نشریه آب و خاک (علوم و صنایع کشاورزی)
جلد ۲۶، شماره ۳، مرداد – شهریور ۱۳۹۱، ص. ۷۵۲-۷۴۳

Journal of Water and Soil
Vol. 26, No. 3, Jul-Aug 2012, p. 743-752

رگرسیون وزن دار جغرافیایی: روشی برای ترسیم نقشه های هم بارش در استان گیلان

خلیل قربانی[۱]

تاریخ دریافت:۹۰/۴/۲۲
تاریخ پذیرش:۹۰/۱۱/۱۰

چکیده

تاکنون روش‌های زیادی برای درون‌یابی و ترسیم نقشه‌های هم بارش ارائه و نتایج متفاوتی از آنها در مناطق مختلف گزارش شده است. روش رگرسیون وزن‌دار جغرافیایی یکی از روش‌هایی است که اخیرا مورد استقبال قرار گرفته است. این روش برای ارزیابی غیریکنواختی متغیر وابسته با برقراری رگرسیون‌های موضعی مناسب است. تحقیقی بر این اساس صورت گرفت تا تغییرات مکانی بارش سالانه در استان گیلان (که تحت تأثیر شرایط محیطی مختلف به‌طور غیریکنواخت توزیع می‌شود) مورد بررسی قرار گیرد. نتایج روش رگرسیون وزن‌دار جغرافیایی با روش‌های معمول درون‌یابی شامل چندجمله‌ای سراسری و موضعی، اسپلاین، وزن عکس فاصله‌ها، کریجینگ و کوکریجینگ مقایسه شد. آمار میانگین ده ساله بارش سالانه ۱۸۵ ایستگاه هواشناسی واقع در استان گیلان و مجاورت آن استفاده و تغییرات مکانی بارش به کمک پارامترهای دیگر مانند ارتفاع از سطح دریا و موقعیت نقاط نسبت به دریا، الگوسازی شد. سپس بهترین روش درون‌یابی بارش در این منطقه معرفی شد. برای ارزیابی دقت هر یک از این روش‌ها از فن اعتبارسنجی متقابل استفاده شد. نتایج به‌دست آمده نشان داد که روش رگرسیون وزن‌دار جغرافیایی با RMSE=147 دارای کمترین مقدار خطا بود و نتایج آن با روش کریجینگ که با RMSE=187 در رتبه دوم قرار دارد اختلاف زیادی داشت. بر این اساس مناسب‌ترین روش برای ترسیم هم‌بارش در استان گیلان روش رگرسیون وزن‌دار جغرافیایی معرفی می‌شود.

واژه های کلیدی: درون‌یابی، رگرسیون وزن‌دار جغرافیایی، کریجینگ، گیلان، هم‌بارش

مقدمه

پارامترهای هواشناسی به‌دلیل هزینه‌بر بودن نمونه‌برداری، به‌صورت نقطه‌ای یا ایستگاهی اندازه‌گیری‌می‌شوند. اما ضرورت ایجاب می‌کند این اطلاعات اندازه‌گیری شده در ایستگاه‌ها برای تحلیل‌های منطقه‌ای، به منطقه مطالعاتی تعمیم داده شوند و توزیع آنها در پهنه جغرافیایی ترسیم شود. روش‌های مختلفی به این منظور توسط محققین مختلف مورد بررسی قرار گرفته است که می‌توان به روش‌های درون‌یابی اشاره کرد. روش‌های درون‌یابی با استناد به این قانون جغرافیایی که هر پدیده‌ای مرتبط است با پدیده دیگری اما پدیده‌های نزدیک نسبت به پدیده‌های دور بهم وابسته‌تراند از یکسری نقاط با مقادیر معلوم در اطراف نقاط با مقادیر مجهول استفاده کرده و مقادیر آنها را تخمین می‌زنند (۱۴). اما چگونگی مشارکت نقاط معلوم در تخمین مقادیر مجهول سبب می‌شود تا روش‌های درون‌یابی به دودسته سراسری[۲] و موضعی[۳] تقسیم‌بندی شوند. کلیه نقاط مشاهده‌ای در روش سراسری در محاسبات نقطه یا نقاط مجهول دخالت داده می‌شوند. در حالیکه در روش موضعی فقط از نقاط مشاهده‌ای موجود در یک محدوده (همسایه نزدیک یا پنجره) استفاده می‌شود. سپس، کلیه نقاط مجهول در محدوده موردمطالعه با حرکت پنجره‌ها به‌صورت گام به گام در کل پهنه تخمین زده می‌شوند.

برخی از روش‌های درون‌یابی با برقراری روابط آمار کلاسیک بین مقادیر نقاط معلوم اقدام به تخمین مقادیر مجهول می‌کنند. آن‌ها را درون‌یاب‌های قطعی می‌نامند. اما برخی از روش‌ها با استفاده از ویژگیهای وابستگی فضایی متغیر مورد نظر اقدام به تخمین مقادیر نقاط مجهول می‌کنند. آن‌ها را درون‌یاب‌های زمین آماری می‌نامند. مطالعات زیادی در زمینه مقایسه روش‌های مختلف درون‌یابی در ترسیم نقشه‌های هم‌مقدار انجام شده است. مهدیزاده (۷) از روش‌های زمین آماری کریجینگ، کوکریجینگ، اسپلاین و روش وزنی عکس

۱- استادیار گروه مهندسی آب، دانشگاه علوم کشاورزی و منابع طبیعی گرگان
Email: ghorbani.khalil@yahoo.com

