درونیابی داده‌های مکانی با حل بازگشتی روش کمترین مربعات متحرک

حمید مهرابی[۱]، بهنام تشیع[*۱]

[۱] استادیار گروه مهندسی نقشه‌برداری - دانشکده عمران و حمل‌ونقل - دانشگاه اصفهان
{h.mehrabi, *b.tashayo}@eng.ui.ac.ir

(تاریخ دریافت شهریور ۱۳۹۷، تاریخ تصویب تیر ۱۳۹۸)

چکیده

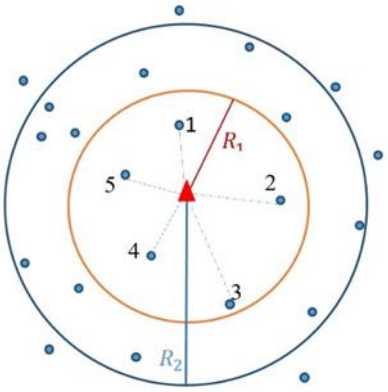
مدیریت و اخذ تصمیمات موثر در بسیاری از حوزه‌های علمی و اجرایی نیازمند سطوح پیوسته و دقیق از داده‌های مکان مرجع است. معمولا چنین داده‌هایی مستقیما قابل اخذ نیستند و از داده‌هایی که به صورت نقطه‌ای جمع‌آوری می‌گردند، تولید می‌شوند. علاوه بر هزینه قابل توجه جمع‌آوری داده‌های نقطه‌ای، این داده‌ها به ویژه در مناطق وسیع یا صعب‌العبور به صورت پراکنده و نامنظم برداشت می‌شوند و در بسیاری از موارد با تغییرات زیاد در مقادیر مواجهند. گمانه‌های تحقیقاتی مربوط به اندازه‌گیری عیار مواد معدنی از جمله داده‌های مکان مرجع نقطه‌ای هستند که دارای ویژگی‌های پراکندگی، بی‌نظمی و همچنین تغییرات قابل توجه در مقادیر (شکستگی) هستند. با وجود اینکه پراکندگی و شکستگی داده‌ها در کارایی روش‌های درونیابی تاثیرگذار است ولی از روش‌های متداول، برای درونیابی انواع داده‌های معدنی استفاده می‌گردد. در این مطالعه، از تلفیق دو روش کمترین مربعات متحرک و کمترین مربعات بازگشتی یک روش درونیابی دقیق، قابل اعتماد و انعطاف‌پذیر برای درونیابی مواد معدنی پیشنهاد شده است. در روش پیشنهادی برای کشف شکستگی از مقادیر خطاهای ظاهری محاسبات کمترین مربعات استفاده شده است. در این روش، برای نقاط مرکزی یک آستانه پیوستگی (عدم شکستگی) بر اساس انحراف معیار مقادیر مشاهداتی تعیین می‌گردد، تا نقاط دارای شکستگی کشف و از محاسبه ارزش مقدار مرکزی حذف شوند. علاوه بر این، با استفاده از ویژگی بازگشتی، اندازه دامنه تاثیر به صورت دینامیک تعیین می‌گردد. در این روش برای هر نقطه مرکزی یک شعاع دامنه تاثیر منحصر به فرد متناسب با مقادیر و ساختار نقاط پیرامونی نقطه مرکزی تعیین می‌شود. دینامیک بودن اندازه شعاع دامنه تاثیر این امکان را فراهم می‌نماید که مقادیر ضرایب چندجمله‌ای و به تبع آن مقدار مرکزی با دقت و قابلیت اعتماد مطلوب تعیین شوند. کارایی روش پیشنهادی از طریق اعمال آن بر روی داده‌های شبیه‌سازی شده و همچنین مقایسه نتایج آن با نتایج حاصل از اجرای روش‌های متداول درونیابی بر روی داده‌های معدنی واقعی مورد ارزیابی قرار گرفته است. RMSE روش پیشنهادی با استفاده از روش اعتبارسنجی متقابل ۱۰تایی برای عنصر کروم معادل ۲۸,۰۲۰ و برای آهن معادل ۱,۰۷۴ است. مقایسه این نتایج با روش‌های متداول درونیابی نشان می‌دهد در شرایط توزیع نامناسب و پراکنده داده‌ها، کارایی روش پیشنهادی در مورد هر دو دسته عناصر دارای غلظت و تغییرات قابل توجه و عناصر دارای غلظت پایین و سطوح تقریباً یکنواخت بالاتر است.

واژگان کلیدی: درونیابی، کمترین مربعات بازگشتی، کمترین مربعات متحرک، توابع پایه شعاعی، داده‌های معدنی

* نویسنده رابط

۱- مقدمه

حوزه‌های مختلف علوم و مهندسی نیازمند توصیف تغییرات متغیرها در یک دامنه مکانی معین می‌باشند. ارائه یک تابع تحلیلی پیوسته برای توصیف تغییرات یک متغیر در سطح دامنه، در حالی که داده‌های مشاهداتی معمولاً گسسته و محدود هستند، پیچیده و حتی غیرممکن است. روش‌های عددی درونیابی با استفاده از داده‌های گسسته، امکان تقریب و توصیف متغیر را به صورت پیوسته در سطح دامنه فراهم می‌سازند. از روش‌های شناخته‌شده درونیابی می‌توان به روش کمترین مربعات متحرک[۱] (MLS) [۱]، روش توابع پایه شعاعی[۲] (RBF) [۲]، روش IDW [۳] و Kriging [۴] اشاره نمود. در تمامی روش‌های مذکور گروهی از نقاط تحت عنوان دامنه تأثیر[۳] (و یا دامنه پشتیبانی[۴])، حول نقطه مرکزی[۵] انتخاب می‌شوند. این نقاط برای تقریب رویه‌ای خطی یا غیرخطی و محاسبه مقادیر مجهول در دامنه تاثیر استفاده می‌شوند. دراین روش‌ها با تغییر اندازه شعاع، می‌توان اندازه محدوده تأثیر را افزایش یا کاهش داد (شکل۱). به این ترتیب درجه آزادی و قابلیت اعتماد برآورد مجهولات به دلخواه تغییر خواهد یافت.



شکل ۱- تغییر شعاع دامنه تاثیر یک نقطه مرکزی از R_1 به R_2 و اثر آن بر تعداد نقاط واقع در محدوده تأثیر

شکل و اندازه دامنه تأثیر و اثر آن بر دقت برآورد مجهولات، موضوع مورد توجه تحقیقات دو دهه گذشته بوده است [۵] . تعداد و چگونگی پراکندگی نقاط داخل دامنه تأثیر از عوامل مؤثر بر دقت مجهولات است [۶, ۷]. Deng و Tang پیشنهاد نمودند که تعداد نقاط واقع در دامنه تأثیر بر اساس برآورد خطای درون‌یابی تعیین شود [۸] و Wang و Liu با ذکر چند مثال عددی نشان داد که تعداد ۶ تا ۱۷ نقطه به دقت کافی برای برآورد مجهولات می‌رسد [۹]. آنچه از بررسی مقالات موجود مشخص است، این است که در تمام تحقیقات انجام‌شده یا تعداد نقاط دامنه تأثیر یا شعاع آن در کل دامنه ثابت در نظر گرفته‌شده است [۱۰-۱۳]. در حالی که اندازه شعاع دامنه تاثیر باید با توجه به نحوه توزیع نقاط واقع در دامنه تأثیر و با در نظر گرفتن پیوستگی دامنه انتخاب شود [۱۴]. این نکته باعث می‌شود که اندازه شعاع دامنه تاثیر برای هر نقطه مرکزی به صورت منحصر به فرد و مجزا از سایر نقاط انتخاب شود. به عبارت دیگر اندازه شعاع دامنه تاثیر و شکل آن از یک نقطه مرکزی به نقطه دیگر متفاوت خواهد بود. تامین شروط فوق منوط به تعیین دینامیک شعاع دامنه تأثیر خواهد بود.

