

مقدمه

یکی از پیامدهای توسعه صنعت، کشاورزی و شهرنشینی افزایش آلودگی آب است. تأمین آب با کیفیت مناسب برای مصارف مختلف از اهمیت بیشتری برخوردار می‌باشد. در هیدرولوژی مدیریت منابع آب حوزه آبخیز، توجه به کیفیت آب و شناسایی منابع آلوده‌کننده آب اهمیت اساسی دارد. انواع کاربری اراضی و الگوی مکانی آنها در سطح حوزه، روی کیفیت آب‌های سطحی و زیرزمینی اثر مستقیم دارد [۲۱]. کاربری اراضی در مفهوم کلی آن به نوع استفاده از زمین در وضعیت موجود گفته می‌شود که در برگیرنده تمامی کاربری در بخش‌های مختلف کشاورزی، منابع طبیعی و صنعت می‌گردد [۱]. چون نوع استفاده از زمین (کاربری اراضی) به دو صورت مثبت و منفی بر کیفیت آب رودخانه تأثیر گذار می‌باشد، لذا تعیین نقش هر یک از انواع کاربری اراضی در پارامترهای کیفیت آب رودخانه‌ها حائز اهمیت می‌باشد.

تاکنون تحقیقات زیادی در خارج از ایران در مورد مدل‌سازی و تحلیل روابط بین کاربری اراضی و پارامترهای کیفیت آب با استفاده از روش‌های آماری پیشرفته انجام شده است که می‌توان به تحقیقات بسنیت و همکاران [۲]، جاروی و همکاران [۱۴]، لیتل و همکاران [۱۷]، محافی و همکاران [۱۸]، رودریگوئز و همکاران [۲۲]، تانگ و همکاران [۲۶]، ویلیامز و همکاران [۳۱]، وولی و همکاران [۳۲] اشاره کرد. به منظور برنامه‌ریزی صحیح مدیریت منابع طبیعی، لازم است اثرات انواع کاربری اراضی روی پارامترهای مختلف کیفیت آب در حوزه‌های آبخیز به طور صحیح مدل‌سازی و درک شوند. امروزه انجام بیشتر مطالعات، به داده‌ها اطلاعات کمی متأثر از مفهوم مکان (فضا) و محیط نیاز می‌باشد. در مدل‌سازی مکانی، میزان و نحوه اثرگذاری پدیده‌های مکانی (مانند کاربری اراضی) دارای اهمیت می‌باشد. نادیده گرفتن اثرات مکانی، منجر به افزایش میزان خطای برآورد در مدل‌سازی می‌شود. تاکنون در ایران به موضوع مدل‌سازی مکانی رابطه بین پارامترهای کیفیت آب و انواع کاربری اراضی حوزه آبخیز توجه جدی نشده است.

در این مقاله، مبانی نظری مدل‌سازی تأثیرات کاربری اراضی روی پارامترهای کیفیت آب با استفاده از مدل رگرسیونی حداقل مربعات معمولی (OLS) و مدل رگرسیون وزنی جغرافیایی (GWR) و ملاک‌های اعتبارسنجی آنها ارائه شده است. در ابتدا به مهمترین پژوهش‌های انجام شده در خارج اشاره می‌شود: جان تو و همکاران

مقدمه‌ای بر روش‌های رگرسیونی چند متغیره OLS و GWR در مدل‌سازی مکانی اثرات کاربری اراضی بر کیفیت آب

مهدی عرفانیان^۱، مریم حسین خواه^۲ و احمد علیچانپور^۳
 تاریخ دریافت: ۱۳۹۲/۵/۲ تاریخ پذیرش: ۱۳۹۲/۶/۱۱

چکیده

کیفیت آب از نظر کشاورزی، شرب، صنعت و امثال آن اهمیت دارد. انواع کاربری اراضی در سطح حوزه آبخیز روی کمیت و کیفیت آب رودخانه از نظر فیزیکی، شیمیایی و بیولوژی تأثیر بسزایی دارند. تاکنون در ایران، مدل‌سازی مکانی اثرات کاربری اراضی روی پارامترهای کیفی آب مورد توجه قرار نگرفته است. با مدل‌سازی مکانی می‌توان نقشه تغییرات مکانی اثرات مثبت یا منفی هر یک از انواع کاربری اراضی را روی پارامترهای کیفیت آب تهیه کرد. بر همین اساس نواحی یا مناطق آلوده‌کننده در سطح حوزه، قابل شناسایی و مدیریت می‌باشد. هدف این مقاله، معرفی روش‌های حداقل مربعات معمولی (OLS) و رگرسیون وزنی جغرافیایی (GWR) می‌باشد. نتایج تحقیقات انجام شده در خارج بیانگر کارایی بالاتر مدل GWR نسبت به OLS می‌باشد. استفاده از مدل GWR در مطالعات و تحقیقات آبخیزداری ایران توصیه می‌شود.