2 -Global
3 - Local

فاصله[1] جهت درون‌یابی داده‌های دما و بارندگی ماهانه و سالانه در حوضه آبریز دریاچه ارومیه استفاده کرد. نتایج ایشان در مورد درون‌یابی دما نشان داد که روش اسپیلاین توان دو و با متغیر کمکی در مقایسه با سایر روش‌ها از دقت بالاتری برخوردار است. روش‌های کوکریجینگ، کریجینگ و وزن عکس فاصله در مراتب بعدی اهمیت قرار می‌گیرند. روش اسپلاین توان دو و بدون متغیر کمکی را نیز برای داده‌های بارندگی پیشنهاد کرد. سپس روش‌های کریجینگ معمولی و وزن عکس فاصله اولویت دارند. اقدسی (۱) با مقایسه چهار روش زمین آماری کریجینگ و کوکریجینگ و روش وزنی عکس فاصله و اسپلاین جهت درون‌یابی داده‌های بارندگی روزانه و ده روزه و ماهانه و سالانه دشت برخوار اصفهان (منطقه‌ای با آب و هوای بیابانی خشک) به این نتیجه رسید که روش اسپلاین با توان دو و بدون متغیر کمکی دارای کمترین خطای درون‌یابی است. مقدار خطا از مقیاس زمانی روزانه به سالانه کاهش می‌یابد. جامعی با مقایسه روش‌های درون‌یابی وزن‌عکس فاصله، اسپلاین، کریجینگ و کوکریجینگ در برآورد منطقه‌ای تبخیروتعرق مرجع در استان خوزستان به این نتیجه رسید که روش کریجینگ بهترین روش درون‌یابی داده‌های تبخیروتعرق مرجع در مقیاس‌های ماهانه و سالانه است (۲). نادی (۸) نشان داد روش پیوند خطی گرادیان خطی و غیر خطی بهترین روش برآورد داده‌های بارندگی ماهانه در منطقه جنگلی چهارباغ گرگان است. روش‌های وزنی عکس فاصله و کریجینگ با روند خارجی در رتبه های دوم و سوم قرار دارند. آتوره و همکاران (۹) سه روش درون‌یابی وزن عکس فاصله، کریجینگ با روند خارجی و شبکه‌های عصبی مصنوعی را برای میان یابی ۱۲ شاخص اقلیمی و ۹ شاخص زیست اقلیمی در منطقه لاتزیو ایتالیا بررسی کردند. آنها از متغیرهای عوارض سطح زمین (شیب ،جهت و...) و موقعیت جغرافیایی به عنوان متغیرهای مستقل استفاده نمودند و به این نتیجه رسیدند که روش کریجینگ با روند خارجی در ۱۷ مورد و روش‌های شبکه عصبی مصنوعی و روش وزنی عکس فاصله به ترتیب در سه و یک مورد بهترین روش‌های برآورد بودند.

استفاده از روش‌های رگرسیونی نیز در پهنه‌بندی پارامترهای هواشناسی متداول بوده است. می‌توان میزان متغیر وابسته را در روش‌های رگرسیونی در صورت برقراری رابطه بین یک متغیر وابسته با یک یا چند متغیر مستقل در کل پهنه تخمین زد. به‌طور مثال خلیلی (۳) نشان داد که گرادیان تغییرات دمای هوا یک بردار در فضای سه بعدی است که می‌توان تغییرات مؤلفه های قائم، نصف النهاری و مداری آن را با الگوهای رگرسیونی ساده خطی بیان نمود.

در الگوهای رگرسیونی نیز استفاده از روش‌های موضعی باعث شده تا جزئیاتی که در روش‌های سراسری نادیده گرفته می‌شوند تا حدی به تصویر کشیده شود. روش رگرسیون وزن‌دار جغرافیایی یکی از روش‌های رگرسیونی موضعی است. گوندوگدا و اسن (۱۱) سه روش کریجینگ، کوکریجینگ و رگرسیون وزن‌دار جغرافیایی را برای پهنه‌بندی میانگین ۲۵ ساله بارش سالانه در ترکیه بررسی کردند. آنها با محاسبه ضریب همبستگی بین مقدار پیش‌بینی شده و الگو شده نتیجه گرفتند که روش رگرسیون وزن‌دار جغرافیایی با ضریب تعیین ۸۶ درصد در مقایسه با روش کریجینگ و کوکریجینگ بترتیب با ضرایب تبیین ۵۱ و ۶۷ درصد کمترین مقدار خطای پیش‌بینی را به‌خود اختصاص می‌دهد. در این تحقیق سعی بر آن شده است تا دقت روش رگرسیون وزن‌دار جغرافیایی با روش‌های درون‌یابی معمول در ترسیم هم‌بارش در استان گیلان ارزیابی شود.

مواد و روش‌ها

منطقه مورد مطالعه

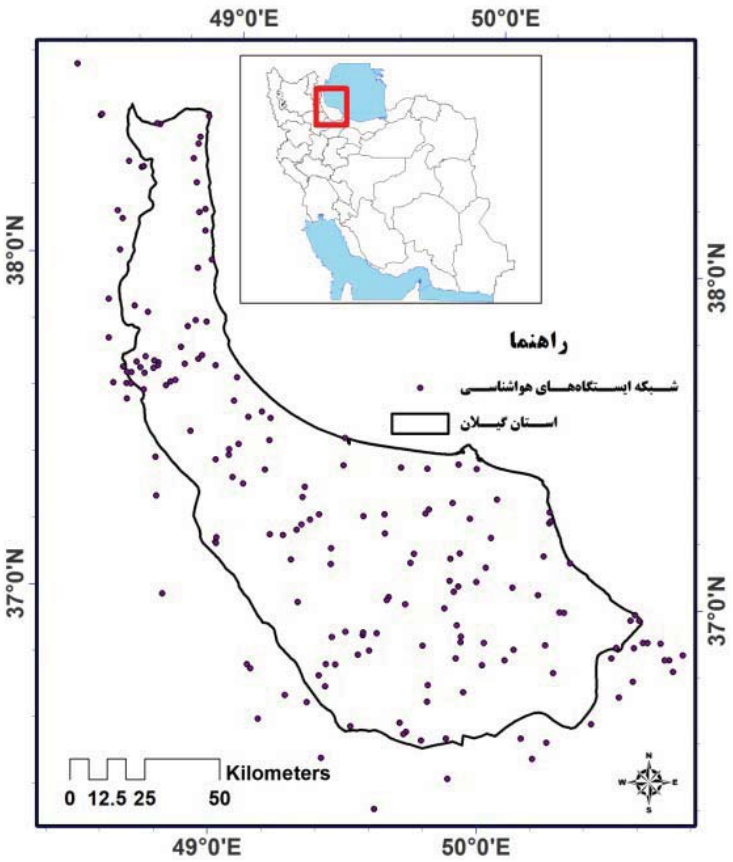
منطقه مورد مطالعه استان گیلان است. این استان از شمال به کرانه‌های جنوب غربی دریای خزر و از جنوب به رشته‌کوه‌های البرز و تالش محدود و مساحت آن ۱۴۷۱۱کیلومتر مربع است. میانگین بیست ساله آمار بارش سالانه ۱۸۵ ایستگاه هواشناسی (شامل ایستگاه‌های همدیدی، باران‌سنجی و تبخیرسنجی) بین سالهای ۱۳۶۹-۱۳۸۹ بعد از پالایش و کنترل کیفیت و بازسازی (استفاده از روش همبستگی با ایستگاه‌های مجاور) در این تحقیق استفاده شد. دامنه بارش ایستگاه‌ها بین ۱۷۴ تا ۱۸۸۷ میلیمتر در سال و ارتفاع ایستگاه‌ها از ۲۰ متر پائین‌تر از سطح دریای آزاد تا ۲۵۳۰ متر بالاتر از آن متغیر است. ایستگاه‌های مطالعاتی از اقلیم‌های متنوعی مطابق جدول ۱ برخوردارند. شکل ۱ موقعیت استان گیلان و ایستگاه‌های مورد استفاده را نشان می‌دهد.

