

## ارزیابی مدل بیان ژن در پیش‌بینی مکانی شوری آب زیرزمینی و مقایسه آن با مدل‌های زمین‌آماري

(مطالعه موردی: دشت مشهد)

سعیده حسین‌آبادی<sup>۱</sup>، حسین خزیمه‌نژاد<sup>۲\*</sup> و عباس خاشعی‌سیوکی<sup>۳</sup>

۱. دانشجوی دکتری منابع آب، دانشکده کشاورزی، دانشگاه بیرجند

۲. دانشیار گروه علوم و مهندسی آب، دانشکده کشاورزی، دانشگاه بیرجند

۳. استاد گروه علوم و مهندسی آب، دانشکده کشاورزی، دانشگاه بیرجند

(تاریخ دریافت ۱۴۰۰/۰۳/۲۲، تاریخ تصویب ۱۴۰۰/۰۶/۱۲)

## چکیده

آب‌های زیرزمینی از منابع مهم بهره‌برداری در مناطق خشک و نیمه‌خشک هستند. به همین دلیل، به منظور حفظ کیفیت آب‌های زیرزمینی و مدیریت بهینه آن، اطلاع از پراکنش مکانی و زمانی آنها با اهمیت است و پایش و پهنه‌بندی آنها به عنوان اصل مهمی در برنامه‌ریزی‌های منابع آب کشور باید مد نظر قرار بگیرد. هدف از انجام پژوهش حاضر، پهنه‌بندی هدایت الکتریکی آب زیرزمینی آبخوان دشت مشهد با استفاده از ۵ روش عکس فاصله (IDW)، تخمین‌گر موضعی (GPI)، تخمین‌گر عام (LPI)، کریجینگ و کوکریجینگ و همچنین، ارزیابی مدل برنامه‌ریزی بیان ژن در پیش‌بینی این پارامتر با استفاده از داده‌های مکانی است. برای انجام پژوهش حاضر از داده‌های ۱۲۲ حلقه چاه مشاهده‌ای در محدوده آبخوان دشت مشهد استفاده شد. برای مقایسه روش‌های استفاده‌شده از سه معیار ارزیابی مجذور میانگین مربعات خطا (RMSE)، میانگین قدرمطلق خطا (MAE) و معیار نش-ساتکلیف (NSE) استفاده شد. ترسیم نیم‌تغییرنما در GS+ نشان داد داده‌های هدایت الکتریکی بهترین برازش را در مدل کروی دارند. نتایج پژوهش حاضر نشان داد از میان روش‌های یادشده، مدل برنامه‌ریزی بیان ژن با خطای  $RMSE=275/54 \mu\text{mos/cm}$  و  $MAE=319/73 \mu\text{mos/cm}$  و  $NSE=0/94$  و پس از آن، روش کوکریجینگ با خطای  $RMSE=573/59 \mu\text{mos/cm}$  و  $MAE=723/15$  و  $NSE=0/72$  بیشترین دقت و روش تخمین‌گر موضعی (GPI) با خطای  $RMSE=996/11 \mu\text{mos/cm}$  و  $MAE=755/56 \mu\text{mos/cm}$  و  $NSE=0/16$  کمترین دقت در این زمینه را داشتند.

کلیدواژگان: آب‌های زیرزمینی، برنامه‌ریزی بیان ژن، زمین‌آمار، هدایت الکتریکی.

## مقدمه

آب از نظر کمی فراوان‌ترین ماده کره زمین است و امروزه کیفیت منابع آبی در سراسر جهان به‌ویژه در کشورهای جهان سوم به دلیل فرایندهای طبیعی و انسانی در حال کاهش است و منابع آب شیرین به وسیله پدیده شوری تهدید می‌شوند. تخلیه فاضلاب‌های شهری، رواناب، دامداری‌های بزرگ، فاضلاب کارخانه‌ها، مکان دفع فاضلاب به عنوان عواملی هستند که بر کیفیت آب زیرزمینی تأثیرات بسیار منفی می‌گذارند [۱]. سفره‌های آب زیرزمینی نقش بسیار مهمی در تأمین آب در مصارف مختلف از جمله آب آشامیدنی، کشاورزی و صنعتی ایفا می‌کنند [۲]. افزایش بی‌رویه جمعیت، محدودیت منابع آب‌های سطحی و بهره‌برداری بیش از اندازه از آب‌های زیرزمینی باعث وارد آمدن خسارت‌های جبران‌ناپذیری به منابع طبیعی کشور در سال‌های گذشته شده است. عوامل یادشده نه تنها باعث کاهش کیفیت منابع آب زیرزمینی شده‌اند، بلکه کیفیت این منابع را نیز کاهش داده‌اند و در این زمینه، مدیریت بهره‌برداری و حفاظت از آب‌های زیرزمینی باید به عنوان یک اصل و پایه در برنامه‌ریزی‌های کشور قرار بگیرد.