کلمات کلیدی: کیفیت آب، کاربری اراضی، OLS، GWR

۱- نویسنده مسئول، استادیار دانشکده منابع طبیعی دانشگاه ارومیه
 Erfanian.ma@gmail.com

۲- دانشجوی کارشناسی ارشد رشته آبخیزداری دانشگاه ارومیه

۳- استادیار دانشکده منابع طبیعی دانشگاه ارومیه

[۲۸] روابط مختلف مکانی بین کاربری اراضی و کیفیت آب را با استفاده از روش رگرسیون وزنی جغرافیایی و روش حداقل مربعات معمولی در شرق ماساچوست ایالت متحده آمریکا بررسی کردند. آنها از پارامتر یا اندیکاتورهای کیفیت آب (متغیر وابسته) شامل یون‌های کلسیم، منیزیم، سدیم، پتاسیم، کلر، سولفات، نیتروژن-هیدروژن نیتريد (NH_3-N)، نیتريد نیتروژن (NO_2-N)، نترات بعلاوه نیتريت نیتروژن (NO_3-N+NO_2-N)، آمونیاک همراه نیتروژن آلی فسفر و فسفات فسفر (PO_4-P)، باقی مانده خشک (TDS)، هدایت یا رسانایی مخصوص (SC) و درصد انواع کاربری اراضی شامل جنگل، کشاورزی اراضی متروکه، مناطق فاقد پوشش گیاهی، زمین تجاری، مناطق صنعتی، حمل و نقل، معدن کاری، تفریحگاه و مناطق مسکونی (متغیرهای مستقل) استفاده کردند. هر دو مدل OLS و GWR با استفاده از شاخص‌های کیفیت آب به عنوان متغیر وابسته و شاخص‌های کاربری اراضی به عنوان متغیر مستقل ارائه شدند. مقایسه نتایج دو مدل نشان داد که ضرایب مدل آماری در سطح منطقه ثابت نبوده و مدل GWR یک ابزار ساده و مفید در بررسی همبستگی مکانی ارائه می‌کند.

جان تو و همکاران [۲۷] روابط مختلف مکانی بین کاربری اراضی و کیفیت آب در منطقه اوهایوی ایالت متحده آمریکا بوسیله مدل رگرسیون وزنی جغرافیایی را مورد ارزیابی قرار دادند. آنها از ۶ تپ کاربری اراضی به عنوان متغیر مستقل شامل اراضی کشاورزی، جنگل، تفریحی، تجاری، صنعتی و مسکونی و ۱۴ پارامتر کیفیت آب به عنوان متغیر وابسته شامل: کلسیم، منیزیم، سدیم، پتاسیم، کلر، سولفات، نیتريد پتاسیم، نیتروژن آمونیاک، نیتريت- نیتروژن، نترات بعلاوه نیتريت- نیتروژن، فسفر بعلاوه فسفات فسفر، باقیمانده خشک و هدایت الکتریکی ویژه استفاده کردند. نتایج آنها نشان داد که ضرایب مدل GWR در مناطق کشاورزی، مثبت و در مناطق شهرنشین منفی می‌باشد و کیفیت آب در مناطق حومه شهر نسبت به قسمت های مرکزی شهر پایین است. بتانی و همکاران [۲۰] در مدل‌سازی اثرات پوشش زمین و توپوگرافی بر روی پارامترهای کیفیت آب در منطقه پرتلند و کلارک واشنگتن، از مدل‌های رگرسیون وزنی جغرافیایی و حداقل مربعات معمولی استفاده کردند. آنها نتیجه گرفتند که در فصول مرطوب، پارامتر کیفیت آب به پوشش اراضی شهری مربوط است ولی در فصول خشک بیشتر به فاکتورهای توپوگرافی مثل ارتفاع و شیب وابسته هستند. به عقیده آنها میزان خودهمبستگی مکانی مدل GWR نسبت به مدل OLS کمتر می‌باشد. راجش و همکاران [۲۴] عوامل مؤثر در تغییر کاربری اراضی را در میزان گسترش و تغییرات مناظر شهری در دره کاتماندو (نیپال) را با استفاده از مدل‌های رگرسیون وزنی جغرافیایی و حداقل مربعات معمولی بررسی کردند. در تحقیق آنها، روابط آماری بین متغیر وابسته (تغییر کاربری زمین) طی سال‌های ۱۹۹۱ تا ۲۰۰۰ میلادی و متغیرهای مستقل (تغییرات جمعیتی، مساحت اراضی زراعی در

