



روشی تازه به منظور درون‌یابی مشاهدات بارندگی با کمک کریجینگ شاخص نرم و الگوریتم ژنتیک

حمیدرضا زورقین*^۱، علی اصغر آل شیخ^۲، عباس علیمحمدی^۳، محمدحسن وحیدنیا^۴

۱. دانشجوی کارشناسی ارشد، سیستم‌های اطلاعات مکانی، دانشگاه صنعتی خواجه نصیرالدین طوسی

۲. دانشیار گروه مهندسی سیستم‌های اطلاعات مکانی، دانشگاه صنعتی خواجه نصیرالدین طوسی

۳. استادیار گروه مهندسی سیستم‌های اطلاعات مکانی، دانشگاه صنعتی خواجه نصیرالدین طوسی

۴. دانشجوی دکتری سیستم‌های اطلاعات مکانی، دانشگاه صنعتی خواجه نصیرالدین طوسی

تاریخ پذیرش مقاله: ۱۳۹۰/۶/۱۱

تاریخ دریافت مقاله: ۱۳۹۰/۲/۱۹

چکیده

درون‌یابی از جمله مهم‌ترین مسائل در علوم مکانی است؛ زیرا همیشه این نیاز وجود دارد که از یک حجم محدود داده، مجموعه‌ای از اطلاعات دقیق فراهم گردد. کریجینگ نیز از کاربردی‌ترین و دقیق‌ترین روش‌های درون‌یابی به شمار می‌آید که به زیر روش‌های مختلفی از جمله کریجینگ معمولی، کریجینگ شاخص و کریجینگ شاخص نرم تقسیم می‌شود. الگوریتم ژنتیک در زمره الگوریتم‌های تکاملی است که برای یافتن یک جواب بهینه از مجموعه‌ای از جواب‌های بالقوه استفاده می‌شود. در این مقاله، از الگوریتم ژنتیک برای آموزش بهینه برخی از پارامترهای کریجینگ شاخص نرم با توجه به تابع هدف برای رسیدن به جواب بهینه استفاده می‌شود. در این تحقیق، از چهار روش کریجینگ معمولی، کریجینگ شاخص، کریجینگ شاخص نرم و کریجینگ شاخص نرم با الگوریتم ژنتیک بر یک مجموعه داده از ۱۵ ایستگاه سینوپتیک استان مازندران استفاده گردید. داده‌های استفاده‌شده، مشاهدات بارندگی در این ایستگاه‌ها مربوط به بهمن ماه سال ۱۳۸۷ بود. با مقایسه مقادیر مشاهداتی و محاسباتی در دو حالت اعتبارسنجی متقابل و جکنایف در چهار روش ذکرشده، ثابت گردید که روش کریجینگ شاخص نرم با الگوریتم ژنتیک دقیق‌ترین روش است. به عنوان معیار این تصمیم‌گیری، از میانگین قدرمطلق اختلاف بین مقادیر مشاهداتی و محاسباتی هر روش استفاده گردید. این کمیت برای روش کریجینگ شاخص نرم با الگوریتم ژنتیک برابر با ۱۰/۳ میلی‌متر است که نسبت به سه روش اول (۱۲/۳، ۱۳/۳ و ۱۲/۸) کمینه است.

کلیدواژه‌ها: درون‌یابی، کریجینگ، عدم قطعیت، کریجینگ شاخص نرم، الگوریتم ژنتیک.

* نویسنده مکاتبه‌کننده: تهران، خیابان ولیعصر، تقاطع میرداماد، دانشگاه صنعتی خواجه نصیرالدین طوسی، تلفن: ۸۷۸۶۲۱۲

۱- مقدمه

(SIK) ^۱ برخلاف روش کریجینگ شاخص داده‌ها را با توجه به توزیع احتمالاتی^۱ شان به اعدادی بین ۰ و ۱ تبدیل می‌کند (سایتو و گووارتس ۲۰۰۲ و ۲۰۰۳ و رنارد و همکاران ۲۰۰۵). در حقیقت، این روش خطاهای اندازه‌گیری را هم به نوعی در فرایند درون‌یابی وارد می‌کند. این تبدیل بر اساس تابع توزیع احتمالاتی نرمال داده‌ها صورت می‌گیرد. حال اگر با استفاده از الگوریتم ژنتیک (GA) ^۲ حدود آستانه مرتبط با این تبدیل طوری آموزش ببینند که منجر به حداقل‌سازی تابع هدف گردند، می‌توان دقت درون‌یابی را افزایش داد.

به عبارت دیگر، در این تحقیق برای بهبود دقت روش کریجینگ شاخص، با ارائه روشی ترکیبی دو مورد از منابع خطای روش کریجینگ شاخص مدل‌سازی شده‌اند. اولین منبع خطا مربوط به خطای اندازه‌گیری مشاهدات است، زیرا در روش کریجینگ شاخص با توجه به تبدیل قطعی مشاهدات به دو حالت ۰ یا ۱، فرض می‌شود که هیچ‌گونه عدم قطعیتی در خصوص نحوه اندازه‌گیری مشاهدات وجود ندارد. فرض مذکور به‌ویژه در شرایطی که دقت دستگاه‌های اندازه‌گیری پایین باشد، دور از واقعیت است. دومین منبع خطا مربوط به انتخاب حدود آستانه است. این حدود یا با استفاده از پارامترهای مربوط به توزیع مشاهدات (اعدادی که توزیع مشاهدات را به قسمت‌های مساوی تقسیم می‌کنند)^{۱۱} انتخاب می‌شوند و یا با توجه به دامنه تغییرات مشاهدات و نیازمندی‌های مسئله به طور معمولی انتخاب می‌گردند. از آنجا که دقت درون‌یابی

از مهم‌ترین موضوعات در علوم مکانی، مسئله درون‌یابی^۱ است؛ چرا که به دست آوردن داده‌ها هزینه‌بر و زمان‌بر است و همواره محدودیتی در خصوص میزان داده‌های قابل اکتساب در هر پروژه وجود دارد و در عمل باید به میزان محدودی بسنده کرد. در درون‌یابی، از این داده‌ها محدود برای تولید و شبیه‌سازی کمیت مورد نظر در سایر نقاط منطقه مورد مطالعه استفاده می‌گردد. درون‌یابی باید تا جایی که ممکن است دقیق باشد و انطباق بیشتری با واقعیت داشته باشد تا حاصل کار قابل اعتماد گردد.

تا کنون روش‌های مختلفی برای درون‌یابی ارائه شده‌اند که هر کدام مزایا و معایبی دارند و سطوح متفاوتی از دقت را نتیجه می‌دهند. روش کریجینگ^۲ یکی از روش‌های پرسابقه در زمینه درون‌یابی است که برای نخستین بار کریج^۳ در سال ۱۹۵۱ آن را ارائه کرد (کریج، ۱۹۵۱). روش وی برآوردی بدون تورش با واریانس کمینه را به دست می‌دهد (ژوانگ و همکاران، ۲۰۰۸). از مهم‌ترین مزایای روش کریجینگ می‌توان به این دو مورد اشاره کرد: استفاده از تابع واریوگرام^۴ برای در نظر گرفتن ساختار مکانی کمیت مورد نظر که منجر به بهبود درون‌یابی می‌شود؛ و برآورد دقت درون‌یابی برای هر نقطه مشخص. اساساً می‌توان گفت که هر دوی این مزایا در پی آماری بودن این روش حاصل می‌شوند (بارو و مک دانل، ۱۹۹۸).