در این مطالعه، روش کمترین مربعات متحرک بازگشتی[۶] (RMLS) برای درونیابی داده‌های معدنی پیشنهاد شده است. این روش با تلفیق روش کمترین مربعات متحرک[۷] (MLS) و کمترین مربعات بازگشتی[۸] (RLS)، شعاع دامنه تاثیر را به صورت دینامیک و نقطه به نقطه تعیین می‌کند. کارایی روش پیشنهادی بر روی داده-های معدنی نمونه‌برداری شده با توزیع و غلظت‌های مختلف از گمانه‌های تحقیقاتی در محدوده وسیعی از شرق استان اصفهان مورد بررسی و ارزیابی قرار گرفته است.

در ادامه مقاله در بخش ۲ مواد و روش‌های مورد استفاده در این مطالعه معرفی شده است. بخش ۳ منطقه مورد مطالعه و داده‌های مورد استفاده در ارزیابی‌ها را توصیف می‌کند. نتایج حاصل از ارزیابی‌ها در بخش ۴ تشریح و مورد بحث قرار می‌گیرند. در نهایت بخش ۵ به بیان نتیجه‌گیری می‌پردازد.

۲- مواد و روش‌ها

در این بخش، ابتدا روش پایه کمترین مربعات متحرک معرفی می‌شود، سپس چگونگی تلفیق کمترین مربعات متحرک و کمترین مربعات بازگشتی برای درونیابی دقیق،

۱ Moving Least Squares
۲ Radial Basis Function
۳ Influence Domain
۴ Support Domain
۵ Central Point

۶ Recursive Moving Least Squares
۷ Moving Least Squares Method
۸ Recursive Least Squares

قابل اعتماد و انعطاف‌پذیر داده‌های معدنی تشریح می‌شود. در نهایت مروری بر چهار مدل متداول درونیابی که به منظور مقایسه و ارزیابی دقیق قابلیت‌های روش پیشنهادی مورد استفاده قرار می‌گیرند، ارائه می‌گردد.

۲-۱- کمترین مربعات متحرک (MLS)

روش کمترین مربعات متحرک (MLS) یکی از پرکاربردترین روش‌های بدون المان است. این روش برای اولین بار برای ایجاد و برازش یک رویه هموار به داده‌های پراکنده مورد استفاده قرار گرفت [۱۵] ولی دامنه کاربرد آن خیلی زود در علوم مختلف فراگیر شد [۱, ۱۶, ۱۷]. شکل کلی تقریب هر تابع اسکالر u با روش MLS به صورت رابطه (۱) بیان می‌شود. این تابع می‌تواند بیانگر عیار نمونه‌های ماده معدنی در چاهک‌های تحقیقاتی به صورت تابعی اسکالر از موقعیت گمانه باشد.

$$u(\boldsymbol{x}) = \sum_{j=1}^{m} \lambda_j(\boldsymbol{x})\, \phi_j(\boldsymbol{x}) = \Phi(\boldsymbol{x})\boldsymbol{\lambda}(\boldsymbol{x}) \qquad (۱)$$

در رابطه (۱)، $\Phi(x)$ توابع پایه چندجمله‌ای، m تعداد جملات چندجمله‌ای و $\lambda(x)$ ضرایب مجهول توابع پایه می‌باشند. بردار مجهولات $\lambda(x)$ با استفاده از روش کمترین مربعات وزن‌دار[۱] (WLS) بر اساس رابطه (۲) برآورد می‌شود.

$$\hat{\boldsymbol{\lambda}}(\boldsymbol{x}) = (\Phi^T W \Phi)^{-1} \Phi^T W u \qquad (۲)$$

در رابطه (۲)، W ماتریس وزن مربوط به نقاط تشکیل دهنده دامنه تاثیر می‌باشد. وزن مقادیر مولفه سوم نقاط می‌تواند مربوط به دقت اندازه‌گیری مولفه باشد. ولی متداول است که این وزن با فاصله هر نقطه از نقطه مرکزی نسبت معکوس داشته باشد.

۲-۲- کمترین مربعات متحرک بازگشتی (RMLS)

فرض پنهان در تمامی روش‌های درونیابی از جمله روش MLS این است که تغییرات مقادیر مشاهداتی در سطح دامنه تأثیر پیوسته است. این در حالی است که ممکن است در سطح دامنه تأثیر، شکستگی (ترانشه)، گسل، تغییر ناگهانی یا هر نوع ناپیوستگی دیگری وجود داشته باشد و پیوسته فرض نمودن آن باعث ایجاد خطا در برآورد مجهولات شود. با توجه به اینکه روش‌های درونیابی متداول ضرایب مجهول را روی کل دامنه تأثیر محاسبه می‌کنند، امکان در نظر گرفتن تغییرات ناگهانی و گسستگی‌های احتمالی را ندارند. در این مطالعه، برای مرتفع ساختن این مشکل از روش کمترین مربعات متحرک بازگشتی (RMLS) که حاصل تلفیق دو روش MLS و RLS است، برای درونیابی داده‌های معدنی استفاده می‌گردد. کمترین مربعات بازگشتی (RLS) یک الگوریتم تطبیقی است که برای کمینه کردن یک تابع هدف مربعی وزن‌دار از باقیمانده‌ها، مجهولات را به صورت بازگشتی برآورد می‌نماید [۱۸, ۱۹].

در روش RMLS به جای در نظر گرفتن یک شعاع تأثیر[۲] ثابت برای همه نقاط مرکزی یا وارد نمودن تعداد مشخصی از نقاط در محدوده تأثیر از روش کمترین مربعات بازگشتی برای برآورد مجهولات استفاده می‌شود. در این روش نقاط به ترتیب فاصله از نقطه مرکزی وارد فرآیند محاسبات می‌شوند، سپس با تعریف یک معیار برای تعیین آستانه ناپیوستگی، نقاطی که باعث رد شدن شرط می‌شوند از فرآیند محاسبات حذف می‌شوند.

مدل خطی معادلات مشاهدات مطابق رابطه (۳) تعریف می‌شود.

$$\boldsymbol{y} = A\boldsymbol{x} + \boldsymbol{\varepsilon} \qquad (۳)$$

در این رابطه، $\boldsymbol{y}$ بردار مشاهدات، $\boldsymbol{x}$ بردار مجهولات، $\boldsymbol{\varepsilon}$ بردار خطاهای تصادفی و $\boldsymbol{A}$ ماتریس ضرایب است. اگر در فرآیند محاسبات همه مشاهدات به‌صورت یکجا وارد معادلات شوند، مجهولات با فرض پایدار بودن ماتریس نرمال از رابطه (۴) برآورد می‌شوند.

$$\hat{\boldsymbol{x}} = (A^T W A)^{-1} A^T W \boldsymbol{y} \qquad (۴)$$

در این رابطه، W ماتریس وزن مشاهدات و $\hat{x}$ بردار مجهولات برآورد شده می‌باشد. حال اگر مشاهده جدید y_k به $(k-1)$ مشاهدات قبلی اضافه شود، می‌توان بردار مجهولات را بدون تکرار همه محاسبات با استفاده از روابط (۵) و (۶) به هنگام نمود.

---

[۱] Weighted Least Squares

[۲] Support Size

$$y_k = A_k x + \varepsilon_k \qquad (۵)$$

$$\hat{\boldsymbol{x}}_k = \hat{\boldsymbol{x}}_{k-1} + K_k(y_k - \boldsymbol{A}_k\hat{\boldsymbol{x}}_{k-1}) \quad k > 1 \qquad (۶)$$

در این روابط $\hat{x}_{k-1}$ و $\hat{x}_k$ مجهولات برآورد شده در مراحل $(k-1)$ و (k) هستند. در رابطه (۶) مقدار $\hat{x}_{(k-1)}$ به تمام مشاهدات قبلی وابسته است و بردار $A_k\,\hat{x}_{(k-1)}$ مقدار پیش‌بینی‌شده y_k است. بنابراین $v_k = y_k -$ $A_k\,\hat{x}_{(k-1)}$ باقیمانده پیش‌بینی‌شده است. مقدار باقیمانده پیش‌بینی‌شده اختلاف بین $\hat{x}_{(k)}$ و $\hat{x}_{(k-1)}$ است. بزرگ شدن ناگهانی بردار باقیمانده‌ها، ابزاری مناسب برای تشخیص گسل‌ها، ترانشه‌ها یا اساساً هر نوع شکستگی یا ناپیوستگی دیگر در دامنه تأثیر خواهد بود. ماتریس K_k (ماتریس بهره‌برداری[۱]) از رابطه (۷) به‌دست می‌آید:

$$K_k = \left(N_{(k-1.k-1)} + A_k^T W_{kk} A_k\right)^{-1} A_k^T W_{kk} \qquad (۷)$$

ماتریس نرمال $N_{(k-1.k-1)}$ در رابطه (۷) از رابطه (۸) محاسبه می‌شود.

$$N_{(k-1.k-1)} = \sum_{i=1}^{k-1} A_i^T W_{ii} A_i \qquad (۸)$$

ماتریس واریانس کواریانس بازگشتی از رابطه (۹) به دست می‌آید.

$$C_{\hat{x}_{(k)}} = (I - K_k A_k)\,C_{\hat{x}_{(k-1)}} \qquad (۹)$$

جذر عناصر قطر اصلی ماتریس $C_{\hat{x}_{(k)}}$ انحراف معیار مجهولات برآورد شده در مرحله k می‌باشند که به عنوان معیار دقت مجهولات در نظر گرفته می‌شوند.