سال ۱۹۹۱، میانگین شیب، مساحت اراضی جنگلی در سال ۱۹۹۱، میانگین فاصله سطوح آب (مانند دریاچه) در سال ۱۹۹۱ و میانگین فاصله از جاده‌ها در سال ۱۹۹۱) با استفاده از مدل‌های OLS و GWR ارائه شد. نتایج آنها نشان داد که مدل GWR از کارایی بیشتری در توضیح تغییرات متغیرها برخوردار بوده و همچنین میزان تغییر کاربری زمین را با باقیمانده استاندارد کمتری نسبت به مدل‌های رگرسیون معمولی (OLS) مدل‌سازی می‌کند. تانگ و همکاران [۲۶] در بررسی رابطه بین کاربری اراضی و کیفیت آب در منطقه اوهایوی ایالت متحده آمریکا به این نتیجه رسیدند که رابطه بین کاربری اراضی با کیفیت آب به دلیل ویژگی‌های حوزه آبخیز و منابع آلودگی در مناطق مختلف یکسان نیست. همچنین اراضی کشاورزی، با نیتروژن کل (TN) همبستگی معنی‌دار مثبت ولی با سدیم و کلسیم همبستگی معنی‌دار منفی می‌باشد. اراضی شهری از جمله مسکونی و تجاری همبستگی مثبتی با چهار شاخص کیفیت آب نشان دادند. با افزایش اراضی کشاورزی، میزان کیفیت آب کاهش یافته و بین افزایش مساحت اراضی جنگلی با مقادیر مشخصه‌های کیفیت آب، همبستگی منفی وجود دارد. فوودی [۶] روابط آماری بین بارندگی و شاخص اختلاف نرمال شده گیاهی (NDVI) را با استفاده از مدل‌های OLS و GWR در شمال افریقا و خاورمیانه در یک دوره ۸ ساله مورد بررسی قرار داد. در تحقیق مذکور، روابط مکانی متفاوتی برای هر سال ارائه شده است. مدل GWR دارای ضریب تعیین حداقل ۰/۹۶۳۳ و برای مدل OLS برابر ۰/۶۷۸۳ بود. ژانگ و همکاران [۳۳] در بررسی تأثیر تغییرات مکانی بر رشد قطر درخت توسط مدل‌های رگرسیون وزنی جغرافیایی و روش حداقل مربعات معمولی به این نتیجه رسیدند که مقدار ضریب تعیین مدل GWR نسبت به OLS بیشتر بوده و مدل GWR باقیمانده استاندارد کمتری نسبت به مدل OLS دارد، بنابراین مدل GWR در بررسی تغییرات مکانی عملکرد بهتری دارد. فوودی [۷] در ارزیابی رابطه بین فراوانی گونه و مجموعه‌های عوامل زیست محیطی (دما، بارش و NDVI) در جنوب افریقا نتیجه گرفت که رابطه معنی‌دار بین انواع متغیرهای مستقل و فراوانی گونه در GWR نسبت به OLS وجود داشته و مدل GWR مفاهیم مدیریت محلی را بهتر ارائه می‌کند.

در آمار فضایی (مکانی)، معمولاً با داده‌هایی روبرو هستیم که جنبه‌های مکانی در آنها مطرح است. لذا قبل از هر چیز باید به تعیین کمیت و مقدار عددی جنبه‌های مکانی پرداخت. همبستگی مکانی ممکن است در متغیرهای کیفیت آب در مکان‌های مختلف حوزه آبخیز وجود داشته باشد. ممکن است مکانی دارای کیفیت آب مشابه با یک مکان نزدیکتر نسبت به مکان‌های دورتر داشته باشد. زیرا مکان‌های نزدیک ممکن است تحت فعالیت‌های انسانی مشترک قرار داشته باشند. به عنوان مثال ویژگی‌های طبیعی از جمله خاک، زمین شناسی، پوشش زمین، و آب و هوا در مکان‌های نزدیک به هم دارای شباهت یا همبستگی مکانی بیشتری می‌باشند.

شاخص‌های کیفیت آب در مکان‌های مختلف یک حوزه و یا حتی در امتداد طولی جریان (رودخانه) دارای تغییرات مکانی می‌باشد. با این حال در بسیاری از تحقیقات از قبیل وولی و همکاران [۳۲] و جاروی و همکاران [۱۴] این نکته در نظر گرفته نشده است. مدل‌سازی اثرات کاربری اراضی روی پارامترهای کیفیت آب توسط مدل‌های رگرسیون چند متغیره OLS و GWR برای اولین بار در ایران توسط نویسندگان مقاله حاضر در تعدادی از حوزه‌های آبخیز استان فارس در حال انجام می‌باشد. متدلوژی و نتایج این تحقیق به عنوان یک گام اساسی برای کارشناسان و محققان کشور در مدیریت کیفیت آب و شناسایی مناطق بحرانی و آلوده کننده آب خواهد بود. بنابراین بخش‌های اجرایی و ترویجی قادر می‌باشند با شناسایی، مدیریت مناطق آلوده‌کننده و آموزش مردم در کاهش اثرات منفی تغییرات کاربری اراضی روی پارامترهای کیفیت آب رودخانه نقش بسیار کلیدی را ایفا کنند. دو فرض اساسی در تحقیق در حال انجام وجود دارد: (۱) انواع کاربری اراضی اثرات مکانی متفاوتی روی پارامتر کیفیت آب دارند و (۲) مدل GWR نسبت به مدل OLS در مدل‌سازی کارایی بالاتری دارد. به منظور برنامه ریزی صحیح مدیریت منابع طبیعی، لازم است تا اثرات انواع کاربری اراضی روی پارامترهای مختلف کیفیت آب در حوزه‌های آبخیز به طور صحیح مدل‌سازی و درک شوند. تهیه مدل‌های رگرسیونی معمولی (OLS) برای یک منطقه برای تخمین هر یک از پارامترهای کیفی آب (متغیر وابسته) و درصد انواع کاربری (متغیرهای مستقل) به عنوان یکی از روش‌های آماری متداول با استفاده از نرم افزارهای آماری مانند SPSS می‌باشد. مدل‌های رگرسیون وزنی جغرافیایی (GWR) یکی از روش‌های نوین برای مدل‌سازی مکانی و آنالیز رگرسیون چند متغیره در محیط GIS می‌باشد. در مدل GWR بر خلاف مدل رگرسیونی معمولی، ضرایب یا پارامترهای مدل آماری در سطح حوزه آبخیز برای یک منطقه مورد مطالعه، ثابت نمی‌باشد. در واقع ضرایب مدل GWR به مختصات مکانی (جغرافیایی) به صورت وزنی وابسته بوده و مقدار و علامت هر یک از ضرایب مدل نسبت به مختصات جغرافیایی نقاط یا پیکسل‌ها در سطح آبخیز تغییر می‌کند. آگاهی از مقدار و علامت هر یک از پارامترها یا ضرایب مدل رگرسیون چند متغیره GWR در مدیریت کاربری اراضی به منظور کاهش اثرات منفی آنها روی پارامترهای کیفیت آب رودخانه‌ها، نقش بسزایی دارد. برای مقایسه و ارزیابی مدل‌های OLS و GWR از ضریب تعیین، معیار اطلاعات اکائیکه و شاخص موران استفاده می‌شود.