روش‌های درون یابی و ترسیم نقشه‌های هم‌مقدار بارش

روش چندجمله‌ای سراسری و موضعی: درون‌یابی به روش چندجمله‌ای سراسری یکی از روش‌های درون‌یابی است که برای سطوحی که تغییرات پیوسته و هموار دارند مناسب است. یک سطح هموار در این روش با استفاده از یک تابع ریاضی (یک چندجمله‌ای با درجات مختلف) به تمام نقاط نمونه ورودی برازش داده می‌شود. درون‌یابی به روش چندجمله‌ای موضعی قابلیت بیشتری دارد. این روش تعدادی تابع چندجمله‌ای (با درجات مختلف) را به نقاط واقع در یک همسایگی تعیین شده برازش می‌دهد. این کار سبب می‌شود که سطوح با دقت بیشتری ارائه شود.

روش وزنی عکس فاصله: مقدار یک متغیر در نقطه‌ای که نمونه‌برداری انجام نشده باشد از روی نقاط مجاورش با استفاده از رابطه ۱ تخمین زده می‌شود:

1 - IDW :Inverse distance weighted



شکل ۱- موقعیت ایستگاه‌ها و منطقه مطالعاتی

جدول ۱- وضعیت اقلیمی ایستگاه‌های مطالعاتی در سیستم اقلیمی دمارتن

اقلیم	خشک	نیمه‌خشک	مدیترانه‌ای	نیمه‌مرطوب	مرطوب	بسیارمرطوب
تعداد ایستگاه‌ها	۷	۱۶	۸	۱۰	۱۴	۹۱

$$Z^* = \frac{\sum_{i=1}^{n} \frac{z_i}{d_i^a}}{\sum_{i=1}^{n} \frac{1}{d_i^a}} \quad (۱)$$

که در آن: Z^*: مقدار برآورد شده بارش، Z_i : مقدار مشاهده شده بارش در اطراف نقطه موردنظر، d_i: فاصله بین نقطه مشاهده شده تا نقطه مورد تخمین، a : عدد صحیحی (بین ۱ تا ۵) و n : تعداد نقاط مشارکت داده شده در درون‌یابی است (۱).

روش اسپیلاین: اسپلاین‌ها، توابع پیچیده‌ای شامل قطعاتی از چند جمله‌هایی با درجات مختلف بین هر دو نقطه از فضا هستند که در محل اتصالات، نرم شده‌اند. اسپیلاین‌ها یکی از بهترین روش‌ها برای نشان دادن تغییرات تدریجی سطوح مانند ارتفاع، عمق سطح ایستابی یا غلظت آلودگی می‌باشد. نتیجه درون یابی با این روش سطحی نرم و هموار می‌باشد (۱).

در روش اسپیلاین نقاط مجهول بر اساس معادله زیر تخمین زده می‌شوند:

$$\hat{Z}(s_0) = \sum_{i=1}^{n} \omega_i \emptyset(\|s_i - s_0\|) + \omega_{n+1} \quad (۲)$$

$\Phi(r)$: تابع پایه شعاعی، $r=\|s_i-s_0\|$ فاصله اقلیدسی بین پیش‌بینی مکان s_0 و هر داده مکانی s_i و $\{\omega_i=1,2,\ldots,n+1\}$وزن‌هایی که تخمین زده می‌شوند.

روش کریجینگ: کریجینگ یک روش تخمین آماری است که مقادیر مجهول را با استفاده از مقادیر معلوم و یک نیم‌تغییرنما، برآورد می‌کند. این روش بر منطق "میانگین متحرک وزن دار" استوار می‌باشد. کریجینگ علاوه بر تخمین مقادیر مجهول، خطای مرتبط با آن تخمین را نیز حساب می‌کند. بنابراین می‌توان فاصله اطمینان تخمین را برای هر مقدار برآورد شده محاسبه کرد (۱).

$$Z^*(x_i) = \sum_{i=1}^{n} \lambda_i . Z(x_i) \qquad (۳)$$

$Z^*(x_i)$: مقدار برآورد در نقطه x_i ، λ_i : وزن مربوط به متغیر z، $z(x_i)$: مقدار مشاهده شده متغیر اصلی است.

نرمال بودن متغیر Z شرط استفاده از این تخمینگر است. اگر Z توزیع نرمال نداشته باشد، باید از کریجینگ غیرخطی استفاده کرد. از آنجا که تخمینگر کریجینگ بهترین تخمینگر خطی نااریب است، لذا باید عاری از خطای سیستماتیک باشد و واریانس تخمین آن نیز حداقل باشد (۱).

روش کوکریجینگ: اگر هدف تخمین با چند متغیر باشد، می‌توان از روش کوکریجینگ استفاده کرد. همچنین در شرایطی مانند اینکه از یک متغیر به اندازه کافی داده در دسترس نباشد که بتوان بر اساس آنها تخمین را با دقت مورد نظر انجام داد، در چنین مواردی با در نظر گرفتن متغیرهای کمکی که داده‌های کافی از آن وجود دارد و بر اساس همبستگی متقابل بین متغیرهای اصلی وثانویه، به روش کوکریجینگ می‌توان فرآیند تخمین را انجام داد (۵).

$$Z^*(x_i) = \sum_{i=1}^{n} \lambda_i . Z(x_i) . \sum_{k=1}^{n} \lambda_k . y(x_k) \qquad (۴)$$

λ_k : وزن مربوط به متغیرکمکی، و $y(x_k)$: مقدارمشاهده شده متغیرکمکی می‌باشد.

روش رگرسیون وزن‌دار جغرافیایی[1]: روش رگرسیون سراسری معمولی، یک رابطه ثابت بین متغیرهای مکانی برای الگوسازی منطقه‌ای فرض می‌کند الگوهای رگرسیون معمولی مانند روش حداقل مربعات معمولی، ناایستایی مکانی متغیرها را به‌حساب نمی‌آورند. مزیت عمده GWR در مقابل الگوهای رگرسیون معمولی، توانایی آن در بررسی کردن ناایستایی مکانی است (۱۳). ناایستایی مکانی نشان می‌دهد که اندازه‌گیری یا تخمین روابط بین متغیرها از محلی به محل دیگر تفاوت می‌کند (۱۲). روش GWR یک فن رگرسیون موضعی است که به‌طور معنی‌داری رگرسیون معمولی را برای استفاده در داده‌های مکانی بهبود داده است. GWR بر مشکل ناایستائی در الگوسازی رگرسیونی با جداسازی موضعی آمارهای سراسری و محاسبه روابط بین متغیرهای موضعی برای هر نقطه به‌صورت جداگانه غلبه می‌کند. پارامترهای موضعی تخمین زده شده می‌توانند در محل‌های نقاط رگرسیونی ترسیم شوند. برخلاف الگوهای رگرسیون معمول که یک معادله رگرسیونی را برای توصیف روابط کلی بین متغیرها برقرار می‌کنند، GWR اطلاعات مکانی‌ای تولید می‌کند که تغییرات مکانی بین روابط متغیرها را بیان می‌کند.