استفاده از روش RMLS این امکان را فراهم می‌نماید که نقاط یکی پس از دیگری، به ترتیب فاصله‌ای که از نقطه مرکزی دارند، وارد فرآیند محاسبات شوند. با این روش تأثیر ورود هر نقطه مشاهداتی به صورت جداگانه روی پارامترهای دلخواه (به عنوان مثال ضرایب تابع چندجمله‌ای درونیابی) قابل بررسی است. از این ویژگی می‌توان برای تشخیص گسل، گسستگی و هر نوع تغییر ناگهانی دیگر در دامنه تاثیر استفاده نمود.

یکی از معیارهایی که می‌تواند در تشخیص تغییرات ناگهانی مورد استفاده قرار گیرد مقادیر باقیمانده‌های ظاهری در برآورد مجهولات می‌باشند. بدینصورت که اگر باقیمانده ظاهری مشاهده‌ای بزرگ باشد، آن مشاهده با انجام تست‌های آماری قابل تشخیص خواهد بود. این باقیمانده بزرگ می‌تواند ناشی از تغییر ناگهانی در مقادیر نمونه‌های مشاهداتی باشد. در روش RMLS تغییرات ناگهانی در سطح دامنه تأثیر (که مشابه نقش خطای بزرگ را در مشاهدات بازی می‌کند) توسط آزمون باقیمانده‌ها قابل کشف است. بنابراین در مناطقی که شکستگی در سطح دامنه تأثیر وجود دارد، می‌توان شعاع را به‌گونه‌ای در نظر گرفت که شکستگی بیرون از دامنه تأثیر قرار گیرد و یا آن نقاط را از محاسبه مقدار نقطه مرکزی کنار گذاشت. اگر آزمون باقیمانده‌های ظاهری نشان داد که تغییرات ناگهانی در سطح دامنه تأثیر وجود ندارد می‌توان نتیجه گرفت که نه مشاهدات آلوده به خطای بزرگ و اشتباه هستند و نه شکستگی در سطح دامنه تأثیر وجود دارد. در این وضعیت برای محدود کردن اندازه شعاع تأثیر می‌توان از معیارهای مختلف اندازه‌گیری دقت استفاده نمود. به عنوان مثال، نرم[۲] $\left(\left\|C_{\hat{x}_{(k)}}\right\|\right)$، اثر[۳] $\left(trace\left(C_{\hat{x}_{(k)}}\right)\right)$، بزرگترین مقدار ویژه (λ_{max})، طیف مقادیر ویژه (A) و دترمینان $(\det(C_x)$ ماتریس واریانس کواریانس مجهولات می‌توانند برای کنترل دقت نتایج و در نتیجه اندازه شعاع دامنه تاثیر استفاده شوند. تمامی معیارهای مذکور در حقیقت معیارهای اسکالر تعریف شده روی ماتریس واریانس کواریانس هستند. در این مطالعه از اثر $\left(trace\left(C_{\hat{x}_{(k)}}\right)\right)$، به عنوان معیار دقت استفاده می‌شود.

از دیگر موارد کاربرد روش RMLS امکان حذف تابع وزن در روش‌های بدون المان است. به ویژه زمانی که هدف تشخیص اشتباهات یا ناپیوستگی‌ها در مجهولات باشد، نیازی به استفاده از تابع وزن وجود ندارد. اعمال تابع وزن روشی برای تعین میزان تأثیر مؤلفه سوم نقاط واقع در دامنه تأثیر بر روی هر پارامتر مجهول و ابزاری برای محدود نمودن شعاع دامنه تأثیر بوده است. در حالی که مفهوم دامنه تأثیر با استفاده از روش کمترین مربعات بازگشتی دچار دگرگونی خواهد شد. در تمام تحقیقات قبلی اندازه دامنه تأثیر برای تمام نقاط، ثابت لحاظ شده است. معیار انتخاب اندازه این دامنه برای تمام نقاط

[۱] Gain Matrix

[۲] Norm
[۳] Trace

مرکزی یا شعاع ثابت و یا تعداد ثابت نقاط واقع در یک دامنه تأثیر بوده است و چگونگی پراکندگی نقاط در دامنه تأثیر به‌هیچ‌عنوان در نظر گرفته نشده است. درحالی‌که در روش RMLS، دامنه تأثیر به حالت دینامیک تعیین می‌شود و اندازه این دامنه از یک نقطه مرکزی به نقطه دیگر می‌تواند متفاوت باشد. در این روش، کوچک یا بزرگ بودن دامنه حایز اهمیت نیست، بلکه اطمینان از پیوستگی دامنه و رسیدن به دقت لازم برای پارامتر مجهول، مشخص‌کننده اندازه دامنه تأثیر است.

با توجه به معیار دقت مورد استفاده، اندازه دامنه تأثیر کاملاً دینامیک بوده و از نقطه‌ای به نقطه دیگر متفاوت است. لازم به ذکر است که شکل دامنه تاثیر در این روش کاملا به وضعیت پراکندگی، پیوستگی و مقادیر مشاهداتی نقاط وابسته است و لزوما یک شکل هندسی مشخص (مانند دایره یا مربع) نخواهد داشت. شکل ۲ روندنمای روش پیشنهادی را نمایش می‌دهد.

۲-۳- روش‌های درونیابی مورد استفاده برای مقایسه

هدف روش پیشنهادی، کنترل مناسب دقت، ضمن ایجاد درجه آزادی و به تبع آن قابلیت اعتماد در نتایج حاصل و امکان تشخیص عدم پیوستگی یا شکست در دامنه تاثیر است. برای تامین این هدف شعاع دامنه تاثیر به صورت دینامیک و نقطه به نقطه و با توجه به توزیع نقاط واقع در دامنه تعیین می‌شود. به منظور ارزیابی کارایی روش پیشنهادی، نتایج حاصل با نتایج چهار روش بدون المان دیگر که به صورت گسترده‌ای برای درونیابی داده‌های معدنی مورد استفاده واقع شده‌اند، مورد مقایسه قرار می‌گیرد. در تمامی روش‌های مذکور یا تعداد نقاط واقع در دامنه تأثیر یا شعاع آن در کل منطقه ثابت در نظر گرفته می‌شود. اولین روش مورد استفاده برای مقایسه، روش وزندهی براساس معکوس فاصله[۱] (IDW) است. در این روش، فرض بر این است که تاثیر هر پدیده متناسب با توانی از معکوس فاصله آن است، بنابراین تاثیر پدیده‌ی مورد نظر با افزایش فاصله، کاهش می‌یابد. در این روش از فاصله به عنوان وزن متغیر معلوم در پیش‌بینی نقاط اندازگیری نشده، استفاده می‌شود. همسایگی در این روش براساس تعداد همسایه‌ها و یا شعاع جستجو تغییر می‌یابد. روش IDW برای داده‌های منظم کارایی مطلوبی دارد ولی زمانی که داده‌ها به صورت نامنظم باشند، با مشکل مواجه می‌شود [۳]. دومین روش مورد استفاده توابع پایه شعاعی[۲] (RBF) است. این روش یک تابع پایه را به کار می‌گیرد که وابسته به فاصله بین نقاط درون‌یابی شده و نمونه‌برداری شده است. توابع پایه‌ای متعددی در این روش مورد استفاده قرار می‌گیرند. RBF برای تولید سطوح هموار از حجم بزرگی از نقاط نمونه مناسب است. با این حال، این روش برای استفاده در شرایطی که تغییرات بزرگی در فاصله کوتاه در نمونه‌ها ایجاد شود و یا زمانی که داده‌های نمونه برداری شده دارای عدم قطعیت باشند، نامناسب است [۲].