مواد و روش‌ها (معرفی مدل‌های آماری)

مدل‌سازی ابزار مهم و مؤثری برای مدیریت کیفیت آب می‌باشد. در مدل‌های رگرسیونی، نوع رابطه بین متغیر وابسته و متغیر یا متغیرهای مستقل بدست می‌آید. هنگامی که مشاهدات و اطلاعات مربوط به متغیر وابسته و متغیرهای مستقل، در مکان یا مختصات

جغرافیایی مختلف جمع‌آوری شده باشند، ارائه یک رابطه رگرسیونی معمولی (OLS) برای توصیف کل مشاهدات به دلیل امکان وجود ناهمگنی مکانی، ممکن است میسر نباشد. برای اینکه مدل بتواند ناهمگنی مکانی را در برآورد پارامترها یا ضرایب مدل لحاظ کند، فرض می‌شود که هر یک از پارامترها یا ضرایب مدل، تابعی از مختصات جغرافیایی می‌باشد. بنابراین اگر ناهمگنی مکانی در نظر گرفته نشود و رابطه‌ای یکسان برای تمام مناطق برآورد شود، تخمینی غیر دقیق از روابط منطقه‌ای خواهد بود. اگر اطلاعات و داده‌های کافی وجود داشته باشد و مرزبندی بین مناطق نیز روشن باشد، ارائه مدلی جداگانه برای هر منطقه مناسب‌تر است. در ارتباط با موضوع این مقاله، مدل‌های OLS و GWR با استفاده از هر یک از پارامترها یا شاخص‌های کیفیت آب (به عنوان متغیر وابسته) و درصد انواع کاربری اراضی (به عنوان متغیر مستقل)، در یک حوزه آبخیز قابل اجرا و مقایسه می‌باشند.

مدل رگرسیونی حداقل مربعات معمولی (SLO)^۱

در بین مدل‌های رگرسیون متداول، روش حداقل مربعات معمولی، ساده‌ترین و مرسوم‌ترین روش است. طرح اولیه این روش معمولاً با SLO نشان داده می‌شود. زیربنای فکری روش حداقل مربعات معمولی این است که ضرایب مدل مقادیری اختیار کنند که مدل رگرسیون نمونه بیشترین نزدیکی را به مشاهدات داشته باشد، به عبارت دیگر کمترین انحراف را از مشاهدات فوق نشان دهد.

در مدل‌سازی مکانی با روش SLO فرض می‌شود که ضرایب یا پارامترهای مدل آماری نسبت به مکان (مختصات جغرافیایی) ثابت می‌باشد. بنابراین مقدار متغیر وابسته که با این مدل تخمین زده می‌شود، برای کل منطقه مورد مطالعه بوده و در نقاط مختلف حوزه نیز مقداری یکسان را تخمین می‌زند که به عنوان نقطه ضعف این روش در مدل‌سازی مکانی محسوب می‌شود. مدل رگرسیونی خطی ساده یک متغیره به شکل زیر می‌باشد

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad (1)$$

که y متغیر وابسته (برآورد شده)، x متغیر مستقل (برآورد کننده)، خطا ε_i یا انحراف مدل در برآورد، β_0 و β_1 پارامترها یا ضرایب مدل می‌باشند. برای تمام سطح حوزه آبخیز، مقادیر β_0 و β_1 ثابت فرض می‌شود. مدل آماری SLO و ماتریس تخمین ضرایب مدل، با روابط زیر بیان می‌شود:

$$y = X\beta + \varepsilon \quad (2)$$

$$\hat{\beta} = (X^T X)^{-1} X^T y \quad (3)$$

که در آن T ترانهاده ماتریس، $(X^T X)^{-1}$ معکوس ماتریس واریانس-کواریانس و X ماتریس متغیرهای مستقل می‌باشد. ضرایب مدل رگرسیونی چند متغیره SLO در سراسر مکان ثابت است. با استفاده

1- Ordinary Least Squares

از این مدل امکان تهیه نقشه تغییرات مکانی پارامترها یا ضرایب مدل وجود ندارد. بعلاوه، این مدل با نرم افزار SIGcrA ناسازگار بوده و همبستگی مکانی را در نظر نمی‌گیرد.