امروزه، شاخه‌های مختلفی از روش کریجینگ گسترش یافته‌اند که هر کدام با توجه به ویژگی‌ها و اهداف خاص درون‌یابی، می‌توانند دقت این روش را بهبود دهند. یکی از این روش‌ها، روش کریجینگ شاخص (IK)^۵ است (ژورنل، ۱۹۸۳). در این روش با تعریف حدود آستانه^۶ و گسسته‌سازی داده‌ها با توجه به این حدود آستانه، درون‌یابی صورت می‌گیرد. این روش، داده‌ها را به صورت دو حالت^۷ یعنی {۰، ۱} تبدیل می‌کند و برای هر نقطه مجهول، احتمال عبور یا عدم عبور کمیت مجهول آن نقطه را از حدود آستانه^۸ تعریف‌شده نتیجه می‌دهد. روش کریجینگ شاخص نرم

1. Interpolation
2. Kriging
3. Krige
4. Variogram
5. Indicator Kriging
6. Threshold
7. Binary
8. Soft Indicator Kriging
9. Probability Distribution
10. Genetic Algorithm
11. Quantile

ارتفاعی استفاده کردند. لی و همکاران (۲۰۰۷) روش کریجینگ شاخص را برای برآورد میزان آرسنیک در آب‌های زیرزمینی در شمال شرقی کشور تایوان به منظور بررسی‌های مربوط به حوزه سلامت به کار بردند. لارک و فرگوسن (۲۰۰۴) از این روش برای برآورد احتمال کمتر یا برابر بودن میزان فسفر از سه حد آستانه در سطح خاک منطقه‌ای در نبراسکای آمریکا بهره بردند. آنها در تحقیق‌شان از ۲۴۷ نقطه‌ی مشاهداتی برای این برآورد در ۱۶۲۲ نقطه کنترل استفاده کردند. چو و همکاران (۲۰۱۰) در تحقیق‌شان از کریجینگ شاخص چندمتغیره برای تعیین احتمالات مربوط به میزان چهار فلز سنگین کرم، مس، نیکل و روی به منظور مطالعه آلودگی خاک در مناطق مرکزی کشور تایوان استفاده کردند. لیون و همکاران (۲۰۰۶) روش کریجینگ شاخص را به منظور تعیین احتمال اشباع آب برای مدل‌سازی مکانی نقاط وقوع اشباع در شمال شرقی آمریکا به کار بردند. سایتو و گووارتس (۲۰۰۲) و (۲۰۰۳) و رنارد و همکاران (۲۰۰۵) نیز از کریجینگ شاخص نرم برای منظور کردن خطای اندازه‌گیری مشاهدات در روند محاسبات روش کریجینگ شاخص استفاده کردند.

گو و همکاران (۲۰۰۷) از کریجینگ فازی^۵ برای درون‌یابی میزان PM_{10} در ایالت کالیفرنیا استفاده کردند. آن‌ها به منظور فازی‌سازی داده‌های اولیه از سه تابع عضویت نیمه‌آماری^۶ استفاده کردند و با تعریف و محاسبه عباراتی نظیر درجه عضویت فازی^۷، واریوگرام فازی^۹ و کریجینگ فازی، درون‌یابی خود را انجام دادند.

ترکیب‌های مختلف این حدود آستانه متفاوت است، بنابراین باید روشی برای انتخاب هدفمند و بهینه این حدود آستانه ایجاد گردد. بدین منظور در تحقیق حاضر، تلفیقی از کریجینگ شاخص نرم و الگوریتم ژنتیک برای افزایش دقت درون‌یابی پیشنهاد شده است. کریجینگ شاخص نرم برای مدل‌سازی خطای نوع اول و الگوریتم ژنتیک برای مدل‌سازی خطای نوع دوم به کار رفته‌اند. روش پیشنهادی، روش کریجینگ شاخص نرم بهینه‌شده با الگوریتم ژنتیک (GSIK)^۱ نام دارد. همچنین، چون این روش برگرفته از روش کریجینگ شاخص است، خروجی آن تابع توزیع تجمعی شرطی (CCDF)^۲ است و برخلاف سایر روش‌های درون‌یابی به یک عدد محدود نمی‌شود. ویژگی تابع مذکور آن است که برآوردی از عدم قطعیت محلی را برای هر نقطه ارائه می‌کند. بنابراین در صورت تلفیق دو روش کریجینگ شاخص نرم و الگوریتم ژنتیک از طریق آموزش پارامترهای روش کریجینگ شاخص نرم با استفاده از الگوریتم ژنتیک می‌توان از مزایای هر دو استفاده کرد و به روشی جامع در درون‌یابی دست یافت. برای اثبات این امر، از روش پیشنهادی بر روی مشاهدات بارندگی بهمن ماه سال ۱۳۸۷ در ۱۵ ایستگاه هواشناسی سینوپتیک استان مازندران استفاده شده و نتایج آن با نتایج سه روش کریجینگ معمولی، کریجینگ شاخص و کریجینگ شاخص نرم مورد مقایسه قرار گرفته است. در ادامه مقاله، ابتدا پیشینه تحقیق مورد بررسی قرار گرفته است. سپس منطقه مطالعاتی و روش‌های مورد استفاده با جزئیات بیشتری معرفی شده‌اند. نحوه پیاده‌سازی تحقیق و نتایج آن در ادامه آورده شده و بالاخره بحث و نتیجه‌گیری ارائه شده است.

۱-۱- پیشینه تحقیق

استفاده از روش‌های کریجینگ دارای سابقه طولانی در علوم زمین آماری است. به طور مثال لوید و آتکینسون (۲۰۰۱) از سه روش کریجینگ معمولی (OK)^۳، کریجینگ با حضور ترند (KT)^۴ و کریجینگ شاخص برای درون‌یابی و بررسی عدم قطعیت در داده‌های

1. Genetic assisted Sofi Indicator Kriging
2. Conditional Cumulative Distribution Function
3. Ordinary Kriging
4. Kriging with a Trend
5. Fuzzy Membership Grade Kriging
6. Particular Matter with a mass median aerodynamic diameter of less than 10 μm
7. Semi-Statistical Membership Function
8. Fuzzy Membership Grade
9. Fuzzy Membership Grade Variogram

در شبکه‌ای متشکل از ۶ ایستگاه ارائه کردند. داده‌های آنها میزان بارندگی روزانه در این ایستگاه‌ها در سال ۲۰۰۲ بود، و برای تعیین وزن هر ایستگاه در میانگین‌گیری وزن دار، از دو پارامتر فاصله و اختلاف ارتفاع استفاده کردند که با بهره‌گیری از دو تابع عضویت فازی، در محدوده بین ۰ و ۱ قرار داشتند. سپس با کمک الگوریتم ژنتیک، پارامترهای این توابع عضویت بهینه شدند. نتایج مطالعه چانگ و همکاران نشان داد که دقت روش آنها در بسیاری از شرایط از دو روش میانگین‌گیری حسابی و چندضلعی‌های تیسسن^۱ بالاتر است. در تحقیق فرجی سبکبار و عزیز (۱۳۸۶)، میزان دقت روش‌های درون‌یابی تین، میانگین‌گیری وزن دار بر اساس معکوس فاصله (IDW)^۲، اسپلاین^۳ و کریجینگ برای درون‌یابی داده‌های بارندگی یک حوزه آبریز با توجه به نتایج اعتبارسنجی متقابل با هم مقایسه شدند. بر اساس نتایج آنها، روش‌های تین و کریجینگ نتایج درون‌یابی بهتری داشتند.

۲- مواد و روش‌ها

در این تحقیق، از اطلاعات مربوط به ۱۵ ایستگاه سینوپتیک هواشناسی استان مازندران استفاده شد. این اطلاعات از وبگاه اداره هواشناسی استان مازندران جمع‌آوری گردید. تمامی برنامه‌های مربوط به روش‌های مختلف موجود در این تحقیق، با استفاده از نرم‌افزار متلب^۱ نوشته شدند. در ادامه، منطقه مطالعاتی و روش‌های انجام‌شده مورد بررسی قرار گرفته‌اند.

شاد و همکاران (۲۰۰۹) برای برآورد میزان PM_{۱۰} در ۱۰ ایستگاه پایش آلودگی هوا در شهر تهران، از سه روش کریجینگ معمولی، کریجینگ فز و کریجینگ فازی - ژنتیک^۱ استفاده کردند. داده‌های آن‌ها مقدار این کمیت در ۵۲ ایستگاه پایش آلودگی هوا بود که ۱۰ ایستگاه به عنوان نقاط کنترل در نظر گرفته شدند. با توجه به اینکه مقدار مشاهداتی PM_{۱۰} در این ۱۰ ایستگاه مشخص بود، آن‌ها با مقایسه مقادیر مشاهداتی و محاسباتی این ۱۰ ایستگاه و محاسبه مجذور متوسط مربع خطا (RMSE)^۲ به برآوردی از دقت درون‌یابی رسیدند. نتیجه کار آنها نشان داد که روش کریجینگ فازی - ژنتیک که بر مبنای بهینه‌سازی پارامترهای تابع عضویت فازی با استفاده از الگوریتم ژنتیک عمل می‌کند، به ترتیب از دو روش کریجینگ فز و کریجینگ معمولی دارای RMSE کمتر و در نتیجه درون‌یابی بالاتر است.