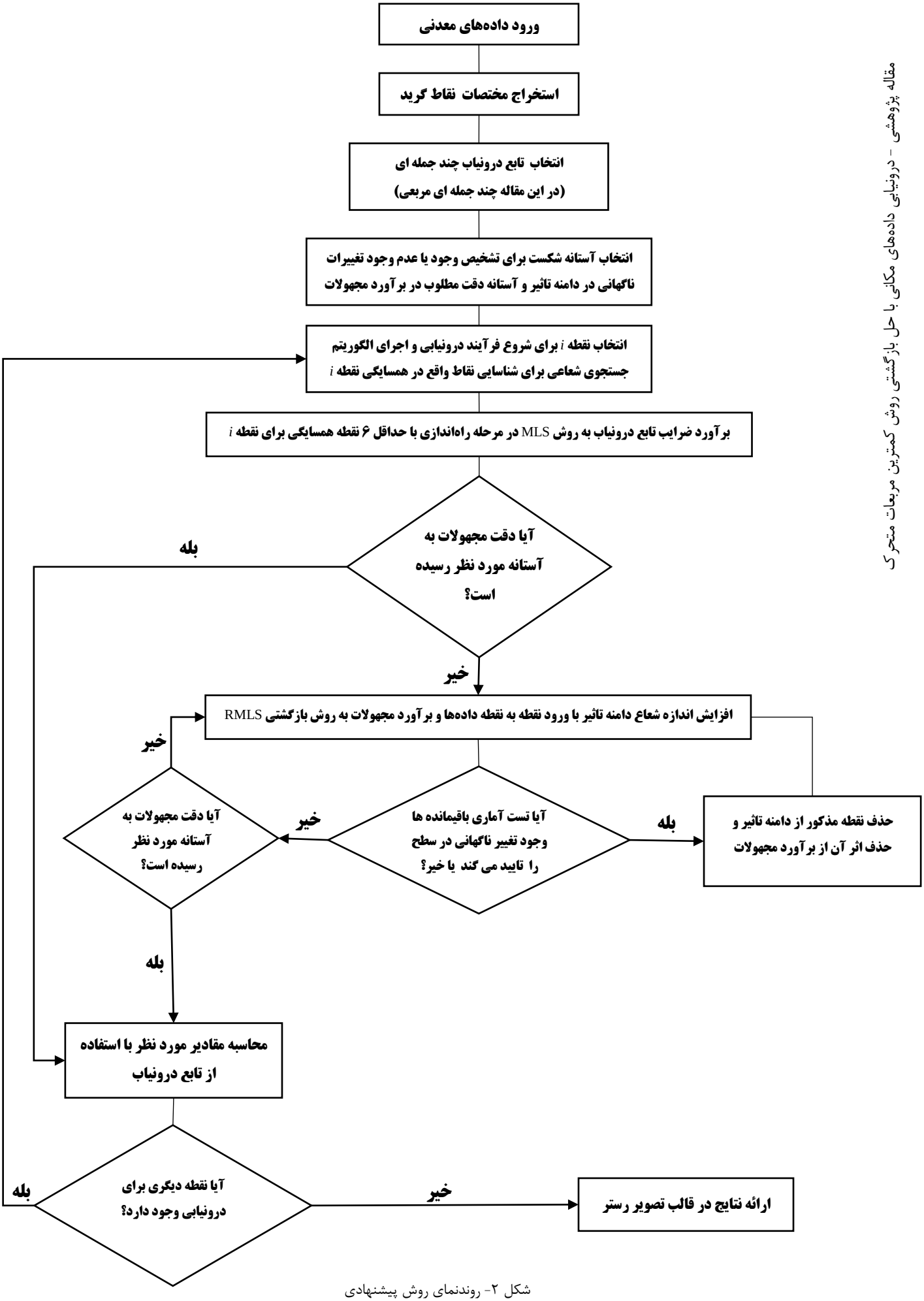
Kriging برخلاف IDW و RBF، یک روش زمین آمار[۳] مبتنی بر خودهمبستگی است. در این روش مقدار وزن هایی که به ارزش نقاط نمونه‌برداری شده ورودی اختصاص داده می شود نه تنها مبتنی بر فاصله بین هر نقطه نمونه‌برداری شده است بلکه علاوه بر آن به نحوه چیدمان و نظم کلی پراکنش مکانی نقاط نمونه برداری شده و همچنین مقدار ارزش‌های آنها بستگی دارد. Kriging ابتدا خودهمبستگی مکانی بین ارزش و موقعیت مکانی نقاط نمونه برداری شده را محاسبه می‌کند، سپس ارزش هر یک از نقاط مورد نظر را بر پایه این محاسبه، به دست می‌آورد [۴]. در این روش فرض بر اینست که ساختار تغییرات همگن بوده و اختلاف مقادیر یک متغیر در نقاط مختلف صرفاً تابعی از فاصله بین آنها است، در حالیکه در بسیاری از موارد اینگونه نیست.

چهارمین روش درونیابی، چند جمله‌ای محلی (LPI) است. در این روش، یک چند جمله‌ای از درجه مشخص (مثلاً مرتبه اول، دوم و سوم) با استفاده از نقاط موجود در همسایگی تعریف شده سطحی را برای همسایگی، برازش می‌کند. مقدار هر نقطه درونیابی شده، مقدار چند جمله‌ای برازش شده در مرکز همسایگی است.

[۱] Inverse distance weighting

[۲] Radial basis function

[۳] Geostatistic



شکل ۲- روندنمای روش پیشنهادی

روش LPI در درونیابی داده‌هایی با تغییر در فواصل کوتاه مورد استفاده قرار می‌گیرد. این روش برای زمانی که داده‌ها به صورت گرید برداشت شده باشند و توزیع داده‌ها در همسایگی نرمال باشد، مناسب است، البته اغلب داده‌ها فاقد چنین ویژگی‌هایی هستند. در روش RMLS با در نظر گرفتن حد آستانه ناپیوستگی و همچنین استفاده از تریس ماتریس واریانس کواریانس تلاش شده تا علاوه بر در نظر گرفتن ساختار و پراکندگی داده‌ها، نقاط ضعف روش‌های موجود، در درونیابی داده‌های دارای تغییرات ناگهانی مرتفع شود. در روش پیشنهادی می‌توان از چندجمله‌هایی با هر درجه دلخواه استفاده نمود. در این مقاله از چندجمله‌ای درجه دو استفاده شده است. با توجه به اینکه چندجمله‌ای درجه دو دارای شش ضریب مجهول می‌باشد (رابطه (۹))، برای برآورد این ضرایب حداقل مختصات ۶ نقطه در دامنه تأثیر مورد نیاز است. به همین دلیل در مرحله راه‌اندازی[۱] ۶ نقطه در همسایگی هر نقطه مرکزی وارد محاسبات می‌شوند. برای درونیابی با سایر روش‌های بدون المان تعداد نقاط معمولاً بین ۶ تا ۱۷ نقطه در نظر گرفته می‌شود [۹]. در این مطالعه، تعداد نقاط همسایگی برای روش‌های مورد استفاده برای مقایسه، ۱۲ نقطه در نظر گرفته شده است.