مدل رگرسیونی وزنی جغرافیایی (GWR)^۱

تحلیل‌های کلاسیک آماری همانند همبستگی و رگرسیون، از جمله روش‌های معمول برای مدل‌سازی اثرات کاربری اراضی روی پارامترهای کیفیت آب می‌باشند. در روش‌های کلاسیک، مقادیر متغیر وابسته و مستقل در سطح وسیعی از حوزه آبخیز جمع‌آوری و تحلیل می‌شوند لذا ضرایب خطای ناشی از دخالت عامل فضا (مکان) در روش‌های آماری متداول مانند OLS بالا می‌باشد. یکی از روش‌های جدید برای دستیابی به دقت بالاتر در تحلیل روابط متأثر از مکان، روش رگرسیون وزنی جغرافیایی می‌باشد [۱۰]. این روش برای نخستین بار در دهه اخیر توسط فوترینگام، چالتون و براندسون، اساتید دانشگاه بریتانیا معرفی گردید و نخستین بار در سال ۱۹۹۱ برای بررسی عوامل تأثیرگذار در قیمت خانه‌ها در لندن مورد استفاده قرار گرفت. این روش کاربردهای متنوعی از جمله آشکار کردن و تحلیل متغیرها در مقیاس محلی داشته و مدل‌سازی مکانی در رشته‌هایی مانند برنامه‌ریزی شهری و منطقه‌ای، محیط‌شناسی، ژئوماتیک، جغرافیا و غیره مورد استفاده قرار گرفته است. بر اساس نتایج تحقیقات براندسون و همکاران [۳]، فوترینگام و همکاران [۹]، فوترینگام و همکاران [۱۰]، فوترینگام و همکاران [۱۴] می‌توان گفت که مدل GWR روشی جدید برای مدل‌سازی فرایندهای ناهمگن مکانی است. ناهمگنی مکانی بیانگر این است که در هر نقطه (مختصات جغرافیایی) رابطه‌ای متفاوت بین متغیر وابسته و مستقل بدلیل وابستگی پارامترها یا ضرایب مدل نسبت به مکان وجود دارد. روش رگرسیونی وزنی جغرافیایی بر این ایده استوار است که پارامترها یا ضرایب مدل را می‌توان در هر نقطه‌ای از فضا یا مکان مورد مطالعه برآورد کرد. در این روش برای برآورد پارامترهای مدل در هر نقطه از مشاهدات اطراف آن نقطه استفاده می‌شود اما به مشاهدات نزدیک وزن بیشتر و به مشاهدات دورتر، وزن کمتری داده می‌شود. اگر فاصله بین مشاهده و نقاط مرجع افزایش یابد، این وزن نیز کاهش می‌یابد [۲۳]. در مدل‌های رگرسیونی وزنی جغرافیایی، پنجره جستجو یا کرنل وقتی ثابت فرض می‌شود که توزیع مشاهدات حول نقاط مرجع در گستره جغرافیایی مورد مطالعه یکسان نباشد [۲۳]. در این صورت اگر تابعی یکسان برای وزن‌دار کردن مشاهدات اطراف نقاط مرجع به کار گرفته شود، این احتمال وجود دارد که برای برخی نقاط در سطح حوزه، از اطلاعاتی کمتر استفاده شده و در نهایت یک مدلی غیر دقیق بدست آید. حتی ممکن است برای برخی نقاط که تجمع مشاهدات حول آن بیشتر است از اطلاعات زیاد و نامرتب استفاده شود و باز از دقت مدل کاسته شود. برای مقابله با این احتمال در روش رگرسیون

وزنی جغرافیایی، پهنای باند برای محاسبه تابع وزن دهی $W_i(u)$ با پراکندگی مشاهدات حول نقطه مرجع تطبیق داده می‌شود. به عبارت دیگر چنانچه مشاهدات پراکنده باشند، پهنای باند بیشتری در نظر گرفته می‌شود و اگر مشاهدات حول نقطه مرجع متراکم باشند، پهنای باند کمتری انتخاب می‌شود [۲۳]. در واقع روش رگرسیون وزنی جغرافیایی همان روش حداقل مربعات معمولی است با این تفاوت که به مشاهدات بر اساس مکان یا مختصات مکانی آنها نسبت به نقاط مرجع وزن داده می‌شود. هر پارامتر یا ضریب مدل GWR دارای یک علامت و مقدار می‌باشد. اگر علامت یک ضریب، مثبت باشد پس افزایش مقدار متغیر مستقل باعث افزایش متغیر وابسته خواهد شد. اگر علامت آن منفی باشد، آن متغیر دارای اثر کاهش‌دهنده روی متغیر وابسته می‌باشد. مدل رگرسیون چند متغیره خطی GWR به صورت زیر می‌باشد:

$$y_i(u) = \beta_{0i}(u) + \beta_{1i}(u)x_{1i} + \beta_{2i}(u)x_{2i} + \dots + \beta_{mi}(u)x_{mi} \quad (4)$$

در صورتیکه مؤلفه وزن مشاهدات در معادله رگرسیون وارد شود، رابطه بردار پارامترها یا ضرایب بتا، به یک رابطه رگرسیونی وزنی جغرافیایی تبدیل می‌شود و مشاهدات نزدیک نسبت به مشاهدات دورتر وزن بیشتری دارند [۲۵]. ضرایب یا پارامترهای مدل با رابطه زیر تخمین زده می‌شوند:

$$\hat{\beta}(u) = (X^T W(u) X)^{-1} X^T W(u) \quad (5)$$

که $W(u)$ ماتریس وزن (ماتریس مربع) در موقعیت u (مختصات متریک)، X^T ترانهاده ماتریس متغیرهای مستقل (X) می‌باشد. ماتریس وزن تابع مختصات متریک u با استفاده از رابطه‌ای زیر قابل تخمین می‌باشد:

$$W_i(u) = e^{-0.5(di(u)/h)^2} \quad (6)$$

که $W_i(u)$ وزن جغرافیایی مشاهدات i در موقعیت مکانی u ، $d_i(u)$ اندازه فاصله مکانی بین مشاهدات i و h پهنای باند انتخابی بر حسب متر می‌باشد [۳].