در زمینه استفاده از روش‌های زمین آماری در برآورد بارندگی، پارادو ایگاز کوئیزا (۱۹۹۸) برای درون‌یابی مشاهدات بارندگی از چهار روش تیسسن^۳، کریجینگ معمولی، کوکریجینگ^۴ و کریجینگ با ترند خارجی^۵ استفاده کرد و به مقایسه نتایج آنها پرداخت. بر اساس نتایج اعتبارسنجی متقابل تحقیق وی، روش کریجینگ با حضور ترند خارجی بهترین خروجی را داشت. گوان و همکاران (۲۰۰۵) برای مطالعه بارندگی کوهستان از حالت بهینه‌شده کریجینگ معمولی استفاده کردند. آنها به منظور وارد کردن تأثیر اطلاعات کوهستانی^۶ و جوی در روند درون‌یابی خود، روش ASOADEK^۷ را معرفی کردند. در حقیقت مدل پیشنهادی آنها ترکیبی از کریجینگ معمولی و یک رگرسیون خطی چند متغیره بود که به دلیل در نظر گرفتن کارآمدتر اطلاعات مربوط به ارتفاع و جهت شیب، عملکرد بهتری در مقایسه با روش کریجینگ معمولی و کوکریجینگ داشت. چانگ و همکاران (۲۰۰۵) برای مطالعه یکی از حوضه‌های آبخیز کشور تایوان و مدل‌سازی آن، روشی را برای برآورد بارندگی

1. Fuzzy Genetic Membership Kriging
2. Root Mean Square Error
3. Thiessen
4. Cokriging
5. Kriging with an external drift
6. Orographic
7. Auto-Searched Orographic and Atmospheric Effects DetrendedKriging
8. Thiessen Polygons
9. Inverse Distance Weighting
10. spline

۲-۱- منطقه مطالعاتی

کریجینگ است (رابطه ۳). در کریجینگ معمولی فرض بر این است که برآوردی از کمیت مورد نظر در سطح منطقه وجود ندارد (وبستر و الیور، ۲۰۰۷). به عبارت دیگر، روش کریجینگ معمولی بر این فرض استوار است که برآوردی از میانگین کمیت مورد نظر در سطح منطقه درون‌یابی موجود نیست.

$$\hat{z}(x) = \sum_{i=1}^N \lambda_i z(x_i)$$

در رابطه (۳) مقادیر z مربوط به نقاط مشاهداتی و \hat{z} مقدار برآورد شده برای نقطه x است.

برای اطمینان از ناریب بودن^۷ برآورد، مجموع وزن‌ها (λ_i) برابر ۱ قرار می‌گیرد. مقادیر وزن‌ها بر اساس دستگاه $n+1$ معادله $n+1$ مجهول رابطه (۴) محاسبه می‌گردند.

رابطه (۴)

$$\sum_{i=1}^n \lambda_i \gamma(x_i, x_j) + \psi(x_o) = \gamma(x_j, x_o)$$

۹

$$\sum_{i=1}^n \lambda_i = 1$$

در رابطه (۴)، $\psi(x_o)$ ضریب لاگرانژ است که برای اطمینان از ناریب بودن برآورد استفاده می‌گردد. همچنین مقادیر $\gamma(x_j, x_o)$ و $\gamma(x_i, x_j)$ به ترتیب مقدار واریوگرام بین نقطه مجهول (x_o) و نقطه مشاهداتی z_j و مقدار واریوگرام بین دو نقطه مشاهداتی x_i و x_j است. پس از به دست آمدن ضرایب و جای‌گذاری آنها در رابطه (۳) مقدار کمیت مورد نظر برآورد می‌شود.

استان مازندران یکی از استان‌های پربارش کشور ایران است. میزان بارش باران در این استان در نقاط مختلف متنوع است. همچنین تعداد ایستگاه‌های سینوپتیک در این استان با توجه به وسعت آن نسبت به استان‌های دیگر دارای وضعیت بهتری است. در جدول ۱، نام و مختصات این ایستگاه‌ها به همراه میزان بارندگی بهمن ماه سال ۱۳۸۷ در آنها آورده شده است. در شکل ۱، موقعیت استان مازندران و پراکندگی انواع ایستگاه‌های هواشناسی در آن مشخص شده است.

۲-۲- کریجینگ معمولی

مهم‌ترین ابزار در زمین‌آمار^۸ واریوگرام است (لوید و آتکینسون، ۲۰۰۱). واریوگرام یا سمی‌واریانس^۹ ($\gamma(h)$) مطابق رابطه (۱) عبارت است از نصف امید ریاضی مربع اختلاف کمیت مورد نظر برای دو نقطه از یک متغیر تصادفی^۴ که با بردار h از هم جدا شده‌اند.

$$\gamma(h) = \frac{1}{2} E[\{z(x) - z(x+h)\}^2]$$

برای $n(h)$ جفت مشاهده می‌توان با استفاده از رابطه (۲)، واریوگرام تجربی^۵ ($\hat{\gamma}(h)$) را محاسبه کرد.

رابطه (۲)

$$\hat{\gamma}(h) = \frac{1}{2n(h)} \sum_{\alpha=1}^{n(h)} \{z_v(x_\alpha) - z_v(x_\alpha + h)\}^2$$

در رابطه (۲) z_v بیانگر مقادیر مشاهداتی در نقاط مختلف (x) است.

واریوگرام بیانگر ساختار مکانی متغیر مورد نظر است. به واریوگرام تجربی حاصل، یک واریوگرام با تابع مشخص ریاضی برآزش داده می‌شود که واریوگرام نظری^۶ نام دارد. کریجینگ یک برآوردگر است که به روش میانگین‌گیری وزن‌دار عمل می‌کند. وزن‌ها بر اساس مدل ریاضی که به واریوگرام تجربی برآزش داده شده‌اند، تعیین می‌گردند (لوید و آتکینسون، ۲۰۰۱).

کریجینگ معمولی متداول‌ترین نوع مورد استفاده از

1. MATLAB®
2. Geostatistics
3. Semivariance
4. Random Variable
5. Experimental Variogram
6. Theoretical Variogram
7. unbiased

۳-۲- کریجینگ شاخص

استفاده از واریوگرام معمولی و روش کریجینگ معمولی دارای یک سری محدودیت‌هاست. اگرچه این روش در شرایط غیرنرمال بودن توزیع متغیر هم نسبتاً کاراست، ولی فرضیات نرمال بودن توزیع متغیر و وابسته نبودن خطای استاندارد کریجینگ معمولی به مقادیر مشاهدات از مهم‌ترین محدودیت‌های آن محسوب می‌شوند. روش کریجینگ شاخص این دو محدودیت را برطرف می‌سازد (لوید و آتکینسون، ۲۰۰۱).

روش کریجینگ شاخص بر مبنای تبدیل غیرخطی داده‌ها با در نظر گرفتن یک سری حد آستانه عمل می‌کند. این روش می‌تواند هر دو محدودیت پیشین را با تبدیل داده‌ها به یک سری متغیرهای دوحالته برطرف سازد. اساس کار در این روش بر این است که ابتدا با توجه به رابطه (۵) مقادیر کمیت مورد نظر دوحالته می‌شوند.

رابطه (۵)

$$I(x; z_k) = \begin{cases} 1 & \text{if } z(x) \leq z_k \\ 0 & \text{otherwise.} \end{cases}$$

پس برای هر نقطه مشاهداتی و به ازای هر حد آستانه (z_k) ، یک مقدار ۰ یا ۱ به دست می‌آید که با $I(x; z_k)$ مشخص می‌گردد. به ازای هر حد آستانه و با توجه به ۰ و ۱ها، واریوگرام‌های شاخص^۱ مطابق رابطه (۶) محاسبه می‌گردند (لوید و آتکینسون، ۲۰۰۱).

رابطه (۶)

$$\hat{\gamma}(h; z_k)$$

در رابطه مذکور، $\hat{\gamma}(h; z_k)$ واریوگرام شاخص و $i(x_\alpha; z_k)$ مقدار شاخص برای نقطه x_α و به ازای حد آستانه z_k است.

با به دست آوردن مقادیر واریوگرام‌های شاخص، ادامه کار مانند کریجینگ معمولی است. بنابراین، برای هر نقطه مجهول و به ازای هر حد آستانه، مقداری بین ۰ و ۱ برآورد می‌شود که بیانگر احتمال کمتر یا مساوی بودن کمیت مجهول نقطه مورد نظر از آن حد آستانه است. بنابراین، محصول روش کریجینگ شاخص، احتمالی است که در آن کمیت یک نقطه مجهول از حد

آستانه تعریف شده کمتر یا بیشتر (در اینجا کمتر یا مساوی) باشد.