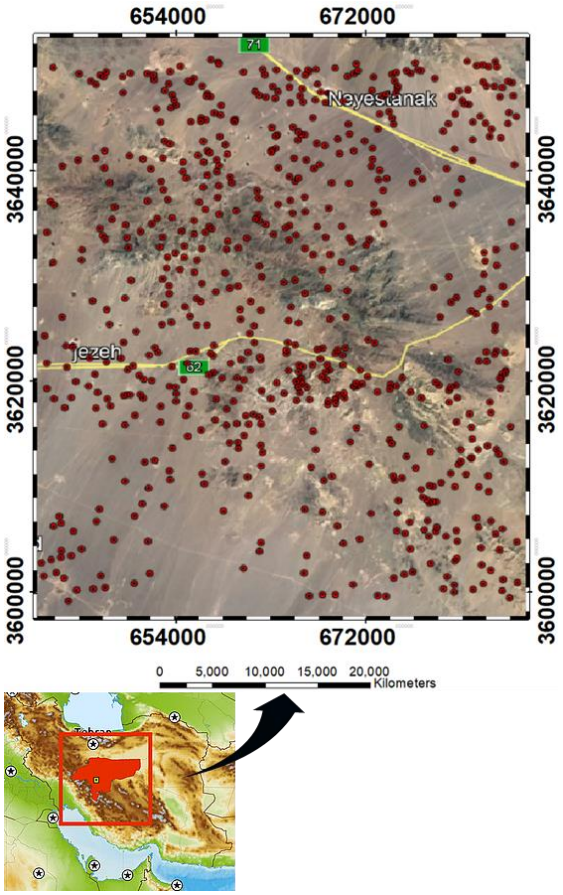
$$f(x.y) = a_0x^2 + a_1y^2 + a_2xy + a_3x + a_4y + a_5 \quad (۹)$$

در این رابطه $a_0.a_1.\ldots.a_5$ ضرایب مجهول و $f(x.y)$ مقادیر مشاهداتی می‌باشند.

۳- منطقه مورد مطالعه

به منظور بررسی وضعیت مواد معدنی در منطقه وسیعی از شرق استان اصفهان تعداد ۶۷۰ گمانه‌ی اکتشافی حفر شده است (شکل ۳).

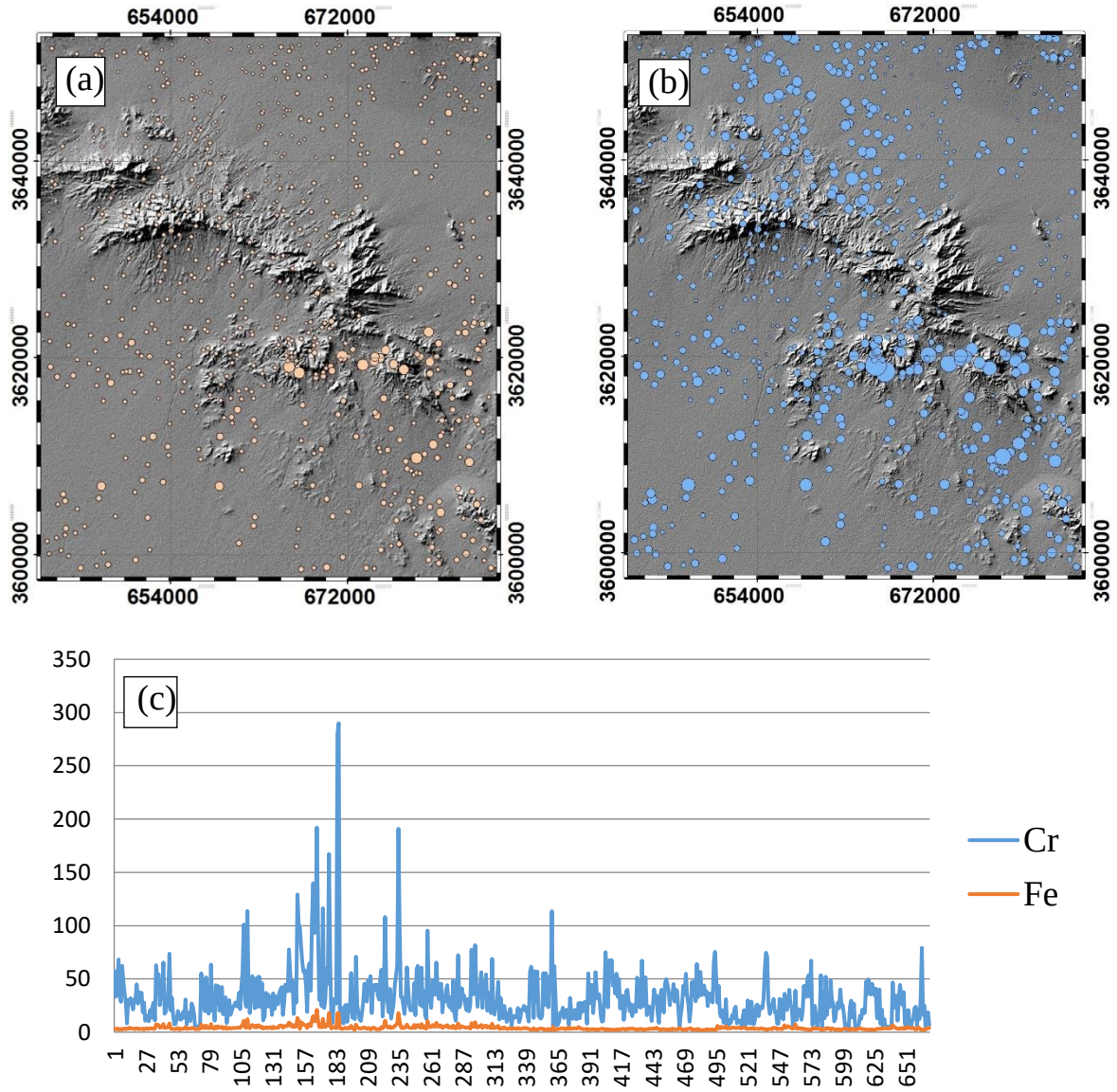
هدف از حفر این گمانه‌های اکتشافی، تعیین مرز دقیق ۱۴ نوع کانسنگ فلزی با سنگ‌های اطراف و تعیین عیار ماده معدنی در اعماق مختلف است. کانسنگ‌های فلزی مواد معدنی مقاوم و دارای رخنمون هستند. به همین دلیل، محل حفر چاه‌های گمانه براساس رخنمون و نقشه-های زمین‌شناسی بزرگ مقیاس، به صورت نامنظم تعیین می‌شود.



شکل ۳- پراکندگی ۶۷۰ گمانه اکتشافی در منطقه شرق استان اصفهان

در بخش‌هایی که تصور می‌شود تغییرات ماده معدنی زیاد است چاه‌های گمانه با تراکم بیشتر و در محل‌هایی که تغییرات اندک است فواصل چاه‌ها بیشتر در نظر گرفته شده است. به بیان دیگر، چاه‌های گمانه برای تعقیب لایه-های ماده‌های معدنی حفر شده‌اند. چاه‌های گمانه، در صورت اثبات عیار کافی از ماده معدنی، مقدمه حفاری‌های بیشتر هستند. با توجه به هزینه‌های بالای حفاری تعیین دقیق توزیع مواد معدنی از اهمیت بالایی برخوردار است. برای تعیین توزیع مواد معدنی در سطح منطقه مورد بررسی از روش‌های درونیابی استفاده می‌شود. به منظور بررسی کارایی روش پیشنهادی، دو کانی کروم و آهن انتخاب شده‌اند. کروم دارای غلظت و تغییرات قابل توجه و آهن دارای غلظت پایین و سطوح تقریباً یکنواخت غلظت در کل منطقه مورد بررسی می‌باشد (شکل ۴). دلیل انتخاب کانی‌های مذکور ایجاد امکان مقایسه روش پیشنهادی با سایر روش‌های متداول درونیابی در هر دو حالت توزیع غلظت نمونه‌ها است.