مدیریت بهینه منابع آبی و حفظ و ارتقای کیفیت آنها نیازمند وجود اطلاعات در زمینه موقعیت، مقدار و پراکنش فاکتورهای شیمیایی آب در یک منطقه جغرافیایی معین است. در محیط GIS روش‌های مختلفی برای درون‌یابی و تهیه نقشه‌های پهنه‌بندی وجود دارد که در نگاه کلی می‌توان آنها را به دو دسته روش‌های قطعی شامل روش‌های فاصله معکوس وزنی، توابع پایه شعاعی و چندجمله‌ای‌های سراسری و محلی و روش‌های زمین‌آماری شامل کریجینگ تقسیم کرد [۳]. زمین‌آمار به عنوان یکی از شاخه‌های آمار عملی امکان بررسی هم‌زمان مقدار متغیر و آرایش مکانی و زمانی مشاهدات را در تحلیل داده‌ها فراهم می‌کند [۴]. با استفاده از تکنیک‌های زمین‌آمار می‌توان سطحی پیوسته از خصوصیات آماری نقاط معلوم را به‌وجود آورد. به بیانی، در زمین‌آمار نمونه‌ها، مستقل از یکدیگر نیستند و نمونه‌های مجاور تا فاصله معینی به طور مکانی به هم وابستگی دارند و می‌توان وابستگی بین نمونه‌ها را به صورت مدل ریاضی (تغییرنما) ارائه داد [۵]. در زمین‌آمار ابتدا به بررسی وجود داشتن یا نداشتن ساختار فضایی بین داده‌ها پرداخته می‌شود و سپس، در صورت وجود ساختار فضایی، تحلیل داده‌ها انجام می‌گیرد [۴]. تخمین زمین‌آمار شامل دو مرحله است. مرحله اول شناخت و مدل‌سازی ساختار

مکانی متغیر ناحیه‌ای است که بررسی آنالیز تغییرنما بررسی می‌شود و مرحله دوم، تخمین متغیر مد نظر به وسیله توابع زمین‌آمار از جمله کریجینگ است [۶].

با توجه به ناشناخته و پیچیده بودن تغییرات مکانی داده‌ها، از روش‌های مختلف درون‌یابی استفاده می‌شود و این روش‌ها با استفاده از فن اعتبارسنجی تقابلی مقایسه می‌شوند و از بین آنها روشی که کمترین خطا را در پیش‌بینی داشته باشد، انتخاب می‌شود. به منظور مدیریت بهینه منابع آبی در مناطق مختلف و انتخاب روش مناسب پهنه‌بندی و تهیه نقشه‌های تغییرات ویژگی‌های کیفی آب‌های زیرزمینی، تا کنون پژوهش‌های فراوانی در داخل و خارج کشور انجام شده است. نیکبخت و دلبری (۲۰۱۴) در پژوهشی با بررسی روش‌های مختلف درون‌یابی برای برآورد سطح آب زیرزمینی در دشت زاهدان بر اساس روش‌های زمین‌آمار، روش کوکریجینگ را به عنوان بهترین روش درون‌یابی معرفی کردند [۷]. شعبانی (۱۳۹۰) در پژوهشی که در دشت نیریز استان فارس انجام داد، به ارزیابی روش‌های زمین‌آمار در تهیه نقشه‌های کیفی و پهنه‌بندی آنها از نظر پارامترهای شوری و نیترات پرداخت. نتایج مطالعه ایشان نشان داد روش کریجینگ ساده نسبت به روش تابع شعاعی، برتری دارد و به‌عنوان روش مناسب برای تهیه نقشه تغییرات شوری و نیترات در دشت نیریز انتخاب شد [۸]. دلبری و همکاران (۱۳۹۲) در پژوهشی به پهنه‌بندی شاخص‌های کیفی (شوری و سدیمی) آب با استفاده از روش‌های زمین‌آمار در دشت کرمان پرداختند که در پژوهش خود از آمار مربوط به ۷۶ حلقه چاه استفاده کردند. نتایج پژوهش یادشده نشان می‌دهد هر دو روش کریجینگ معمولی و لاگ کریجینگ از دقت مشابهی برای تخمین شاخص‌های شوری و سدیمی برخوردارند [۹]. ولی‌اللهی (۲۰۱۹) در پژوهشی نشان داد روش کریجینگ نمای باعث پهنه‌بندی فضایی بهتر کیفیت آب زیرزمینی در حوضه دریاچه ارومیه در دوره مطالعه شده می‌شود [۱۰]. کیوسی و همکاران (۲۰۰۹) به بررسی و تحلیل تغییرات مکانی نیترات و شوری آب زیرزمینی حوضه عمان زرگا با استفاده از کریجینگ معمولی و کریجینگ شاخص پرداختند. نتایج پژوهش آنها نشان می‌دهد مقادیر نیترات در ۷۳ درصد از منطقه مطالعه شده بیشتر از ۵۰ میلی‌گرم در لیتر است [۱۱]. آرسلان (۲۰۱۲) در پژوهشی تغییرات مکانی و زمانی شوری آب زیرزمینی دشت بارفا واقع در شمال کشور ترکیه را بر اساس داده‌های ۹۷ حلقه چاه در یک دوره ۷ ساله با روش کریجینگ مطالعه کرد. نقشه‌های