مشخصات مدل رگرسیونی GWR شامل موارد زیر می‌باشد: (۱) پارامترهای یا ضرایب مدل نسبت به مکان تغییر می‌کنند، (۲) قابلیت ارائه نتایج مدل‌سازی به صورت نقشه‌های مکانی وجود دارد، (۳) سازگار با نرم افزار ArcGIS است و (۴) همبستگی مکانی را در نظر می‌گیرد.

معیارهای کارایی مدل‌های رگرسیونی OLS و GWR

برای مقایسه میزان اعتبار یا کارایی مدل‌های رگرسیونی چند متغیره، از معیارهای زیر برای مقایسه نتایج آنها استفاده می‌شود:

۱- ضریب تعیین (R^2)

این ضریب، میزان درصد واریانس متغیر وابسته که توسط

1- Geographically Weighted Regression

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{i,j} z_i z_j}{S_0 \sum_{i=1}^n z_i^2} \quad (10)$$

که w_{ij} وزن مکانی بین دو متغیر i و j تعداد کل صفات (متغیرها)، S_0 مجموع تمام وزن‌های مکانی، z_i انحراف معیار متغیر i از میانگین و z_j انحراف معیار متغیر j از میانگین می‌باشد. مقدار آماره موران I از -1 و $+1$ متغیر است. هر چه مقدار I به صفر نزدیکتر باشد، بیانگر الگوی تصادفی است. براساس تحقیق ایشواو و همکاران [۱۳]، مقدار نزدیک به 1 نشان می‌دهد که نواحی با ارزش‌های مشابه دارای الگوی خوشه‌ای (Clustered) هستند و مقدار -1 نشان می‌دهد که نواحی با ارزش‌های غیرمشابه (Dispersed) در کنار یکدیگر قرار دارند.

نتیجه‌گیری

در مقاله حاضر، بر ضرورت استفاده از مدل GWR در مطالعات و تحقیقات آبخیزداری ایران تاکید شده است. همانطور که قبلاً گفته شد، پژوهش جامع با عنوان مدل‌سازی مکانی اثرات کاربری ارضی روی پارامترهای کیفیت آب در تعدادی از حوزه‌های آبخیز استان فارس توسط نویسندگان مقاله در حال انجام است. لذا در این مقاله ترویجی، مدل‌های رگرسیونی چند متغیره OLS و GWR برای افزایش دانش کارشناسان و محققان ایران در مدیریت جامع حوزه‌های آبخیز معرفی شدند. به عنوان مثال، روابط آماری بین متغیرهای مستقل (درصد انواع کاربری ارضی) و متغیر وابسته (پارامترهای کیفیت آب) با استفاده از مدل OLS و GWR در حوزه‌های آبخیز دارای آمار بلند مدت کیفیت آب و نقشه معتبر پوشش ارضی - کاربری ارضی قابل بررسی است.

این مدل به دلیل ماهیت کل نگر، در آنالیز رابطه میان متغیرهای مستقل و وابسته از دقت کافی برخوردار نیست. اگر رابطه بین متغیرهای مستقل و وابسته در بخشی از منطقه مورد مطالعه مثبت و در بخشی دیگر منفی شود، مدل رگرسیون معمولی OLS قادر به تشخیص رابطه دقیق بین متغیرهای مستقل و وابسته نمی‌باشد. این مدل به دلیل ماهیت کل نگر، در آنالیز رابطه میان متغیرهای مستقل و وابسته از دقت کافی برخوردار نیست. رگرسیون وزنی جغرافیایی زمانی که همبستگی مکانی بین متغیرها وجود داشته باشد از اهمیت و کارایی زیادی برخوردار است. پارامترهای این مدل، برای هر نقطه از نمونه (پیکسل) بدست می‌آید که مقدار و علامت ضرایب یا پارامترهای مدل در پیکسل‌ها یا نقاط مختلف حوزه، متفاوت می‌باشد. لذا نتایج مدل را می‌توان در محیط GIS به صورت نقشه‌های رستری نمایش داد. علامت منفی ضریب یا پارامتر مدل برای یک متغیر مستقل مشخص، بیانگر نقش کاهنده آن متغیر روی مقدار متغیر وابسته است. علامت مثبت نیز بیانگر اثر افزایشی آن روی مقدار متغیر وابسته است. رگرسیون وزنی جغرافیایی با

متغیر(های) مستقل تبیین می‌شود را بیان می‌کند. به عبارت دیگر با محاسبه این ضریب می‌توان گفت که چند درصد از کل واریانس Y توسط متغیر (های) مستقل X بیان می‌شود. مقدار عددی این ضریب از صفر تا یک تغییر می‌کند. مقدار صفر یعنی استفاده از متغیر(های) مستقل در برآورد متغیر وابسته هیچ نقشی ندارد و مقدار یک بیانگر تخمین 100 درصد واریانس متغیر وابسته توسط متغیر(های) مستقل می‌باشد. اگر انحراف معیار متغیرهای X و Y به ترتیب به صورت S_x و S_y باشد و کواریانس آنها با علامت Cov_{yx} نشان داده شود، ضریب تعیین از رابطه زیر قابل محاسبه است:

$$R^2 = \frac{S_{xy}^2}{S_{xx} S_{yy}} \quad (7)$$

۲ - روش معیار اطلاعات آکائیکه (AIC)