بنابراین نقشه‌های تولید شده از این تحلیل‌ها نقش کلیدی در توصیف و تفسیر غیرایستایی مکانی بین متغیرها بازی می‌کند (۱۲).

GWR ابزاری مفید و عملی برای ارزیابی غیریکنواختی یک متغیر وابسته است. غیریکنواختی مکانی هنگامی که ساختار فرآیند الگوسازی شده تغییر می‌کند بوجود می‌آید. GWR معادلات جداگانه‌ای را با مشارکت متغیرهای مستقل و وابسته‌ای که در داخل یک نوار (فاصله) از هر پدیده قرار می‌گیرند تشکیل می‌دهد. شکل و اندازه پهنای باند برای نوع کرنل، روش پهنای باند، فاصله و تعداد پدیده‌ها توسط کاربر به عنوان ورودی در نظر گرفته می‌شود. GWR یک معادله رگرسیونی جداگانه را برای هر مشاهده به‌جای واسنجی یک معادله، تولید می‌کند، بنابراین این امکان را می‌دهد تا مقادیر پارامتر به‌صورت پیوسته در فضای جغرافیایی تغییر کند. هر یک از معادلات با استفاده از یک وزن متفاوت از مشاهدات مشتمل در مجموع داده‌ها واسنجی می‌شود. چون معادله رگرسیونی به‌طور غیر مستقل برای هر مشاهده واسنجی شده است، لذا یک پارامتر پیش‌بین جداگانه، مقدار Z و R^2 برای هر مشاهده محاسبه می‌کنند. جزئیات کامل GWR توسط فادرینگهام و همکاران (۱۰) ارائه شده است:

تحلیل‌های رگرسیونی برای الگوبندی رابطه بین یک متغیر با یک یا چند متغیر دیگر مطابق روابط ۵ و ۶ استفاده می‌شود.

$$y_i = \beta_{0i} + \beta_{1i}x_{1i} + \beta_{2i}x_{2i} + \cdots + \beta_{kn}x_{kn} + \varepsilon_i \qquad (۵)$$

$$y_i = \beta_0 + \sum_k \beta_k x_{ik} + \varepsilon_i \qquad (۶)$$

y_i بارش درون‌یابی شده در موقعیت i، مقدار β_0 عرض از مبدأ، β_{ik} برابر است با k^{th} پارامتر موضعی در i^{th} موقعیت، X_{iK} نشان دهنده k^{th} متغیر مستقل در i^{th} موقعیت و n بیانگر موقعیت قبلی است. وزن اختصاص داده شده به هر یک از مشاهدات در GWR بر اساس یک تابع تنزل فاصله در مرکز مشاهده i است.

الگوی رگرسیون وزنی جغرافیایی GWR موقعیت مکانی نمونه‌ها را در نظر می‌گیرد و این امکان را می‌دهد تا پارامترهای تخمین‌زده شده به‌صورت موضعی تغییر کند. یک الگو GWR می‌تواند به‌صورت رابطه ۷ نوشته شود (۱۰):

$$y_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i)x_{ik} + \varepsilon_i \qquad (۷)$$

بردار (u_i , v_i) مختصات i امین موقعیت را نشان می‌دهد، $\beta_0 (u_i , v_i)$ و $\beta_k (u_i , v_i)$ پارامترهای تخمین زده شده برای i امین موقعیت هستند که مقادیر آنها با موقعیت تغییر می‌کند. x_{ik} و ε_i بترتیب متغیرهای مستقل و میزان خطا در موقعیت i می‌باشند.

پارامترهای الگوی رگرسیون چند متغیره خطی بر اساس حداقل مربعات معمولی به‌صورت ماتریس ۸ تخمین زده می‌شود (۱۰):

$$\hat{\beta} = (X^T X)^{-1} X^T Y \qquad (۸)$$

X ماتریس مشاهدات و Y بردار پاسخ است. پارامترها در GWR با استفاده از یک تابع وزنی به‌صورت رابطه ۹ تخمین زده می‌شود:

1 -GWR: Geographically weighted regression

$$\hat{\beta}(u_i, v_i) = (X^T W((u_i, v_i))X)^{-1} X^T W((u_i, v_i))Y \quad (۹)$$

$W(u_i , v_i)$ وزن‌هایی انتخاب شده‌اند به‌طوری‌که آنهایی که به نقطه موردمطالعه نزدیک‌تر هستند نسبت به نقاط دورتر تأثیر بیشتری بر نتایج داشته باشند.

معیارهای ارزیابی روش‌های درون‌یابی: در ارزیابی دقت و اعتبار روش‌های درون‌یابی از روش اعتبارسنجی متقابل (حذفی) استفاده شد. این روش بر این اساس است که هر بار یک نقطه مشاهده‌ای حذف شده و برای آن از روی نقاط مجاور مقداری برآورد شود، سپس مقدار واقعی به محل قبلی برگردانده و برای کلیه نقاط شبکه این عمل صورت می‌گیرد. در نهایت در هر روش با توجه به مقادیر مشاهده شده و برآورد شده میانگین خطای مطلق[1]، میانگین خطای اریبی[2] و ریشه مجذور مربعات خطا[3] از روابط ۱۰ تا ۱۲ محاسبه می‌شود:

$$MBE = \sum_{i=1}^{n} (Z^*(x_i) - Z(x_i))/n \quad (۱۰)$$

$$MAE = \sum_{i=1}^{n} |Z^*(x_i) - Z(x_i)|/n \quad (۱۱)$$

$$RMSE = \left(\sum_{i=1}^{n} (Z^*(x_i) - Z(x_i))^2 / n \right)^{1/2} \quad (۱۲)$$

نتایج و بحث

نتایج روش‌های مختلف درون‌یابی مورد مطالعه در این تحقیق با استفاده از فن اعتبارسنجی متقابل ارزیابی و مورد مقایسه قرار گرفت. مقدار خطای درون‌یابی به‌صورت اختلاف بین مقادیر مشاهده شده و مقادیر الگوسازی شده محاسبه شد تا برای محاسبه معیارهای ارزیابی خطا شامل میانگین خطای اریب، ریشه مجذور مربعات خطا و میانگین خطای مطلق استفاده شود. همچنین با برقراری رابطه رگرسیونی بین مقادیر مشاهده شده و مقادیر الگوسازی شده، ضرایب همبستگی بین این دودسته متغیر نیز محاسبه شدند. نتایج این بررسی در جدول ۱ خلاصه شده است. موارد زیر در این مقاله بررسی شده‌اند:

