmental-legal problems' Evaluation", *Journal of Environmental Science and Technology*, Vol. 16(4), pp. 147-163, 2015 (Persian).
- [56] J. Tavakolinia, Zh. Sajjadi, and K. Shamspuya, "Sprawl of population and activities, Case study: Eslamshahr County", *Journal of Geography*, Vol.39, pp. 120-136, 2013 (Persian).
- [57] G. Mortezaian Abkenar, H. Ghasemi Movaghar, and S. M. Sajjadi, *Assimilation of Tehran province demographic information (1996-2011) based on the state divisions of 2011*, Statistics and information office, General Governor of Tehran province, 2012 (Persian).
- [58] I. Gollini, B. Lu, M. Charlton, C. Brunsdon, and P. Harris, "GWmodel: An R package for exploring spatial heterogeneity using geographically weighted models", *Journal of Statistical Software*, Vol.63, pp. 1-50, 2013.
- [59] S. Chatterjee, and A. S. Hadi, *Regression analysis by example*, John Wiley & Sons, 2015.
- [60] C. Xu, M. Liu, C. Zhang, S. An, W. Yu, and J. M. Chen, "The spatiotemporal dynamics of rapid urban growth in the Nanjing metropolitan region of China", *Landscape ecology*, Vol.22, pp. 925-937, 2007.
- [61] R. G. Congalton "A review of assessing the accuracy of classifications of remotely sensed data", *Remote Sensing of Environment*, Vol.37, pp. 35-46, 1991.
- [62] A. Hagen - Zanker, "An improved Fuzzy Kappa statistic that accounts for spatial

autocorrelation” , International Journal of Geographical Information Science, Vol. 23(1), pp. 61-73, 2009.

[63] A. Hagen-Zanker, “Fuzzy set approach to assessing similarity of categorical maps”, International Journal of Geographical Information Science, Vol. 17(3), pp. 235-249, 2003.

[64] H. Visser, and T. De Nijs, “The map comparison kit”, Environmental Modelling & Software, Vol. 21, pp. 346-358, 2006.

[65] P. D. Allison, Logistic regression using SAS®: theory and application, Second Edition. Cary, NC, SAS Publishing, 2012.



Evaluating the Capability of Geographically Weighted Regression in Improvement of Urban Growth Simulation Performance Using Cellular Automata

Babak Mirbagheri^{*1}, Abbas Alimohammadi²

1- PhD student of GIS, Faculty of Geodesy and Geomatics Engineering, K.N. Toosi University of Technology

2- Associated Professor, Center of Excellence in Geospatial Information Technology, Faculty of Geodesy and Geomatics Engineering, K.N. Toosi University of Technology

Abstract

Geographically Weighted Logistic Regression (GWLR) is a local version of logistic regression (LR) which estimates different relationships between independent and dependent variables at each location. In this research, local model (GWLR) is used for defining CA transition rules and evaluating GWLR capabilities in terms of enhancing urban development prediction accuracy. Also, a new parameter named “Edge Expansion Coefficient” was defined for the determination of tradeoff between two important urban development processes: edge expansion and spontaneous growth. Moreover, in order to assess the prediction accuracy, fuzzy Kappa statistic was applied along with the traditional Kappa coefficient. The developed CA model in this study was run for the prediction of urban development in south west of Tehran metropolitan area during 2004-2013 period. The results of the study showed that, using GWLR model for defining CA’s transition rules, one can significantly increase urban development prediction’s accuracy compared to that of predicted urban development by CA model based on logistic regression (Logistic-CA). The prediction accuracies of the proposed model in this research and the Logistic-CA were 0.54 and 0.30, respectively, as measured by Kappa coefficient. Also, the prediction accuracies of the proposed model were calculated to be 0.68 and 0.76 when measured in terms of fuzzy Kappa statistic with halving distances of 50 and 100 meters in exponential distance decay function, respectively.

Key words: Cellular Automata, Geographically Weighted Regression, Logistic Regression