معیار اطلاعاتی آکائیکه، معیاری برای سنجش میزان کارایی نسبی است و نشان می‌دهد که استفاده از یک مدل آماری به چه میزان باعث از دست رفتن اطلاعات می‌شود. به عبارت دیگر، این معیار تعادلی میان دقت مدل و پیچیدگی آن برقرار می‌کند. بر اساس تحقیقات فوودی و همکاران [۷]، وانگ و همکاران [۳۰]، مقدار کم این معیار بیانگر این است که مقدار تخمین زده شده توسط مدل، به مقدار مشاهده‌ای یا واقعیت زمینی نزدیک‌تر است. اگر اختلاف AIC بین دو مدل کمتر از عدد ۳ (به عنوان حد آستانه) باشد می‌توان نتیجه گرفت که که دو مدل وضعیت تقریباً یکسانی دارند [۸]. ذکر این نکته ضروری است که روش استاندارد برای انتخاب حد آستانه AIC وجود ندارد. معیار تصحیح شده آکائیکه از رابطه ۸ بدست می‌آید:

$$AICc = AIC + \frac{2k(k+1)}{n-k-1} \quad (8)$$

$$AIC = 2k - 2h(L) \quad (9)$$

که K تعداد پارامتر مدل و L مقدار حداکثر درست‌نمایی می‌باشد.

- شاخص موران (Moran's I)

یکی از آزمون‌هایی متداول برای تشخیص خود همبستگی مکانی، آزمون موران می‌باشد که برای تشخیص درجه خوشه‌بندی یا پراکندگی کاربرد دارد. مقدار آماره موران نشان می‌دهد که نواحی مجاور در یک لایه رستری، دارای ارزش‌های مشابه و یا غیرمشابه هستند. آماره I براساس مفهوم کواریانس و وزن‌دهی براساس میزان مجاورت نقاط نمونه (مشاهدات زمینی)، به صورت زیر محاسبه می‌شود:

47: 431-443.

5- Fassnacht., K.S, Gower., S.T, MacKenzie., M, Nordheim., E. and Lillesand., T. 1997. Estimating leaf area index of North Central Wisconsin forests using the Landsat thematic mapper. *Remote Sensing of Environment*, 61: 229-245.

6- Foody., G.M. 2003. Geographical weighted as a further refinement to regression modeling: an example focused on the NDVI-rainfall relationship. *Remote Sensing of Environment*, 88: 283-293.

7- Foody., G.M. 2004. Spatial non-stationarity and scale dependency in the relationship between species richness and environmental determinants for the sub-Saharan endemic avifauna. *Global Ecology and Biogeography*, 13: 315-320.

8-Fotheringham., A.S, Brunsdon., C. and Charlton., M. 2002. Geographically Weighted Regression: the analysis of spatially varying relationships, Chichester: Wiley.

9-Fotheringham., A.S, Brunsdon., C. and Charlton., M. 1996. The geography of parameter space: an investigation of spatial non-stationarity. *International Journal of Geographical Information Systems*, 10: 605-627

10-Fotheringham., A.S, Charlton., M.E. and Brunsdon., C. 1997. Spatial variations in school performance: a local analysis using geographically weighted regression. *Geogr Environ Model*, 5: 43-66.

11-Fotheringham., A.S. Charlton., M. and Brunsdon., C. 1997. Two techniques for exploring non-stationarity in geographical data. *Geographical Systems*, 4: 59-82

12-Gerulo., G.R, Hall., R.J, Franklin., S.E. and Smith., L. 2002. Empirical relations between Landsat TM spectral response and forest stands near Fort Simpson, Northwest Territories, Canada. *Canadian Journal of Remote Sensing*, 28(1): 68-79.

13-Ishizawa., H. and Stevens., G. 2007. Non-English language neighborhoods in Chicago, Illinois. 200(36): 1042-64.

14-Jarvie., H.P, Oguchi., T. and Neal., C. 2002.

تولید داده‌های مکانی این امکان را بوجود می‌آورد تا تغییر مکانی در روابط بین متغیرها مورد بررسی قرار گیرد. ارائه نقشه‌های رستری مخالف از مقدار ضرایب مدل، در شناسایی مناطق با اثرات منفی یا مثبت معنی دار روی هر یک از پارامترهای کیفیت آب، بسیار با ارزش و کاربردی می‌باشد.

رهیافت ترویجی و پیشنهادات

تحقیقات انجام شده در خارج نشان می‌دهد که مدل GWR یک ابزار ساده و مفید در تشخیص میزان همبستگی مختلف مکانی می‌باشد. مدل‌سازی مکانی اثرات کاربری اراضی روی پارامترهای کیفیت آب به منظور مدیریت حوزه‌های آبخیز در مقیاس منطقه‌ای، و حتی جهانی حائز اهمیت است. مدل‌سازی مکانی با استفاده از روش GWR یک گام اساسی برای کارشناسان و محققان کشور در مدیریت کیفیت آب، شناسایی مناطق بحرانی یا منابع آلوده کننده آب تلقی می‌شود. با کاربرد این مدل در هر حوزه آبخیز، بخش‌های تحقیقاتی، اجرایی و ترویجی قادر خواهند بود تا با اجرای طرح مناسب آبخیزداری، مدیریت برنامه تغییرات کاربری اراضی، ارائه روش‌های مناسب بیولوژیکی و برنامه‌های آموزشی و ترویجی برای مشارکت مؤثر آبخیزنشینان در کاهش اثرات منفی تغییرات کاربری اراضی، نقش اساسی در کاهش آلودگی کیفیت آب رودخانه‌ها داشته باشند. در پایان به کلیه محققان ایران توصیه می‌شود که در مدل‌سازی مکانی حوزه‌های آبخیز مانند اثرات کاربری اراضی روی کیفیت آب، زمین لغزش و فرسایش خاک، از متدولوژی ارائه شده در این مقاله استفاده کنند.