اگر برای مجموعه‌ای از حدود آستانه این احتمالات محاسبه گردند، می‌توان تابع توزیع تجمعی شرطی را برای هر نقطه مجهول از رابطه (۷) به دست آورد. رابطه (۷)

$$F(u; z | (n)) = \text{prob}\{z(u) \leq z | (n)\}$$

تابع F احتمال کمتر بودن از یا مساوی بودن مقدار z در نقطه u را نسبت به مقدار حد آستانه z با توجه به n نقطه همسایه که در درون‌یابی شرکت کرده‌اند، نتیجه می‌دهد. تابع مذکور ابزاری آماری برای مدل‌سازی محلی عدم قطعیت در نقطه مجهول است (گووآرتس، ۲۰۰۱ و ۲۰۰۵). افزون بر این، با کمک تابع F ، می‌توان یک مقدار محاسباتی را برای کمیت مورد نظر برآورد کرد.

۴-۲- کریجینگ شاخص نرم

فرض می‌شود با توجه به رابطه (۵) و حد آستانه ۵۰ میلی‌متری بخواهیم تبدیل کریجینگ شاخص را انجام دهیم. در این صورت، هیچ تفاوتی در نتیجه تبدیل اعداد ۵۰، ۴۸ و ۴۰ میلی‌متر وجود نخواهد داشت و همگی به عدد ۱ تبدیل می‌گردند. این مورد مانند آن است که هیچ‌گونه عدم قطعیتی در خصوص این مقادیر اندازه‌گیری شده وجود ندارد - به‌ویژه در مواقعی که دقت دستگاه‌های اندازه‌گیری پایین است چنین تصویری نادرست می‌نماید. برای این منظور، سایتو و گووآرتس در سال ۲۰۰۲، کریجینگ شاخص نرم را پیشنهاد کردند.

کریجینگ شاخص نرم برخلاف کریجینگ شاخص، عمل تبدیل‌سازی را محدود به دو حالت صفر یا یک نمی‌کند؛ بلکه با در نظر گرفتن یک منحنی نرمال برای هر نقطه مشاهداتی (که با میانگین (μ) و انحراف معیار (σ) تعریف می‌گردد) مقداری بین ۰ تا ۱ را برای هر

1. Indicator Variogram

است؛ بدین ترتیب که به ازای هر حد آستانه، واریوگرام‌های شاخص بر اساس مقادیر زمحاسبه می‌شوند و درون‌یابی بر اساس آنها صورت می‌پذیرد. بنابراین، محصول این روش نیز احتمالات مربوط به عبور یا عدم عبور از یک سری حدود آستانه تعریف شده برای هر نقطه مجهول است؛ با این تفاوت که در این روش به علت ماهیت متفاوت تبدیل مشاهدات، عدم قطعیت مربوط به اندازه‌گیری مشاهدات هر نقطه نیز وارد محاسبات مربوط به این احتمالات می‌گردد و خطاهای اندازه‌گیری هم مورد بررسی قرار می‌گیرند (سایتو و گووارتس، ۲۰۰۲).

حد آستانه (Z_k) با توجه به رابطه (۸) به هر نقطه اختصاص می‌دهد (سایتو و گووارتس، ۲۰۰۲). در حقیقت، تبدیل در روش کریجینگ شاخص نرم، با توجه به دقت اندازه‌گیری مشاهده که با پارامتر σ تعریف می‌گردد، صورت می‌پذیرد.

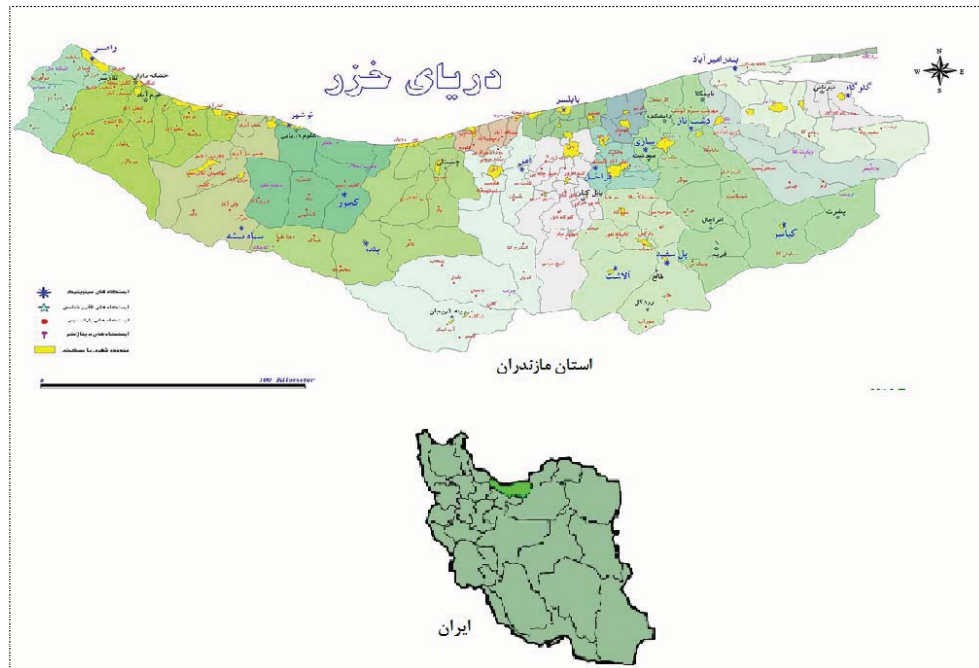
رابطه (۸)

$$j(u_\alpha; z_k) = \text{prob}\{Z(u_\alpha) \leq z_k | (n)\} = G\{z_k - \bar{z}(u_\alpha) / s(u_\alpha)\}$$

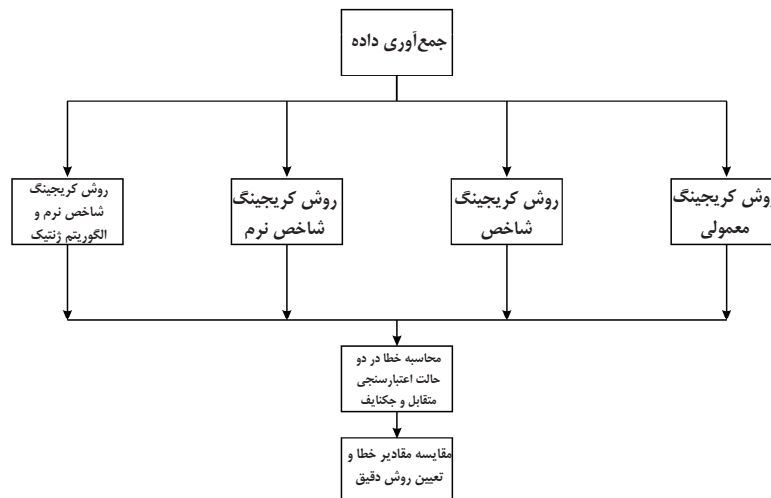
در رابطه (۸)، $G(\cdot)$ همان تابع توزیع تجمعی استاندارد نرمال است. مقادیر z که بین ۰ و ۱ قرار دارند، نمایانگر عدم قطعیت مربوط به عبور یا عدم عبور از هر حد آستانه‌اند. ادامه کار مانند کریجینگ شاخص

جدول ۱. مشخصات ایستگاه‌های سینوپتیک استان مازندران و میزان بارندگی مربوط به بهمن ماه سال ۱۳۸۷ در آنها

نام ایستگاه	X (UTM)	Y (UTM)	ارتفاع (متر)	بارندگی (میلی متر)
نوشهر	544691.25	4056163.5	-20.9	115.5
ساری	679006	4046815.2	22.07	119.3
قراخیل	663145.3	4033182.5	14.7	88.6
بابلسر	647353.06	4065080.8	-21	77.4
دشت ناز	695207	4056418	16	106.1
بندر امیرآباد	721113	4081104	-20	118.7
رامسر	467926.6	4083839.8	-20	130.7
آمل	631701.4	4037084.8	29	79.6
پل سفید	684586.94	3994767	610	108.8
کیاسر	729179.2	4012474.5	1294	113.3
آلاشت	662136	3990993	1900	92.5
گلوگاه	750040	4068577	-10	122.8
سیاه بیشه	526952.7	4011719.5	1855.4	78
کجور	539540.7	4009549	1550	58.1
بلده	543195.75	3997365	2120	43



شکل ۱. استان مازندران و نحوه پراکندگی انواع ایستگاه‌های هواشناسی در آن



شکل ۲. روش اجرای تحقیق

استفاده از الگوریتم ژنتیک و تعریف یک تابع هدف، حدود آستانه مربوط به تبدیل مشاهدات در کریجینگ شاخص نرم طوری بهینه‌سازی می‌گردند تا مقدار تابع هدف تعریف‌شده کمینه گردد.