درون‌یابی چندجمله‌ای سراسری: مقدار بارش را در کل منطقه مطالعاتی در این روش با برقراری یک رابطه چند جمله‌ای بین مختصات طول و عرض جغرافیایی ایستگاه‌ها با بارش ایستگاه‌ها، درون‌یابی می‌کند. نتایج این بررسی نشان می‌دهد که مقدار RMSE که خطای پیش‌بینی را نشان می‌دهد با افزایش درجه چندجمله‌ای از یک به سه کاهش یافته و ضریب همبستگی رابطه چند جمله‌ای افزایش پیدا می‌کند (جدول ۲ و شکل ۲). این روش درون‌یابی در بین کل روش‌های درون‌یابی از دقت پائین‌تری برخوردار است. اما نتایج آن روند تغییراتی را از بارش نشان می‌دهد که با افزایش فاصله از دریا میزان بارش کاهش می‌یابد (شکل ۲).

درون‌یابی چندجمله‌ای موضعی: این روش برخلاف روش درون‌یابی چندجمله‌ای سراسری از چند رابطه رگرسیونی بین مختصات طول و عرض جغرافیایی ایستگاه‌ها با بارش ایستگاه‌ها استفاده می‌کند تا مقدار بارش را در کل منطقه مطالعاتی درون‌یابی کند. این عمل سبب افزایش دقت درون‌یابی شده است که نشان دهنده وجود روند موضعی تغییرات بارش در منطقه می‌باشد (جدول ۲ و شکل ۲).

وزن عکس فاصله‌ها و اسپیلاین: بین این دو روش درون‌یابی، اختلاف زیادی مشاهده نمی‌شود و نتایج آن با درون‌یابی به روش چندجمله‌ای موضعی با درجه یک برابری می‌کند. مقدار میانگین خطای اریب در این دو روش مثبت است ولی در روش چندجمله‌ای موضعی با درجه یک مقدار آن منفی است. مقدار مثبت میانگین خطای اریب نشان دهنده بیش‌برآورد الگو و مقدار منفی آن نشان دهنده کم‌برآورد الگو می‌باشد (جدول ۱ و شکل ۲).

کریجینگ و کوکریجینگ: استفاده از روشهای زمین‌آماری مستلزم بررسی وجود ساختار مکانی در بین داده‌هاست که با واریوگرافی می‌توان به آن پی برد. اما یکی دیگر از شروط استفاده از این تجزیه و تحلیل نرمال بودن داده هاست. میانگین بارش سالانه که در این تحقیق مورد استفاده قرار گرفت با توجه به آزمون کولوموگراف -اسمیرنوف نرمال تشخیص داده شد و نتایج حاصل از تحلیل واریوگرافی نیز ساختار مکانی بین داده‌ها را نشان می‌دهد (جدول ۳).

نتایج درون‌یابی به روش کریجینگ و کوکریجینگ نشان می‌دهد که دقت روش کریجینگ در مقایسه با روش‌های قبلی ذکر شده افزایش یافته است به‌طوریکه مقدار RMSE از ۱۹۲ در روش اسپلاین به مقدار ۱۷۸ در روش کریجینگ کاهش یافته است. با دخالت دادن پارامتر کمکی ارتفاع و استفاده از روش کوکریجینگ، نه تنها دقت روش درون‌یابی افزایش پیدا نکرد، بلکه دقت ان کاهش پیدا کرد و مقدار RMSE در روش کوکریجینگ به ۱۸۸ رسید (جدول ۲).

رگرسیون وزن‌دار جغرافیایی: روش رگرسیون وزن‌دار جغرافیایی یک روش آماری است که برای مطالعه الگوهای موضعی سازگار شده است. روش GWR وزن‌های نسبی بیشتری به مشاهدات نزدیک‌تر و وزن کمتر یا صفر به آنهایی که در دوردست هستند اختصاص می‌دهد. به عبارت دیگر، GWR فقط از مشاهداتی که از لحاظ جغرافیایی نزدیک هستند برای تخمین ضرایب موضعی استفاده می‌کند. این شیوه وزن دهی براساس این تفکر است که استفاده از مشاهدات نزدیک از لحاظ جغرافیایی بهترین روش برای تخمین ضرایب موضعی است.

1 -MAE: Mean absolute error
2 - Mean biases error
3 - Root mean square error

جدول ۲- نتایج ارزیابی خطا روش‌های مختلف درون‌یابی

الگو	علامت اختصاری الگو	متغیرهای دخالت داده شده در درون‌یابی	معیارهای ارزیابی خطا RMSE	MAE	MBE	ضریب تعیین R^2
چندجمله‌ای سراسری با درجات ۱، ۲ و ۳	GP1	بارش	۳۵۸/۸	۲۹۵/۲	۱/۶	۰/۲۳
	GP2	بارش	۲۶۱/۴	۲۱۸/۱	-۰/۱	۰/۵۹
	GP3	بارش	۲۵۱/۶	۲۰۴/۳	-۱/۴	۰/۶۲
چندجمله‌ای موضعی با درجات ۱، ۲ و ۳	LPI1	بارش	۱۹۳/۹	۱۴۷/۸	-۶/۷	۰/۷۷
	LPI2	بارش	۲۲۰/۶	۱۳۷/۷	۹/۹	۰/۷۱
	LPI3	بارش	۲۰۷/۹	۱۶۰/۹	-۶/۹	۰/۷۴
اسپیلاین	RBF	بارش	۱۹۲/۱	۱۴۴/۸	۷/۴	۰/۷۸
وزن عکس فاصله	IDW	بارش	۱۹۴/۴	۱۳۸/۴	۶	۰/۷۷
کریجینگ معمولی	OK	بارش	۱۷۸/۱	۱۳۲/۹	۱/۴	۰/۸۱
کوکریجینگ	CO_Kriging	بارش+ارتفاع	۱۸۸/۱	۱۴۱/۸	۳/۵	۰/۷۹
رگرسیون وزن دار جغرافیایی	GWR_Z	بارش+ارتفاع	۱۵۰/۲	۱۱۵/۳	-۶/۵	۰/۸۶
	GWR_ZD	بارش+ارتفاع+فاصله از دریا	۱۴۷/۱	۱۱۳/۱	-۳/۶	۰/۸۷
	GWR_ZA	بارش+ارتفاع+ سمت دریا	۲۱۳/۵	۱۷۲/۹	-۱۵/۹	۰/۷۲
	GWR_ZDA	بارش+ارتفاع+ فاصله و سمت دریا	۱۷۷/۷	۱۳۷/۶	۲/۹	۰/۸۱