منابع

1-Ahmadi., R. 1996. Intensification of land use in creating a mass movement in the forest, Master's thesis, Faculty of Humanities, Tarbiat Modarres University, 12p. (in Persian)

2-Basnyat., P. Teeter., L.D. Flynn., K.M. and Lockaby., BG. 1999. Relationship between landscap characteristics and nonpoint source pollution inpoint to coastal estuaries. *Environ Manag*, 23: 539-49.

3-Brunsdon., C. Fotheringham., A.S and Charlton., M. 1996. Geographically weighted regression: a method for exploring spatial non-stationarity. *Geographical Analysis*, 28(4): 281-298

4-Brunsdon., C. Fotheringham, S. and Charlton., M. 1998. Geographically weighted regression modeling spatial non-stationarity. *The Statistician*,

- 24-Thapa., R.B. and Murayama., Y. 2009. Land Use Change Factors in Kathmandu Valley: A GWR Approach. Japan, University of Tsukuba.
- 25-Tobler., W.R. 1970. A Computer movie simulating urban growth in the Detroit region. *Economic Geography*, 46(2): 234-24.
- 26-Tong., S.T.Y. and Chen., W. 2002. Modeling the relationship between land use and surface water qualities. *Journal of Environ Manage*, 66: 377-93.
- 27-Tu., j. 2011. Spatially varying relationships between land use and water quality across an urbanization gradient explored by geographically weighted regression. *Appl*. 31: 376-392.
- 28-Tu., J. and Xia., ZG. 2008. Examining spatially varying relationships between land use water qualities using geographically weighted regression: model design and evaluation. *Science of the Total Environment*, 407: 358-378.
- 29-Turner., D.P, Cohen., W.B, Kennedy., R.E, Fassnacht., K.S. and Briggs., J.M. 1999. Relationships between leaf area index and Landsat TM spectral vegetation indices across three temperate zone sites. *Remote Sensing of Environment*, 70: 52-68.
- 30-Wang., Q, Ni., J. and Tenhunen., J. 2005. Application of a geographically-weighted regression analysis to estimate net primary production of Chinese forest ecosystems. *Glob Ecol Biogeogr*, 14: 379-93.
- 31-Williams., M, Hopkinson., C, Rastetter., E, Vallion., J. and Claessens., L. 2005 Relationships of land use and stream solute concentrations in the Ipswich River basin, Northeastern Massachusetts. *Water Air Soil Pollution*, 161: 55-74.
- 32-Woli., K.P, Nagumo., T, Kuramochi., K. and Hatano., R. 2004. Evaluating river water quality through land use analysis and N budget approaches in livestock farming areas. *Science Total Environment*, 329: 61-74.
- 33-Zhang., L. and Shi., H. 2004. Local modeling of tree growth by geographically weight regression. *Forest Science*, 50: 225-244.
- Exploring the linkages between river water chemistry and watershed characteristics using GIS- based catchment and locality analyses. *Reg Environ change*, 3: 36-50.
- 15-Jedary Eyvazi., G, Mogimy., I, Mohammadi., M.H. and Isaei, A. 2011. (Case Study: Kor River and Lake Doroudzan). *Journal of Geography and Environmental Planning*. 21th edition. (In Persian).
- 16-Jenifer., L. Bruce Michelle., M, Green., w. and Suzan., Z.J. 2004. *Water Resources Investigations In Wisconsin*. US Geological Survey.
- 17-Little., J.L, Saffran., K.A. and Fent., L. 2003. Land use and water quality relationships in the lower Little Bow River Watershed- Alberta, Canada. *Water Qual Res J Can*, 38: 563-84.
- 18-Mehaffey., M.H, Nash., M.S, Wade., T.G, Ebert., D.W, Jones., K.B. and Rager., A. 2005. Linking land cover and water quality in New York City's water supply watersheds. *Environ Monit Assess*, 107: 29-44
- 19-Montgomery., D.C, Peck., E.A. and Vining., G.G. 2001. *Introduction to Linear Regression Analysis*. 3 edition. John Wiley & Sons, Inc, New York.
- 20-Part., B. and Chang., H. 2012. Effect of land cover topography and built structure on seasonal water quality at multiple spatial scales. *Journal of Hazardous Materials*, 209-210: 48-58.
- 21-Riky., M. 2001. *Kash Plain Groundwater Quality*. Master's thesis in the field of groundwater the Department of Irrigation and Reclamation Engineering Faculty of Natural Resources Tehran University. (in Persian).
- 22-Rodrigues., W, August., P.V, Wang., Y, Paul., J.F, Gold., A. and Rubinstein., N. 2007 Empirical relationships between land-use/cover estuarine conditions in the Northeastern United States. *Landsc Ecol*, 22: 403-17.
- 23-Souri., D. Moniry Javid., S. Housing pricing model, using geographically weighted regression method.