۵-۲- کریجینگ شاخص نرم بهینه‌شده با الگوریتم ژنتیک
در این روش که روش پیشنهادی تحقیق حاضر است، با

ژنتیک طوری هدایت می‌شود که مقدار f کمینه گردد. برای ارزیابی دقیق‌تر روش‌ها، افزون بر حالت اعتبارسنجی متقابل، از روش جک‌نایف نیز در دو آزمایش استفاده شد. در هر آزمایش، ۳ ایستگاه به طور کامل از فرایند محاسبات حذف شدند و داده‌های آنها به عنوان نقاط کنترل در نظر گرفته شدند. در ادامه، روش‌های مختلفی که بر روی این داده‌ها اعمال شده‌اند، مورد بررسی قرار گرفته‌اند و در پایان، نتایج در قالب جدول ۲ ارائه شده‌اند.

۳-۱- کریجینگ معمولی

روش کریجینگ معمولی تنها روشی است که ورودی آن داده‌های اندازه‌گیری شده هستند و هیچ‌گونه تبدیلی بر روی داده‌ها صورت نمی‌پذیرد. به همین دلیل می‌توان ادعا کرد که سرعت این روش از سایر روش‌ها بیشتر است (آتکینسون و لوید، ۱۹۹۸). برای برآزش مدل ریاضی به واریوگرام‌های تجربی، از تابع کروی^۸ - که یکی از متداول‌ترین مدل‌هایی است که به کار می‌رود - استفاده شده است. البته تعداد اندک مشاهدات در این تحقیق، باعث شده است که واریوگرام به طور مناسب برآورد نشود و تنها راه برطرف کردن این مشکل، افزایش ایستگاه‌های سنجش بارندگی در سطح منطقه است. در همین زمینه بوستر و اولیور (۱۹۹۳) نشان داده‌اند که برای برآورد قابل اطمینان واریوگرام به حداقل ۱۵۰ نقطه مشاهداتی نیاز است. البته، این یافته‌ها بیانگر این نکته نیستند که روش کریجینگ برای تعداد کمتر داده قابل اجرا نیست (گووآرتس، ۱۹۹۹). اما از آنجا که هدف این تحقیق،

الگوریتم ژنتیک یکی از روش‌های بهینه‌سازی است که بر مبنای نظریه داروین عمل می‌کند. در این روش، با تعریف یک جمعیت تصادفی از کروموزوم‌ها^۱ و عملیات مربوط به انتخاب^۲، تلفیق^۳ و جهش ژنتیکی^۴، جمعیت اولیه طوری هدایت می‌شود که جمعیتی که جواب بهینه در آن قرار دارد، حاصل گردد. تعیین ارزش هر جواب بر اساس یک تابع هدف از پیش تعیین شده است. در این روش، هر جواب بالقوه به صورت یک کروموزوم یا بردار است که از مجموعه‌ای از ژن‌ها^۵ که همان اعضای جواب‌ها هستند، تشکیل می‌گردد.

در شکل ۲، مراحل صورت گرفته در این تحقیق نمایش داده شده است.

۳- پیاده‌سازی و نتایج

به منظور بررسی خطای برآورد مربوط به هر روش درون‌یابی در این تحقیق، هر مشاهده به طور موقت از روند محاسبات خارج می‌شود و مقدار محاسباتی آن پس از اجرای برنامه به دست می‌آید. این روش اعتبارسنجی متقابل^۶ نام دارد و برخلاف روش جک‌نایف^۷ است که در آن تعدادی از مشاهدات به طور کامل از روند محاسبات خارج می‌شوند و در انتها پس از یک بار اجرای برنامه، مقدار محاسباتی آنها به دست می‌آید (گووآرتس، ۲۰۰۱ و ۲۰۰۵). اگرچه در روش اعتبارسنجی متقابل زمان انجام محاسبات افزایش می‌یابد ولی در عوض از تمامی مشاهدات استفاده می‌شود. از طرفی به دلیل اینکه تعداد مشاهدات در این تحقیق نسبتاً اندک بود، روش اعتبارسنجی متقابل گزینه مناسبی برای بررسی دقت به نظر می‌رسید. در تمامی روش‌ها، خطای برآورد به صورت رابطه (۹) تعریف شده است.

$$f = \frac{\sum_{i=1}^{15} |x_i^o - x_i^c|}{15} \quad \text{رابطه (۹)}$$

در رابطه (۹)، x^o و x^c به ترتیب مقادیر مشاهداتی و محاسباتی برای هر نقطه هستند. الگوریتم

1. Chromosome
2. Selection
3. Crossover
4. Mutation
5. Gene
6. Cross-Validation
7. Jackknife
8. Spherical

سرانجام با توجه به روابط کریجینگ معمولی (روابط ۳ و ۴) که به ازای هر حد آستانه بر مشاهدات تبدیل یافته مربوط به آن حد آستانه اعمال شدند و همچنین رابطه (۷) برای هر نقطه مشاهداتی، CCDF محاسبه شد.

برای برآورد یک مقدار محاسباتی از CCDF که محصول اصلی روش کریجینگ شاخص (و همین طور روش کریجینگ شاخص نرم و کریجینگ شاخص نرم بهینه شده با الگوریتم ژنتیک) است، به عنوان نمونه، لویید و آتکینسون (۲۰۰۱) برآورد E-type را معرفی کرده اند که از رابطه (۱۰) به دست می آید.

رابطه (۱۰)

$$\hat{Z}_E(x) = \sum_{k=1}^K \bar{z}_k \cdot [F(x; z_{k+1} | (n)) - F(x; z_k | (n))]$$

در این رابطه، مقدار \bar{z}_k ، برابر میانگین مقادیر حدود آستانه در بازه $[z_k, z_{k+1}]$ و K برابر تعداد بازه‌هایی است که به ازای آنها مقادیر \bar{z}_k محاسبه می‌گردند. همچنین تابع F ، مربوطه به تابع CCDF به دست آمده از روش کریجینگ شاخص برای دو حد آستانه z_k و z_{k+1} است. در این تحقیق، از فرایند میانگین گیری وزن دار برای برآورد یک مقدار محاسباتی از CCDF استفاده گردید. در این فرایند، عبارت $dF(x, k) = \frac{F(x, z_{k+1}) - F(x, z_k)}{z_{k+1} - z_k}$ که برای K زوج حد آستانه z_k و z_{k+1} مربوط به بارندگی محاسبه گردید و همان مشتق تابع CCDF (تابع چگالی احتمال^۱) است به عنوان وزن در نظر گرفته شد. سپس مطابق رابطه (۱۱)، این وزن‌ها برای برآورد مقدار محاسباتی بارندگی برای نقطه مجهول x به کار رفتند.