۱ Initial Stage



شکل ۴- پراکندگی و عیار ماده معدنی آهن (a)، پراکندگی و عیار ماده معدنی کروم (b)، مقایسه عیار کروم و آهن در چاهک‌های تحقیقاتی (c)

۴- نتایج و بحث

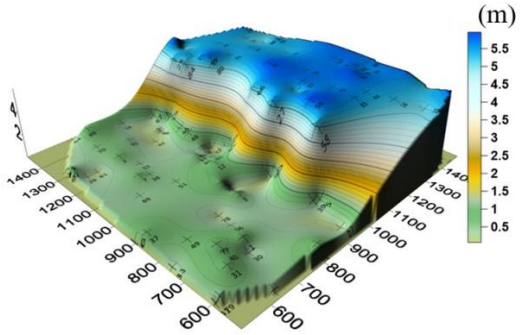
به منظور بررسی دقت و کارایی روش پیشنهادی، دو آزمایش بر روی مجموعه داده‌های شبیه‌سازی شده و واقعی انجام پذیرفته است.

در آزمایش اول، از داده‌های شبیه‌سازی شده برای ارزیابی قابلیت روش RMLS در تشخیص شکستگی در دامنه تاثیر و تعیین اندازه بهینه شعاع دامنه استفاده شده است. با توجه به اینکه وضعیت داده‌های شبیه‌سازی شده کاملا مشخص است، ابزار مناسبی برای بررسی ویژگی‌های روش پیشنهادی هستند.

در فرایند شبیه‌سازی یک شکستگی در مقادیر مشاهداتی (روی خط $x = 1000$) ایجاد شده است (شکل ۵). نقطه ۱۸ در موقعیتی قرار گرفته که این شکستگی (تغییر ناگهانی در مقادیر مشاهداتی) از داخل دامنه تاثیر آن می‌گذرد.

در روش پیشنهادی برای کشف شکستگی از مقادیر باقیمانده‌های ظاهری استفاده شده است. در محاسبات کمترین مربعات مقادیر باقیمانده‌های ظاهری بیانگر تغییرات مشاهدات می‌باشد. بنابراین اگر تغییر ناگهانی در سطح دامنه تأثیر اتفاق بیافتد مقدار باقیمانده ظاهری مربوط به آن مشاهدات نیز بزرگ خواهد شد. در نتیجه آزمون باقیمانده‌های ظاهری یک معیار مناسب برای

تشخیص شکستگی در سطح دامنه تأثیر خواهد بود. به این ترتیب، نقاطی که در سطح دامنه دچار شکستگی می‌شوند در محاسبات برازش رویه شرکت داده نمی‌شوند.



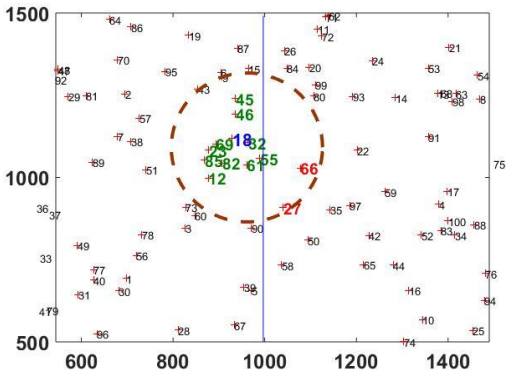
شکل ۵- وقوع تغییر ناگهانی در مقادیر مشاهداتی $x = 1000$

در این مقاله آستانه پیوستگی (عدم شکستگی) مقادیر مشاهداتی برابر ۵٪ انحراف معیار مشاهدات برای هر نقطه گرهی در نظر گرفته شده است. این بدین معناست که اندازه دینامیک دامنه تاثیر برای هر نقطه گرهی کاملاً متناسب با نحوه پراکندگی نقاط موجود در دامنه تاثیر و بر اساس انحراف معیار مقادیر داده (متناسب با واریوگرام مقادیر داده) تعیین می‌شود. مطالعات نویسندگان نشان می‌دهد که در نظر گرفتن همزمان نوع پراکنش مشاهدات و میزان تغییرات مقادیر داده (واریوگرام مقادیر داده) در سطح دامنه تاثیر، ویژگی منحصر به فرد روش RMLS می‌باشد. نتایج حاصل از اجرای الگوریتم RMLS بر روی داده‌های شبیه‌سازی شده در شکل ۶ نمایش داده شده است. در این شکل نقاط با $x < 1000$ (نقاط سبز) که دارای مقادیر مشابه با مقدار نقطه مرکزی هستند در دامنه تاثیر آن وارد شده‌اند ولی نقاط با $x > 1000$ (نقاط قرمز) از دامنه تاثیر حذف شده‌اند. به این ترتیب، روش RMLS قادر است تنها از مقادیر نقاطی برای برآورد نقطه مرکزی استفاده نماید که دارای تغییرات ناگهانی نسبت به آن نقطه نیستند.

روش پیشنهادی برای محدود کردن اندازه دامنه تاثیر از تریس[۱] $\left(trace\left(\boldsymbol{C}_{\hat{\boldsymbol{x}}_{(k)}}\right)\right)$ ماتریس واریانس کواریانس مجهولات استفاده می‌نماید (رابطه ۱۰). با این روش، شعاع دامنه تاثیر با در نظر گرفتن ساختار و مقادیر نقاط در اطراف نقطه مرکزی تغییر می‌کند. برای این منظور، آستانه همگرایی تریس ماتریس واریانس کواریانس مجهولات بر اساس آهنگ تغییرات این پارامتر محاسبه می‌شود و زمانی که این مقدار بیش از پنج درصد باشد عملیات تکرار و ورود نقاط جدید قطع می‌شود:

$$\frac{Trace(C_X(k))-Trace(C_X(k-1))}{Trace(C_X(k))} < 0.05 \qquad (۱۰)$$

در رابطه فوق k شمارنده ورود نقطه جدید در فرآیند بازگشتی است.



شکل ۶- دامنه تاثیر نقطه ۱۸ با در نظر گرفتن نقاط سبز رنگ و صرفنظر کردن از نقاط قرمز رنگ

در تمامی آزمایش‌ها از روش اعتبارسنجی متقابل ۱۰ تایی[۲] استفاده شده است. در این روش داده‌ها به ۱۰ زیرمجموعه تقسیم می‌شوند و از این ۱۰ زیرمجموعه هر بار ۹ زیرمجموعه برای آموزش[۳] و یک زیرمجموعه برای آزمایش[۴] مورد استفاده قرار می‌گیرد. به همین ترتیب این فرایند ۱۰ مرتبه تکرار می‌گردد و نتایج نهایی حاصل میانگین این ۱۰ مرتبه تکرار می‌باشند.

جدول (۱) نتایج حاصل از اجرای ۵ روش درونیابی را برای دو مجموعه داده کروم و آهن نمایش می‌دهد. در این جدول خطای جذر میانگین مربعات (RMSE)، میانگین و میانه خطاهای ظاهری برای هر مجموعه داده آورده شده است. RMSE به تنهایی معیار مناسبی برای نمایش کارایی روش‌های مختلف نیست زیرا وجود یک یا چند داده پرت در مشاهدات می‌تواند مقدار آن را به طرز چشمگیری تغییر دهد. میانگین و میانه خطاهای ظاهری شاخص‌های گرایش به مرکز هستند. این شاخص‌ها توزیعی از داده‌های اندازه‌گیری شده را خلاصه و توصیف

۱ Trace

۲ 10 fold Cross-validation
۳ Training
۴ Test

می‌کنند. به این ترتیب، هر چه این شاخص‌ها به صفر نزدیکتر باشند، نشان‌دهنده نزدیک به صفر بودن خطای هر روش و کارایی آن بر روی کل مجموع داده است.

براساس نتایج ارائه شده در جدول (۱) روش پیشنهادی بهترین عملکرد را در مقایسه با سایر روش‌ها داشته است. روش RMLS در مقایسه با سایر روش‌ها از یک سو شاخص RMSE را بهبود داده است، از سوی دیگر میانگین و میانه خطاهای ظاهری کوچکتری دارد. به عبارت دیگر داده‌های پرت تاثیر کمتری در دقت نهایی روش پیشنهادی در مقایسه با سایر روش‌های درونیابی دارند. تفاوت نتایج در مورد کانی کروم محسوس‌تر است. دلیل این تفاوت، مقادیر قابل توجه و تغییرات زیاد این عنصر در منطقه مطالعاتی است.