جدول ۳- نتایج برازش الگوهای مختلف واریوگرام

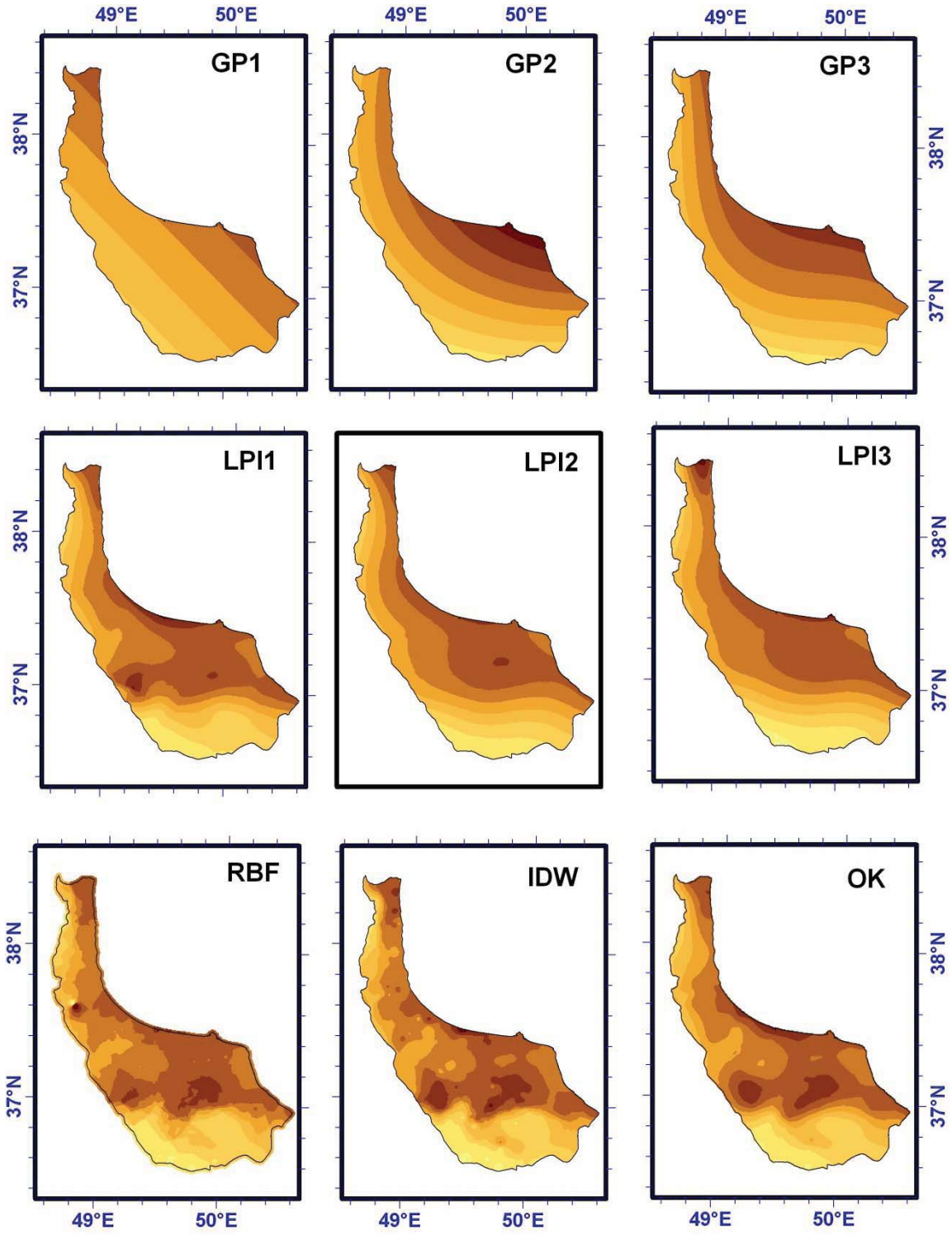
الگو واریوگرام	واریانس بدون ساختار(اثر قطعه‌ای) C0	آستانه (واریانس ساختار دار +واریانس بدون ساختار) C0+C	معیار استحکام ساختار مکانی C/(C0+C)	ضریب تعیین R^2
Spherical	۰/۷۹	۳/۵۶	۰/۷۷	۰/۵۷
Exponential	۰/۷۳	۳/۵	۰/۷۹	۰/۵۸
Linear	۰/۸	۳/۶	۰/۷۸	۰/۵۶
Linear to Sill	۰/۸	۳/۶۲	۰/۷۸	۰/۵۶
Gaussian	۰/۹۹	۳/۷۵	۰/۷۴	۰/۴۴

روش GWR نه تنها اثرات موقعیت خود متغیرها را بر روی متغیر مستقل بلکه اثرات موقعیت‌های همسایگی را نیز در نظر می گیرد. این موضوع باعث شده است تا دقت این روش نسبت به دیگر روش‌های درون‌یابی برای پهنه‌بندی مقدار بارش در منطقه مطالعاتی به‌طور معنی‌داری افزایش یابد به‌طوری‌که باعث کاهش RMSE تا میزان ۱۴۷ و افزایش ضریب تعیین تا حدود ۸۷ درصد شود.