رابطه (۱۱)

$$\hat{z}(x) = \frac{\sum_{k=1}^K \bar{z}_k \times dF(x, k)}{\sum_{k=1}^K dF(x, k)}$$

1. Probability Density Function (PDF)

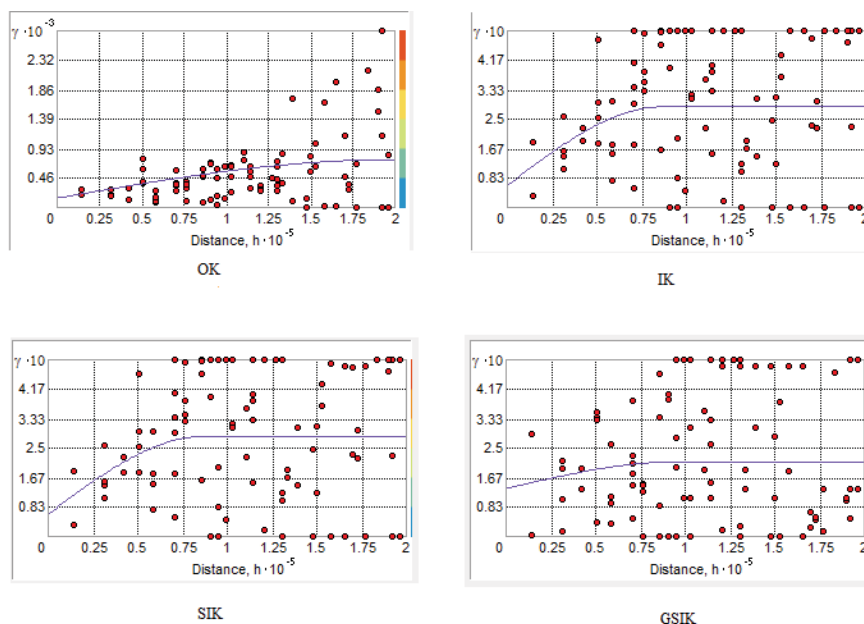
تعیین دقیق ترین روش با توجه به شرایط یکسان داده‌ها بوده است، از همین تعداد اندک مشاهدات برای برآورد واریوگرام تجربی و برازش واریوگرام نظری استفاده شده است. همچنین طبق آزمایش‌هایی که برای این تحقیق صورت گرفت، روش‌های زمین آماری با همین واریوگرام باز هم بسیار دقیق تر از روش غیرزمین آماری میانگین گیری وزن دار بر اساس معکوس فاصله بودند. شکل ۳، بیانگر یک نمونه از واریوگرام تجربی و واریوگرام نظری برازش داده شده بر آن برای هر روش است. این واریوگرام‌ها مربوط به شرایطی هستند که نقطه هفتم در حالت اعتبارسنجی متقابل از روند محاسبات خارج شده است. در دو روش کریجینگ شاخص و کریجینگ شاخص نرم، واریوگرام‌ها مربوط به حد آستانه ۱۰۴ میلی متری هستند و برای روش کریجینگ شاخص نرم بهینه شده با الگوریتم ژنتیک، واریوگرام ترسیم شده مربوط به حد آستانه بهینه شده ۱۱۳/۳۸۵ میلی متر است. در اجرای روش کریجینگ معمولی، پس از محاسبه وزن‌ها مطابق رابطه (۴) با استفاده از رابطه (۳)، مقدار برآورد شده برای هر ایستگاه محاسبه شده است.

۳-۲- کریجینگ شاخص

در این روش، برای تبدیل دو حالت داده‌ها، ابتدا باید یک سری حدود آستانه تعریف شود. این حدود آستانه باید به گونه‌ای تعریف گردند که محدوده مشاهدات را دربرگیرند (لازم به ذکر است که واحد تمامی حدود آستانه، میلی متر است):

$$T = [40 \quad 56 \quad 72 \quad 88 \quad 104 \quad 120 \quad 136]$$

مطابق این حدود آستانه و رابطه (۵)، تبدیل دو حالت صورت پذیرفت. سپس مطابق رابطه (۶)، واریوگرام شاخص به ازای هر حد آستانه و برای مشاهدات تبدیل شده محاسبه گردید. در این روش و دو روش بعدی، همانند روش کریجینگ معمولی از تابع کروی برای برازش بر واریوگرام‌های تجربی استفاده شد.



شکل ۳. واریوگرام‌های تجربی و نظری برای چهار روش کریجینگ معمولی (OK)، کریجینگ شاخص (IK)، کریجینگ شاخص نرم (SIK) و کریجینگ شاخص نرم بهینه‌شده با الگوریتم ژنتیک (GSIK)

ازای هر حد آستانه، واریوگرام‌ها محاسبه شدند و درون‌یابی کریجینگ انجام گرفت. همانند روش قبل، خروجی برای هر نقطه یک تابع CCDF است، و در پایان به روشی مشابه آنچه در بخش ۳-۲ بیان شد، برآورد از این توابع برای محاسبه دقت درون‌یابی انجام گرفت.

در رابطه (۱۱)، میانگین دو حد آستانه Z_k و Z_{k+1} است که در این تحقیق برحسب میلی‌متر است. بنابراین، با این روش میانگین‌گیری وزن‌دار، مقادیر محاسباتی بارندگی برای ایستگاه‌ها که به منظور محاسبه دقت درون‌یابی به آن‌ها نیاز بود، به دست آمدند.

۳-۴- کریجینگ شاخص نرم بهینه‌شده با

الگوریتم ژنتیک

آخرین روشی که مورد استفاده قرار گرفته است، استفاده از کریجینگ شاخص نرم به همراه الگوریتم ژنتیک است. در این روش، حدود آستانه که تأثیری مستقیم در فرایند درون‌یابی دارند، به عنوان متغیرهای ورودی الگوریتم ژنتیک (یا همان کروموزوم‌ها) انتخاب شد. به ازای هر ترکیبی از حدود آستانه که به وسیله الگوریتم ژنتیک ایجاد گردید، واریوگرام‌های مربوط به حدود آستانه آن ترکیب محاسبه شد و مطابق بخش ۳-۳، درون‌یابی به روش کریجینگ شاخص نرم صورت

۳-۳- کریجینگ شاخص نرم

در این روش، به جای تبدیل باینری یا سخت، از تبدیل نرم بر اساس تابع توزیع نرمال استفاده می‌شود. بنابراین برای هر نقطه مشاهداتی یک تابع توزیع نرمال تعریف شد. برای مقدار میانگین (μ) توزیع، مقدار مشاهداتی هر نقطه در نظر گرفته شده است. با توجه به اینکه خطای اسمی دستگاه‌های سنجش بارندگی برای مشاهدات ماهانه برابر ۱ میلی‌متر است (وبگاه اداره هواشناسی استان یزد)، پارامتر σ هم برابر ۱ میلی‌متر قرار گرفت. سپس با توجه به این پارامترها، حدود آستانه و رابطه (۸)، تبدیل مشاهدات به مقادیر بین ۰ و ۱ انجام پذیرفت. مجدداً برای اعداد به‌دست‌آمده و به

اعتبارسنجی متقابل است که از تمامی داده‌ها در آن استفاده شده است.

مقادیر حدود آستانه بهینه‌شده برای روش ترکیبی کریجینگ شاخص نرم و الگوریتم ژنتیک به صورت زیر محاسبه گشت.

$$T = \text{[بهینه شده]} \\ = [25/951 \quad 59/172 \quad 67/221 \quad 87/15 \quad 112/285 \quad 125/969 \quad 122/452]$$

برای حالت جکنایف، در قسمت اول، سه ایستگاه پل سفید، کجور و بندر امیرآباد و در قسمت دوم، سه ایستگاه آمل، کیاسر و سیاه‌بیشه به عنوان نقاط کنترل در نظر گرفته شدند. برای محاسبه خطا در این حالت، رابطه (۹) باید برای سه نقطه در نظر گرفته شود. نتایج در جدول‌های ۳ و ۴ آورده شده‌اند.

برای مقایسه بهتر بین نتایج حاصل از هر روش، از آماره R^2 هم استفاده شد. این آماره برای بررسی ارتباط و همبستگی^۵ بین دو متغیر تعریف می‌گردد و مقدار آن بین صفر (عدم همبستگی) و ۱ (همبستگی کامل) است. برای دو متغیر y و y_{fit} ، این آماره مطابق رابطه (۱۳) تعریف می‌گردد:

$$TSS = \sum_{i=1}^n (y_i - \text{mean}(y))^2 \\ RSS = \sum_{i=1}^n (y_i - y_{fit_i})^2 \\ R^2 = 1 - \frac{RSS}{TSS} \quad \text{رابطه (۱۳)}$$

بنابراین هر چه این آماره برای یک روش بیشتر باشد بیانگر دقیق‌تر بودن آن روش است. در جدول ۵، مقادیر R^2 برای روش‌های مختلف انجام‌شده در این تحقیق در حالت اعتبارسنجی متقابل آورده شده است.

گرفت. پس از به دست آوردن مقادیر محاسباتی برای آن ترکیب حدود آستانه، یک مقدار برای دقت ترکیب برآورد شد. این روند برای ترکیب‌های مختلف حدود آستانه تکرار شد، تا حدود آستانه بهینه‌شده که به ازای آن دقت درون‌یابی به بهترین مقدار خود در یک جمعیت رسید، تعیین گردید. سرانجام، ترکیب بهینه‌شده حدود آستانه به همراه مقادیر محاسباتی و دقت مربوط به آن به عنوان خروجی روش ترکیبی ارائه شده‌اند.