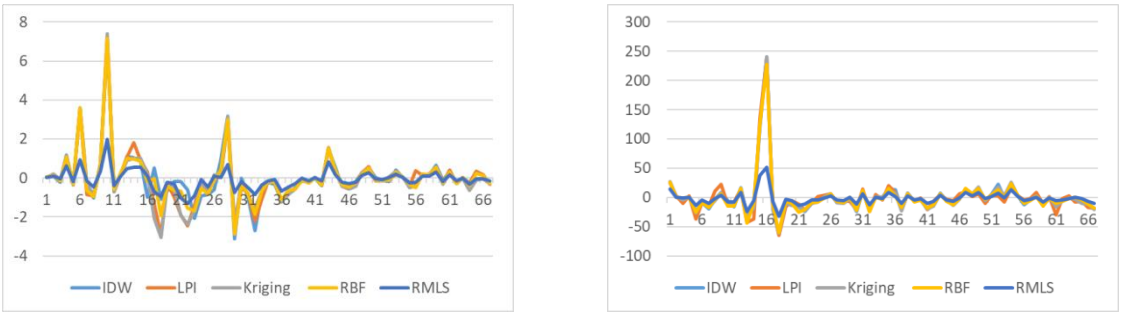
جدول ۱- میانگین RMSE مدل‌های مختلف برای مجموعه داده‌های کروم و آهن

مدل	مجموعه داده کروم (Cr)			مجموعه داده آهن (Fe)		
	RMSE	Mean[1]	Median[2]	RMSE	Mean	Median
LPI	۳۷,۳۳۹	۰,۴۱۸	-۳,۸۰۳	۱,۳۷۴	-۰,۰۵۴	-۰,۱۵۰
IDW	۳۶,۲۱۹	۰,۷۴۴	-۳,۹۱۵	۱,۳۰۱	**۰,۰۰۹**	-۰,۱۸۴
Kriging	۳۶,۰۷۴	۰,۴۵۷	-۴,۶۷۲	۱,۳۷۸	-۰,۰۷۳	-۰,۱۳۲
RBF	۳۵,۰۲۰	۰,۲۴۲	-۳,۶۸۷	۱,۲۸۳	-۰,۰۱۵	-۰,۱۸۰
RMLS	**۲۸,۰۲۰**	**۰,۲۰۱**	**-۲,۸۷۴**	**۱,۰۷۴**	-۰,۰۱۷	**-۰,۰۹۴**

[1] میانگین باقیمانده‌های ظاهری

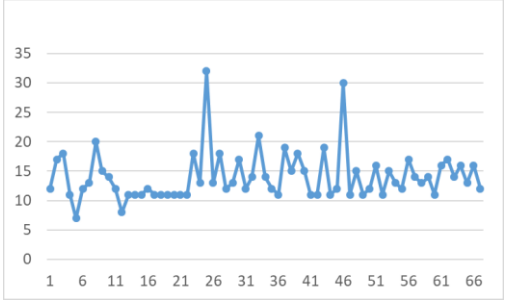
[2] میانه باقیمانده‌های ظاهری

روش RBF از میان سایر روش‌ها به صورت جزئی عملکرد بهتری داشته، با این وجود علیرغم تفاوت‌های ذاتی میان روش‌های مورد استفاده برای مقایسه، اختلاف میان نتایج آن‌ها در دو مجموعه داده چندان معنادار نیست.



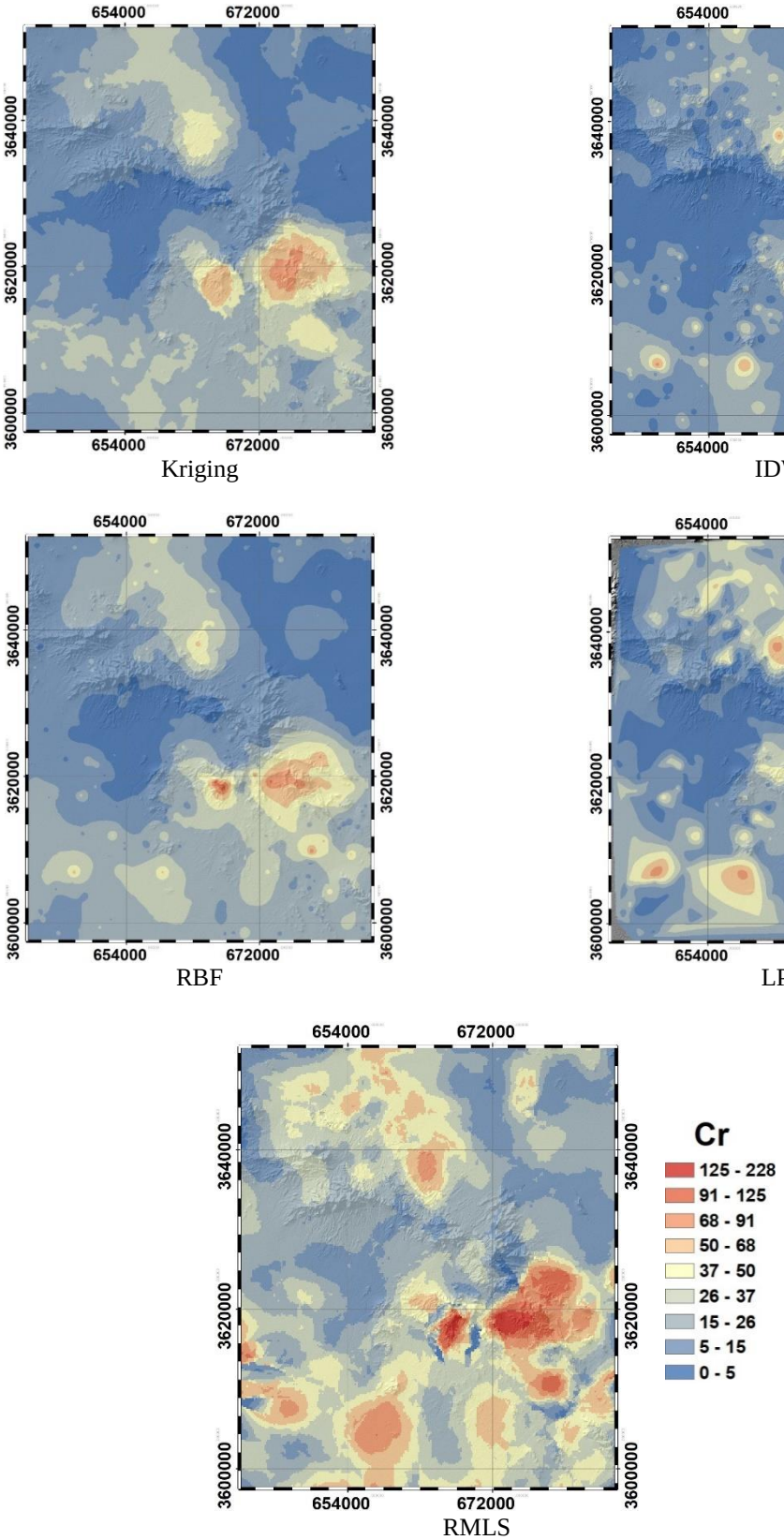
شکل ۷- خطای روش‌های درونیابی مختلف بر روی نقاط کنترل در یکی از آزمایش‌ها برای مجموعه داده کروم (سمت راست) و آهن (سمت چپ)

همان‌گونه که در شکل (۷) مشخص است مقادیر خطاهای روش پیشنهادی برای یکی از اجراهای آزمایشی کوچکتر از سایر روش‌های مورد مقایسه است. می‌توان نتیجه‌گیری کرد قابلیت‌های روش پیشنهادی در کشف پرش‌ها و شکستگی‌ها در دامنه تاثیر و حذف برخی نقاط دامنه از محاسبات و همچنین استفاده از یک دامنه تاثیر دینامیک منجر به افزایش کارایی در مقایسه با سایر روش‌ها شده است. شکل (۸) تعداد بهینه نقاط وارد شده در دامنه تاثیر هر نقطه مرکزی از مجموعه داده آزمایشی برای یکی از اجراها را نمایش می‌دهد. همان‌گونه که در شکل مشخص است تعداد نقاط حاصل از روش RMLS بین ۷ تا ۳۲ نقطه می‌باشد.