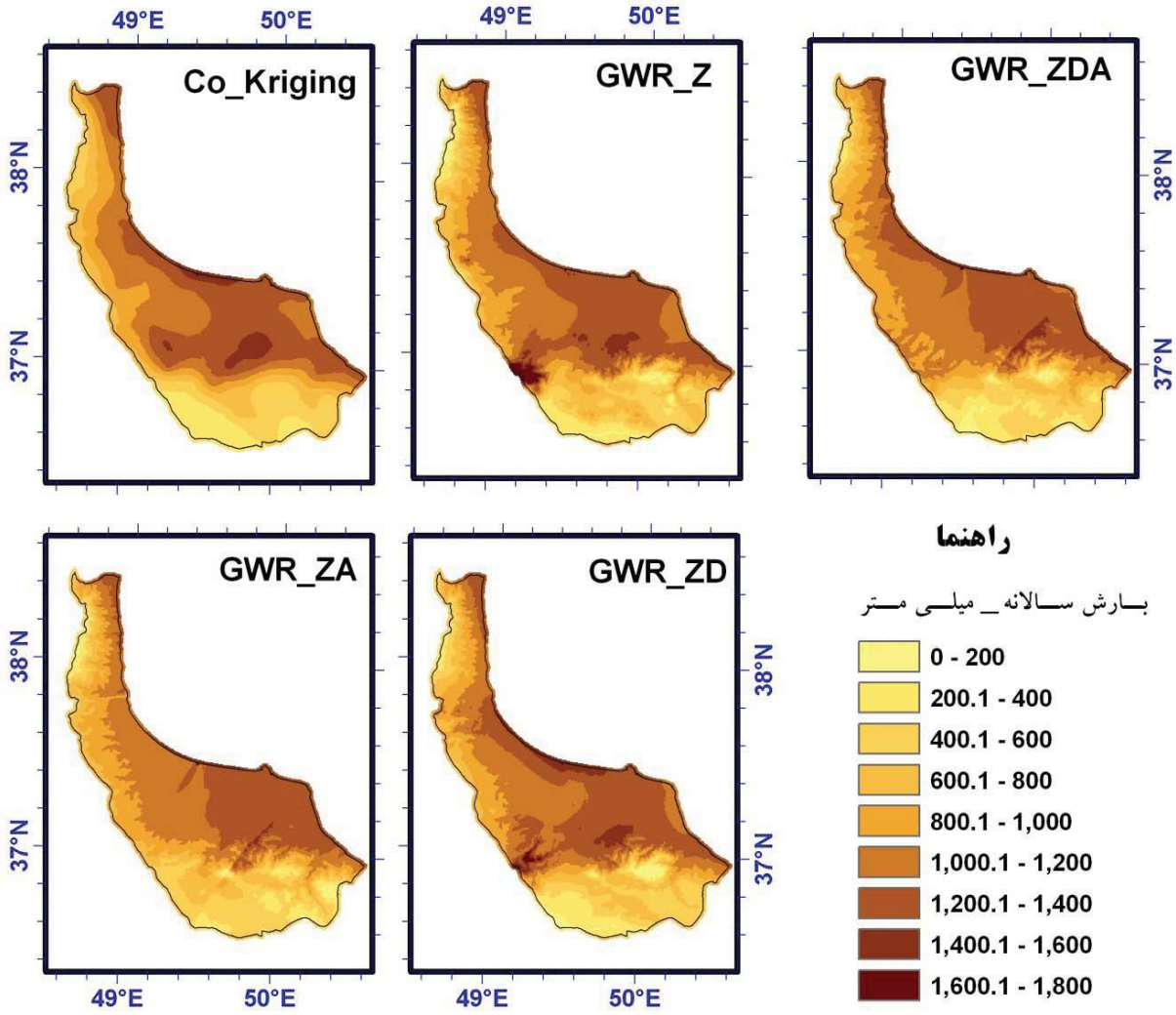
دخالت دادن پارامترهای مؤثر در معادلات رگرسیون‌در روش GWR سبب تغییر دقت الگو شده است، پارامترهای ارتفاع (Z)، فاصله از دریا (D)، زاویه آزیموت سمت دریا (A) در این تحقیق وارد معادلات رگرسیونی شدند تا نتایج آنها در تخمین مقدار بارش در منطقه مطالعاتی مورد ارزیابی قرار گیرد. نتایج این بررسی نشان داد که دو پارامتر ارتفاع و فاصله از دریا از عوامل مؤثر در تغییرات مکانی بارش در منطقه می‌باشند به‌طوری‌که الگوی رگرسیونی وزن‌دار جغرافیایی حاصل از آنها (GWR_ZD) کمترین مقدار خطا با RMSE=147 و بیشترین همبستگی با ضریب تعیین ۸۷ درصد را دارد. دخالت دادن پارامتر زاویه آزیموت سمت دریا نه تنها باعث کاهش خطای پیش‌بینی بلکه باعث افزایش خطای پیش بینی می‌شود (الگوهای GWR_ZA و GWR_ZDA) ولی خود پارامتر ارتفاع به تنهایی ۸۶ درصد واریانس بارش را در منطقه توجیه می‌کند. الگو GWR_Z نیز که تنها پارامتر بارش را در الگوی رگرسیونی وزن‌دار جغرافیایی شرکت داده است بعد از الگو GWR_ZD کمترین مقدار خطا با RMSE=150 را دارا می‌باشد (جدول ۲ و شکل ۲).

مطالعات زیادی در زمینه ارزیابی روش‌های مختلف درون‌یابی قطعی و زمین آمار برای برآورد بارش انجام شده است. با وجود شباهتهایی که در نتایج مطالعات مختلف بدست آمده، اختلافات زیادی نیز در آنها به چشم می‌خورد. در دو منطقه متفاوت، مهدیزاده (۷) در حوضه آبریز دریاچه ارومیه و اقدسی (۱) در منطقه دشت برخوار اصفهان در ارزیابی روش‌های مختلف درون‌یابی بارش به این نتیجه رسیدند که روش اسپلاین نسبت به روش‌های کریجینگ، کوکریجینگ و وزن عکس فاصله‌ها برآورد بهتری را دارد، ولی نادی و

همکاران (۸) بـرای درون‌یـابی داده‌هـای بارنـدگی ماهانـه در منطقـه چهارباغ گرگان به این نتیجه رسیدند که در همه ماه‌هـای سـال روش درون‌یابی کریجینگ نسبت بـه روش‌هـای درون‌یـابی وزنـی عکـس فاصله و اسپیلاین برتری ندارد.



شکل ۲- نقشه‌های هم‌بارش ترسیم شده با استفاده از روش‌های مختلف درون‌یابی



ادامه شکل ۲- نقشه‌های هم بارش ترسیم شده با استفاده از روش‌های مختلف درون‌یابی

لذا انتخاب یک روش درون‌یابی مناسب بستگی به متغیر مورد نظر دارد و عوامل دیگری همچون منطقه مورد مطالعه، تعداد نقاط مشاهده، تراکم و پراکندگی داده‌ها، همگن یا غیرهمگن بودن نقاط و سایر عوامل فیزیکی و شیمیایی در انتخاب روش بهینه برآورد، مؤثرند (۱). در مورد روش رگرسیون وزن‌دار جغرافیایی، دخالت دادن عوامل مؤثر بر تغییرات مکانی بارش باعث می‌شود تا دقت درون‌یابی افزایش یابد. بطوریکه در این تحقیق استفاده از متغیرهای ارتفاع و دوری از دریا به عنوان دو عامل مؤثر در تغییرات مکانی بارش باعث شده است تا روش رگرسیون وزن‌دار جغرافیایی برآورد بهتری نسبت به روش‌های قطعی و زمین آمار در درون‌یابی بارش سالانه داشته باشد. از بین مطالعاتی که صورت گرفته است و روش رگرسیون وزن‌دار جغرافیایی را با دیگر روش‌های درون‌یابی مورد مقایسه قرار داده‌اند می‌توان به مطالعات گوندوگدا و اسن (۱۱) در ترکیه اشاره کرد که برتری روش رگرسیون وزن‌دار جغرافیایی را نسبت به روش‌های کریجینگ، کوکریجینگ در میان‌یابی بارش نشان می‌دهد و با نتایج این تحقیق نیز کاملا همخوانی دارد.