برای بررسی دقیق‌تر، روش مذکور در سه حالت بر روی داده‌ها انجام شد. در حالت اول برای هر نقطه مشاهداتی، تمامی داده‌ها به جز نقطه مورد نظر وارد فرایند درون‌یابی شدند و به اصطلاح حالت اعتبارسنجی متقابل انجام گرفت. در دو حالت بعدی، با استفاده از روش جکنایف، سه نقطه مشاهداتی از داده‌ها حذف شدند و مقدار محاسباتی این سه نقطه با مقادیر مشاهداتی آنها مقایسه شد. دلیل دوبار استفاده از حالت جکنایف این بود که تأثیر انتخاب سه نقطه کنترل در نتایج نهایی هم مورد بررسی قرار گیرد. برای ترکیب اولیه حدود آستانه، همان مقادیر بخش ۳-۲ در نظر گرفته شدند تا در الگوریتم ژنتیک بهینه شوند.

در به کارگیری الگوریتم ژنتیک (که بر روی اعداد حقیقی صورت گرفت)، تعداد جمعیت^۱ برابر ۲۰، نرخ فرزندان نخبه^۲ که به جمعیت بعدی منتقل می‌شوند برابر ۲ و نسبت فرزندان^۳ که از تلفیق والدین می‌شوند برابر ۰/۸ قرار داده شد.

همچنین قیدی^۴ که در بهینه‌سازی ژنتیک مورد استفاده قرار گرفت. به این صورت بود که حدود آستانه (T_i) دارای ترتیب افزایشی بودند (رابطه ۱۲).

$$T_1 < T_2 < T_3 < T_4 < T_5 < T_6 < T_7 \quad \text{رابطه (۱۲)}$$

در انتها، تمامی روش‌های ذکر شده بر روی مجموعه داده‌های موجود انجام شد. در جدول ۲، نتایج به‌دست‌آمده آورد شده‌اند. این جدول مربوط به حالت

1. Population Size
2. Elite count
3. Crossover Fraction
4. Constraint
5. Correlation

روشی تازه به منظور درون‌یابی مشاهدات بارندگی با کمک کریجینگ شاخص نرم و الگوریتم ژنتیک

جدول ۲. مقادیر مشاهداتی و محاسباتی مربوط به روش‌های مختلف برای ایستگاه‌های سینوپتیک استان مازندران

مقدار محاسباتی (SIK&GA)	مقدار محاسباتی (SIK)	مقدار محاسباتی (IK)	مقدار محاسباتی (OK)	مقدار مشاهداتی	نام ایستگاه
92.8	85.9	82.7	84.7	115.5	نوشهر
100.3	99.4	99.5	97.1	119.3	ساری
102.1	98.9	97	101.7	88.6	قراخیل
100.4	95.7	95.2	98.1	77.4	بابلسر
119.2	114.2	111.1	118.4	106.1	دشت ناز
111.4	117.4	116.9	112.3	118.7	بندر امیرآباد
113.1	104.9	105.2	107.2	130.7	رامسر
79.5	79.5	81.5	76.1	79.6	آمل
108.6	102.1	102.5	102.5	108.8	پل سفید
108	112.8	112.8	115.8	113.3	کیاسر
80.8	99.1	100	95.9	92.5	آلاشت
115.7	105.6	104.4	115.6	122.8	گلوگاه
77.7	76.8	84.5	68.8	78	سیاه‌بیشه
61.4	73.1	73	63.2	58.1	کجور
52.9	74.7	74.8	61.4	43	بلده
f=10.3	f=12.8	f=13.3	f=12.3		

جدول ۳. نتایج حاصل از روش جکنایف برای سه ایستگاه انتخابی اول

مقدار محاسباتی (SIK&GA)	مقدار محاسباتی (SIK)	مقدار محاسباتی (IK)	مقدار محاسباتی (OK)	مقدار مشاهداتی	نام ایستگاه
105.4	103.2	97.5	102.4	108.8	پل سفید
58.1	66.6	71.1	63.1	58.1	کجور
115	118.2	115.4	112.2	118.7	بندر امیرآباد
f=2.4	f=4.8	f=9.2	f=5.9		

$$T(\text{بهینه‌شده}) = [41.225 \ 51.434 \ 67.843 \ 92.71 \ 112.114 \ 121.18 \ 131.121]$$

جدول ۴. نتایج حاصل از روش جکنایف برای سه ایستگاه انتخابی دوم

مقدار محاسباتی	مقدار محاسباتی	مقدار محاسباتی	مقدار محاسباتی	مقدار	نام ایستگاه
(SIK&GA)	(SIK)	(IK)	(OK)	مشاهداتی	
85.8	89.6	91.8	76.9	79.6	آمل
113.3	108.5	108.2	116.1	113.3	کیاسر
78	76.7	76.8	68.9	78	سیاه بیشه
f=2	f=5.4	f=6.2	f=4.9		

$$T(\text{بهینه شده}) = [44.801 \quad 52.024 \quad 70.991 \quad 82.001 \quad 100.001 \quad 118.65 \quad 131.759]$$

جدول ۵. مقادیر آماره R^2

روش	R^2
کریجینگ معمولی	0.6346
کریجینگ شاخص	0.5369
کریجینگ شاخص نرم	0.5571
کریجینگ شاخص نرم و الگوریتم ژنتیک	0.7330

۴- بحث و نتیجه گیری

در مقایسه با روش‌های دیگر به کار رفته در این تحقیق دارای دقت درون‌یابی بالاتری است. بنابراین نقش الگوریتم ژنتیک در بهبود دقت درون‌یابی از طریق آموزش خودکار حدود آستانه دخیل در تبدیل روش کریجینگ شاخص نرم مشخص می‌گردد. بنابراین، می‌توان چنین نتیجه‌گیری کرد که استفاده هم‌زمان کریجینگ شاخص نرم و الگوریتم ژنتیک باعث مدل‌سازی دو منبع خطای روش کریجینگ شاخص شده است، به نحوی که دقت روش ترکیبی در این تحقیق از روش کریجینگ معمولی هم بهتر شده است. افزون بر این، می‌توان چنین نتیجه گرفت که الگوریتم ژنتیک ضمن اینکه در افزایش دقت در روش‌هایی که خروجی آنها یک عدد است (مانند تحقیق چانگ و همکاران (۲۰۰۵)، شاد و همکاران (۲۰۰۹) و معصومی و همکاران (۲۰۱۱) مؤثر است، باعث افزایش دقت در روش کریجینگ شاخص نرم (که خروجی آن تابع CCDF است) نیز می‌گردد.

در هر سه حالت درون‌یابی، روش کریجینگ معمولی دقت بالاتری در مقایسه با روش کریجینگ شاخص داشته است، اما دلیل برتری همبستگی روش

همان‌طور که در جداول ۲، ۳، ۴ و ۵ مشاهده می‌شود، روش ترکیبی کریجینگ شاخص نرم و الگوریتم ژنتیک بیشترین دقت درون‌یابی را در تحقیق حاضر داشته‌اند. مطابق جدول ۲، مقدار تابع f که بیانگر خطای درون‌یابی است، در روش ترکیبی کمینه است، و این موضوع نشان‌دهنده دقت بالاتر این روش در حالت اعتبارسنجی متقابل است. در حالت جکنایف، روش ترکیبی که در آن از دو ترکیب مختلف نقاط کنترل استفاده شده، دارای دقت بالاتری است؛ زیرا مطابق جداول ۳ و ۴، مقدار f برای این روش در مقایسه با روش‌های دیگر کمینه است. همچنین برای بررسی بیشتر، از آماره R^2 هم استفاده شده است. دو متغیری که این آماره برای آنها محاسبه شده، مقادیر مشاهداتی و محاسباتی نقاط در حالت اعتبارسنجی متقابل است. بنابراین، مقدار بیشتر این آماره را می‌توان بیانگر همبستگی بیشتر مقادیر مشاهداتی و محاسباتی و در نتیجه دقت درون‌یابی بالاتر دانست. مطابق جدول ۵، مقدار این آماره برای روش ترکیبی بیشینه است. بنابراین باز هم نتیجه‌گیری می‌شود که روش ترکیبی

پیشنهادی بالاترین دقت را داشته است، ولی هیچ‌کدام از روش‌ها دارای دقت بالایی نیستند. به نظر می‌رسد دلیل این موضوع کم بودن تعداد مشاهدات نسبت به وسعت منطقه است، که یکی از مشکلات اصلی درون‌یابی به شمار می‌آید. نکته دیگر اینکه چگونگی تنظیم پارامترهای الگوریتم ژنتیک مانند تنظیمات مربوط به پارامترهای تلفیق و جهش، نرخ فرزندان نخبه، توابع انتخاب، معرفی قیود و جمعیت اولیه هم نقش بسزایی در چگونگی عملکرد الگوریتم ژنتیک دارد که در انتخاب آنها باید دقت فراوانی کرد. بررسی تأثیرات انتخاب این پارامترها در نتایج نهایی می‌تواند به عنوان موضوعی جدید در آینده مورد بررسی قرار گیرد.