شکل ۸- تعداد نقاط موجود در دامنه تاثیر نقاط مجموعه داده آزمایشی کروم

این بدین معناست که اندازه دامنه تاثیر هر نقطه مرکزی با در نظر گرفتن شرایط نقاط پیرامونی آن نقطه و به صورت کاملاً دینامیک و مستقل از نقاط مرکزی دیگر تعیین می‌شود. نتایج درونیابی چهار روش فوق‌الذکر و روش RMLS برای ماده معدنی کروم در شکل (۹) ارائه شده است.



شکل ۹- نتایج درونیابی ماده معدنی کروم در منطقه مورد مطالعه برای روش‌های درونیابی مختلف

۵- نتیجه‌گیری

در این مقاله، روشی برای درونیابی داده‌های معدنی مربوط به چاهک‌های اکتشافی ارائه گردیده است. روش RMLS از تلفیق دو روش کمترین مربعات متحرک و کمترین مربعات بازگشتی برای ارتقا دقت، قابلیت اعتماد و انعطاف‌پذیری روش‌های درونیابی فعلی بهره برده است.

روش پیشنهادی دارای دو ویژگی منطبق بر شرایط خاص داده‌های معدنی است. شرایط مذکور شامل توزیع نامنظم و تغییرات زیاد مقادیر عیار مواد معدنی (شکستگی) است. ویژگی اول روش پیشنهادی، حل بازگشتی مسئله کمترین مربعات متحرک است. با استفاده از حل بازگشتی، اندازه دامنه تاثیر هر نقطه مرکزی به صورت دینامیک تعیین می‌گردد. به این ترتیب، برای هر نقطه مرکزی یک شعاع دامنه تاثیر منحصر به فرد با توجه به ساختار نقاط پیرامونی تعیین می‌شود. ساختار نقاط پیرامونی از طریق ماتریس واریانس کواریانس در محاسبات وارد می‌شود. به این ترتیب، در این روش برخلاف روش‌های متداول درونیابی به جای اینکه تابع وزن ناشی از فاصله نقش اصلی در کنترل نقاط وارد شده به فرایند درونیابی را داشته باشد، چگونگی توزیع نقاط و مقادیر عددی آن‌ها کنترل کننده نقاط وارد شده به محاسبات می‌باشد. برای انتخاب اندازه بهینه دامنه تأثیر نیز می‌توان از معیارهای مربوط به دقت، نظیر تریس واریانس کواریانس مجهولات استفاده نمود. در نظر گرفتن ساختار و پراکندگی نقاط در تعیین دامنه تاثیر، علاوه بر ارتقا دقت، توجیه فیزیکی بالاتری نسبت به روش‌های متداول دارد.

لازم به ذکر است با وجود دینامیک بودن دامنه تاثیر هر نقطه مرکزی، زمان و حجم محاسبات افزایش محسوسی نخواهد داشت. زیرا در برآورد مجهولات به روش پیشنهادی با افزایش شعاع دامنه تاثیر، فقط یک ترم تصحیحی محاسبه می‌شود و نیازی به تشکیل مجدد ماتریس ضرایب و ماتریس نرمال نیست. علاوه بر این، در این روش به دلیل بازگشتی بودن برآورد مجهولات نیازی به کامل بودن مرتبه[۱] ماتریس نرمال ضرایب برای همه نقاط واقع در دامنه تأثیر نیست، بلکه فقط در مرحله راه‌اندازی، نیازمند مرتبه کامل است. بنابراین، برخلاف روش کمترین مربعات متحرک معمولی که در آن توزیع نامناسب نقاط باعث بدوضع[۲] شدن حل مسئله برای برآورد مجهولات می‌شود، روش پیشنهادی نسبت به توزیع نامناسب نقاط در دامنه تاثیر حساسیت نخواهد داشت.

ویژگی دوم روش پیشنهادی قابلیت تشخیص شکستگی‌ها است. در روش RMLS نقاطی که در دامنه تأثیر نقطه مرکزی قرار می‌گیرند به ترتیب فاصله وارد فرایند محاسبات می‌شوند و در صورت ایجاد تغییر ناگهانی در مقادیر باقیمانده‌های ظاهری و به تبع آن در پارامترهای مجهول، قابل تشخیص خواهند بود. در نظر نگرفتن این نقاط در محاسبه مقدار نقطه مرکزی منجر به افزایش دقت محاسبات شده است.

کارایی روش پیشنهادی با استفاده از داده‌های شبیه‌سازی شده و واقعی مورد ارزیابی قرار گرفت. استفاده از داده‌های شبیه‌سازی شده قابلیت روش پیشنهادی در تشخیص شکستگی‌ها و تعیین دینامیک دامنه تاثیر مبتنی بر ساختار داده‌ها را مشخص می‌سازد.

به منظور بررسی تاثیر ویژگی‌های روش پیشنهادی، نتایج حاصل از آن با روش‌های متداول درونیابی شامل LPI، IDW، Kriging و RBF بر روی داده‌های آهن و کروم مقایسه گردید. نتایج عددی حاصل از این مقایسه نشان داد که روش پیشنهادی علاوه بر افزایش دقت محاسبات، حساسیت کمتری نسبت به داده‌های پرت دارد.

---

۱ Full rank

۲ Ill-posedness

مراجع

[1] Lancaster, P. and K. Salkauskas, Surfaces generated by moving least squares methods. Math. Comput., 1981. 37(141-158).

[2] Buhmann, M.D., Radial basis functions: theory and implementations. Vol. 12. 2003: Cambridge university press.

[3] Isaaks, E.H. and R.M. Srivastava, An introduction to applied geostatistics. 1989, Oxford university press.

[4] Krige, D.G.J.J.o.t.S.A.I.o.M. and Metallurgy, A statistical approach to some basic mine valuation problems on the Witwatersrand. 1951. 52(6): p. 119-139.

[5] Tsai, C.L., et al., Comparison of the performance of SSPH and MLS basis functions for two-dimensional linear elastostatics problems including quasistatic crack propagation. Computational Mechanics, 2013. 51(1): p. 19-34.

[6] Liu, G.R., Meshfree methods: moving beyond the finite element method. 2010, Boca Raton London New York: CRC press.

[7] Zhuang, X., C. Heaney, and C. Augarde, On error control in the element-free Galerkin method. Engineering Analysis with Boundary Elements, 2012. 36(3): p. 351-360.

[8] Deng, X. and Z.A. Tang, Moving Surface Spline Interpolation Based on Green's Function. Math. Geosci., 2011. 43: p. 663-680.

[9] Wang, J.G. and G.R. Liu, A point interpolation meshless method based on radial basis functions. International Journal for Numerical Methods in Engineering, 2002. 54(11): p. 1623-1648.

[10] Belytschko, T., et al., Element-free Galerkin methods for static and dynamic fracture. International Journal of Solids and Structures, 1995. 32(17): p. 2547-2570.

[11] Dolbow, J. and T. Belytschko, An introduction to programming the meshless Element F reeGalerkin method. Archives of Computational Methods in Engineering, 1998. 5(3): p. 207-241.

[12] Liu, G.-R., Meshfree methods: moving beyond the finite element method. 2009: Taylor & Francis.

[13] Li, J. and A.D. Heap, Spatial interpolation methods applied in the environmental sciences: A review. Environmental Modelling & Software, 2014. 53: p. 173-189.

[14] Mehrabi, H. and B. Voosoghi, Recursive moving least squares. Engineering Analysis with Boundary Elements, 2015. 58: p. 119-128.

[15] Shepard, D., A two-dimensional function for irregularly spaced data, in ACM National Conference. 1968. p. 517-524.

[16] Nayroles, B., G. Touzot, and P.J.C.m. Villon, Generalizing the finite element method: diffuse approximation and diffuse elements. 1992. 10(5): p. 307-318.

[17] Belytschko, T., et al., Fracture and crack growth by element free Galerkin methods. 1994. 2(3A): p. 519.

[18] Teunissen, P.J.G., Dynamic data processing recursive least squares. 2001, Netherland: Delft University press

[19] Simon, D., Optimal State Estimations. 2006, New Jersey: John Wiley & Sons, Inc. Hoboken.