نتیجه‌گیری

روش رگرسیون وزن‌دار جغرافیایی یک روش آماری است که برای مطالعه الگوهای موضعی سازگار شده است بنابراین این روش در مناطقی که از الگوهای تغییرات مکانی مختلفی از بارش پیروی می‌کنند روش مناسبی برای پهنه‌بندی و تخمین داده‌های مکانی بحساب می‌آید. نتایجی که از این تحقیق به‌دست آمد نیز نشان داد که روش رگرسیون وزن‌دار جغرافیایی برتری معنی‌داری نسبت به روش‌های قطعی و زمین‌آماری در تحلیل مکانی تغییرات بارش در منطقه مطالعاتی دارد به‌طوریکه با دخالت دادن دو پارامتر ارتفاع و فاصله از دریا به کمک این روش می‌توان با مقدار خطایی برابر RMSE=147 بارش را در منطقه مطالعاتی پهنه‌بندی کرد در حالیکه

روش کریجینگ که بیشترین دقت را در بین روش‌های چندجمله‌ای سراسری و موضعی، وزن عکس فاصله‌ها، اسپیلاین و کوکریجینگ داشت با RMSE=178 اختلاف قابل ملاحظه‌ای را با روش رگرسیون وزن‌دار جغرافیایی دارد. بر این اساس می‌توان نتیجه گرفت که روش رگرسیون وزن دار جغرافیایی می‌تواند ابزاری مناسب برای توجیه تغییرات موضعی بارش و مناسب‌ترین روش در ترسیم نقشه های همبارش سالانه در استان گیلان به حساب آورده شود.

منابع

۱- اقدسی ف. ۱۳۸۳. ارزیابی چند روش زمین آماری ترسیم میدان عددی بارندگی روزانه و سالانه (مطالعه موردی دشت برخوار). پایان نامه کارشناسی ارشد. دانشگاه تهران.

۲- جامعی م. ۱۳۸۷. ارزیابی روشهای درون یابی در برآورد منطقه ای تبخیر و تعرق مرجع و مقایسه با نتایج موجود حاصل از تصاویر ماهواره ای در دشت های مرکزی و شمالی خوزستان. پایان نامه کارشناسی ارشد. دانشگاه آزاد اسلامی واحد علوم و تحقیقات.

۳- خلیلی ع. ۱۳۷۵. تغییرات سه بعدی میانگین های سالانه درازمدت دمای هوا در گستره ایران. مجله نیوار شماره ۳۲.

۴- رحیمی س. و مهدیان م. ۱۳۸۴. بررسی روش های توزیع مکانی بارندگی روزانه و ماهانه در حوضه دریای خزر. نشریه پژوهش و سازندگی در منابع طبیعی ۶۹: ۶۳-۷۲.

۵- مدنی ح. ۱۳۷۳. مبانی زمین آمار. انتشارات دانشگاه صنعتی امیرکبیر.

۶- مهدیان م. و غیاثی ح. ۱۳۸۲. بررسی روش های مختلف میان یابی داده های بارندگی ماهانه در ناحیه مرکزی ایران. نشریه علوم و فنون کشاورزی و منابع طبیعی: ۳۳-۴۴.

۷- مهدیزاده م. ۱۳۸۱. ارزیابی روش های زمین آماری برای برآورد دما و بارندگی در حوضه آبریز دریاچه ارومیه. پایان نامه کارشناسی ارشد. دانشگاه تهران.

۸- نادی م. ۱۳۸۹. استفاده از تکنیک های های مختلف پهنه بندی داده های اقلیمی برای تعیین مهمترین فاکتورهای مؤثر بر رویش درختان ناحیه مرتفع چهار باغ گرگان. پایان نامه کارشناسی ارشد. دانشگاه تهران.

9- Attorre F., and Alfo M. 2007. Comparison of interpolation methods for mapping climatic and bioclimatic variables at regional scale. International journal of climatology 27: 1825-1843.

10- Fotheringham A.S., Brunsdon C., and Charlton M.E. 2002. Geographically weighted regression. Chichester: John Wiley & Sons.

11- Gundogdu I., and ESEN O. 2010. The importance of secondary variables for mapping of meteorological data. 3rd international conference on cartography and GIS. 15-20 June, 2010, Nessebar, Bulgaria.

12- Mennis J. 2006. Mapping the Results of Geographically Weighted Regression. The Cartographic Journal, 43(2): 171-179.

13- Propastin P., and Kappas M. 2008. Reducing uncertainty in modeling the NDVI–precipitation relationship: a comparative study using global and local regression techniques. GISci Remote Sens 45:47–67

14- Tobler W.R. 1970. A computer movie simulating urban growth in the Detroit region, Economic Geography 46: p.234-240.

Journal of Water and Soil
Vol. 26, No. 3, Jul-Aug 2012, p. 743-752

نشریه آب و خاک (علوم و صنایع کشاورزی)
جلد ۲۶، شماره ۳، مرداد – شهریور ۱۳۹۱، ص. ۷۵۲-۷۴۳

Geographically Weighted Regression: A Method for Mapping Isohyets in Gilan Province

Kh. Ghorbani[1]
Received: 13-7-2011
Accepted: 30-1-2012

Abstract

So far several methods have been developed for mapping and interpolation of isohyets.one of the recently accepted methods is geographically weighting regression which is suitable for evaluation of spatial heterogeneity of dependent variable by using local regressions. In order to evaluate annually precipitation spatial variation, this study was conducted in Gilan province which precipitation is distributed non-uniform due to different environmental conditions. The results of geographically weighting regression method were compared with another interpolation methods including global polynomial, local polynomial, inverse distance weighting (IDW), spiline, kriging and co-kriging and . In this study, average of 20 years annually precipitation data of 185 meteorological observations over Gilan Province and its neighboring stations was used for modeling of spatial distribution variations of mean annual precipitation by using other variables like elevation and position of points to the sea level. Cross validation technique was used to assessment accuracy of each interpolation methods. The result showed that geographically weighting regression method had minimum error with RMSE=147 and had significant difference with the kriging method which was in the second rank with RMSE=187. Finally the best method for mapping isohyets in Gilan province is geographically weighting regression method.

Keywords: Geographically weighted regression, Gilan, Interpolation, Isohyets, Kriging

1- Assistant Professor of Water Engineering Deepartment, Gorgan University of Agriculture and Natural Resources, Gorgan
Email: ghorbani.khalil@yahoo.com