۵- منابع

- Atkinson, P. M. , Lloyd C. D., 1998, **Mapping Precipitation in Switzerland with Ordinary and Indicator Kriging**, Journal of Geographic Information and Decision Analysis, Vol. 2, 65-76.
- Burrough, P. and McDonnell, R., 1998, **Principles of Geographic Information Systems**, Oxford University Press, New York.
- Chang, C. L. , Lo, S. L. , Yu, S. L., 2005, **Applying Fuzzy Theory and Genetic Algorithm to Interpolate Precipitation**, Journal of Hydrology, 1-4, Vol. 314., 92-104.
- Chu, H. J. , Lin, Y. P. , Jang, C. S. , Chang, T. K., 2010, **Delineating the Hazard Zone of Multiple Soil Pollutants by Multivariate Indicator Kriging and Conditioned Latin Hypercube Sampling**, Geoderma, 3-4 , Vol. 158., 242-251.

کریجینگ معمولی در مقایسه با روش کریجینگ شاخص صرفاً دقت بالاتر نیست. هر کدام از این دو روش نقاط قوت و ضعفی دارند. اگر چه روش کریجینگ معمولی، روش استواری است^۱ اما این روش بر مبنای نرمال بودن توزیع مشاهدات عمل می‌کند و به نوعی داده‌ها باید دارای توزیع همگن باشند تا این روش بتواند به خوبی عمل کند. اما روش کریجینگ شاخص به هیچ پیش‌فرض اولیه‌ای برای توزیع داده‌ها احتیاج ندارد، زیرا بر روی تمامی داده‌ها یک تبدیل غیرخطی صورت می‌پذیرد (لوید و آتکینسون، ۲۰۰۱). همچنین اگرچه روش کریجینگ شاخص به علت همین تبدیل داده‌ها از روش کریجینگ معمولی کندتر است، اما می‌تواند برآوردی از خطای محلی را برای هر نقطه به وسیله CCDF به دست دهد. شاید بتوان دلیل دقت کمتر روش کریجینگ شاخص نسبت به روش کریجینگ معمولی در این تحقیق را در کم بودن تعداد مشاهدات جست‌وجو کرد، زیرا با توجه به گسسته‌سازی داده‌ها در روش کریجینگ شاخص بر اساس حدود آستانه، تأثیر تعداد اندک مشاهدات در این روش در مقایسه با روش کریجینگ معمولی بیشتر است.

نکته دیگر، بالاتر بودن دقت روش کریجینگ شاخص نرم در مقایسه با روش کریجینگ شاخص است. این بهبود دقت، بر اثر داخل کردن خطای اندازه‌گیری مشاهدات به وسیله تابع توزیع نرمال نقاط مشاهداتی در فرایند درون‌یابی است. این نوع تبدیل باعث انطباق بیشتر درون‌یابی با واقعیت می‌گردد. بنابراین نقش تبدیل نرم داده‌ها به جای تبدیل سخت یا دو حالتی آنها در مدل‌سازی خطای اندازه‌گیری مشاهدات و به دنبال آن در افزایش دقت ثابت می‌گردد. با آموزش و بهینه‌سازی روش کریجینگ شاخص به کمک الگوریتم ژنتیک، می‌توان به بهترین سطح دقت در میان تمامی روش‌های ذکر شده در پژوهش حاضر رسید. البته ذکر این نکته ضروری است که گرچه روش

1. Robust

- Goovaerts, P., 1999, **Geostatistics in Soil Science: State-of-the-art and Perspectives**, Geoderma, 89(1-2), 1-45.
- Goovaerts, P., 2001, **Geostatistical Modelling of Uncertainty in Soil Science**, Geoderma, 1-2, Vol. 103, 3-26.
- Goovaerts, P., 2005, **Geostatistical Modeling of the Spaces of Local**, Spatial and Response Uncertainty for Continuous Petrophysical Properties [Book Section]// Stochastic modelling and geostatistics: Principles, methods, and case studies / book auth. Coburn, T. C., Yarus, J. M. and Chambers, R. L.
- Guan, H., Wilson, J. L., Makhnin, O., 2005, **Geostatistical Mapping of Mountain Precipitation Incorporating Autosearched Effects of Terrain and Climatic Characteristics**, Journal of Hydrometeorology, 6(6), 1018.
- Guo, D., Danni, R., Thiart, C., 2007, **Predicting Air Pollution Using Fuzzy Membership Grade Kriging**, Computers, Environment and Urban Systems, 1, Vol. 31, 33-51.
- Journel, A. G., 1983, **Nonparametric Estimation of Spatial Distributions**, Mathematical Geology, 445-468.
- Juang, K.W., Liao, W. J., Liu, T. L., Tsui, L., Lee, D. Y., 2008, **Additional Sampling Based on Regulation Threshold and Kriging Variance to Reduce the Probability of False Delineation in a Contaminated Site**, Science of The Total Environment, 1, Vol. 389, 20-28.
- Krige, D. G., 1951, **A Statistical Approach to Some Basic Mine Valuation Problems on the Witwatersrand**, Journal of the Chemical, Metallurgical and Mining Society of South Africa, 119-139.
- Lark, R. M., Ferguson, R. B., 2004, **Mapping Risk of Soil Nutrient Deficiency or Excess by Disjunctive and Indicator Kriging**, Geoderma, 1-2, Vol. 118, 39-53.
- Lee, J. J., Jang, C. S., Wang, S. W., Liu, C. W., 2007, **Evaluation of Potential Health Risk of Arsenic-affected Groundwater Using Indicator Kriging and Dose Response Model**, Science of The Total Environment, 1-3, Vol. 384, 151-162.
- Lloyd, C. D., Atkinson, P. M., 2001, **Assesing Uncertainty in Estimates with Ordinary and Indicator Kriging**, Computer & Geosciences.
- Lyon, S.W., Lembo, J. A. J., Walter, M. T., Steenhuis, T. S., 2006, **Defining Probability of Saturation with Indicator Kriging on Hard and Soft Data**, Advances in Water Resources, 2, Vol. 29, 181-193.
- Masoomi, Z., Mesgari, M.S., Menhaj, M. B., 2011, **Modeling Uncertainties in Sodium Spatial Dispersion Using a Computational Intelligence-based Kriging Method**, Computers & Geosciences. In Press, Corrected Proof.
- Meteorological Devices2 [Online]//Yazd Meteorological Organization. <http://www.yazdmet.ir/SC.php?type=static&id=142>.

- Pardo-Igúzquiza, E., 1998, **Comparison of Geostatistical Methods for Estimating the Areal Average Climatological Rainfall Mean Using Data on Precipitation and Topography**, International Journal of Climatology, 18, 1031-1047.
- Saito, H. , Goovaerts, P, 2002, **Accounting for Measurement Error in Uncertainty Modelling and Decision-Making Using Indicator Kriging and P-field Simulation: Application to a Dioxin Contaminated Site**, Environmetrics.
- Shad, R. , Mesgari, M. S., Abkar, A., Shad, A., 2009, **Predicting Air Pollution Using Fuzzy Genetic Linear Membership Kriging in GIS**, Computers, Environment and Urban System, 6 , Vol. 33.
- Webster, R. , Oliver M. A., 1993, **How Large a Sample is Needed to Estimate the Regional Variogram Adequately?** In Soares, A. (Ed.), Geostatistics Troia '92, Kluwer Academic Publishers, Dordrecht, 155-166.
- Webster, R. , Oliver, M. A., 2007, **Geostatistics for Environmental Scientists**, John Wiley & Sons, Ltd.
- <http://www.yazdmet.ir/SC.php?type=static&id=142>.