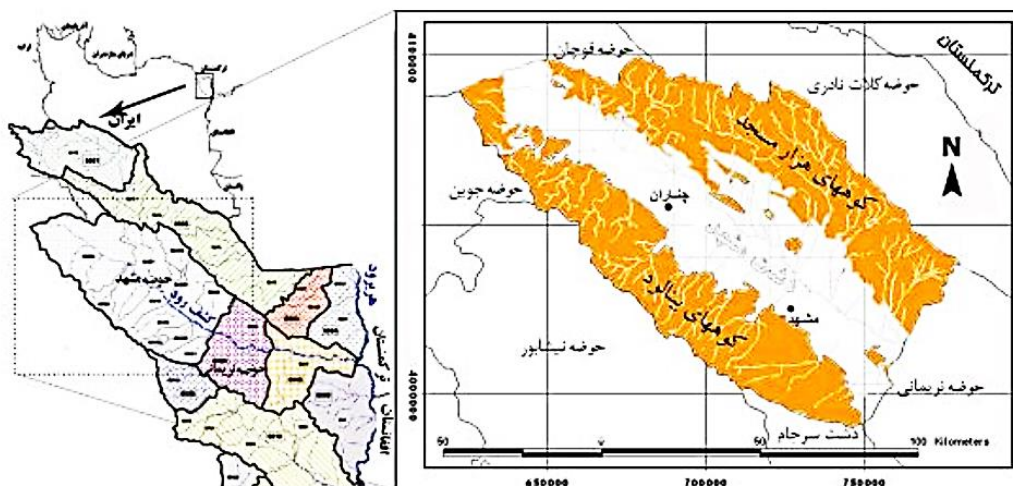
با توجه به اینکه در منطقه مطالعه‌شده تا کنون پژوهشی در خصوص استفاده از برنامه‌ریزی بیان ژن با استفاده از داده‌های جهانی (UTM)، در پیش‌بینی شوری صورت نگرفته است. بنابراین، هدف پژوهش حاضر، استفاده از این روش برای تخمین پارامتر شوری آب زیرزمینی و مقایسه با روش‌های مختلف زمین‌آماری است.

### مواد و روش‌ها

دشت مشهد حدود ۱۶۵۰۰ کیلومترمربع مساحت دارد که در شمال استان خراسان رضوی واقع شده است. این منطقه بین طول جغرافیایی ۵۸°، ۲۰' تا ۶۰°، ۸' و عرض جغرافیایی ۳۵°، ۴۰' تا ۳۶°، ۳' واقع شده است. در شکل ۱ موقعیت جغرافیایی دشت مشهد در کشور نشان داده شده است. دشت یادشده از شمال، با ارتفاعات هزارمسجد؛ از جنوب، با کوه بینالود؛ از شمال غربی، با حوضه رودخانه اترک و از جنوب شرقی، با حوضه رودخانه جمرود محدود می‌شود. آب‌وهوای منطقه به روش اصلاح‌شده دوماتن، خشک و سرد است. میانگین بارندگی آن ۶۲/۱ میلی‌متر و مقدار تبخیر- ترقق پتانسیل<sup>۱</sup>، ۳۴۸۳ میلی‌متر است.

تغییرات شوری نشان داد شوری آب زیرزمینی در سال ۲۰۰۴ با مقدار بیش از ۵ ds/m در ۳۱ درصد مساحت منطقه مطالعه‌شده، در سال ۲۰۱۰ به ۹ درصد مساحت منطقه کاهش یافته است [۱۲].

در دهه‌های اخیر، استفاده از روش‌های هوشمند گسترش فراوانی یافته است. برنامه‌ریزی بیان ژن (GEP)، یکی از روش‌های نوین هوش مصنوعی و تعمیم‌یافته الگوریتم ژنتیک است که بر اساس تئوری داروین ارائه شده و در سال ۲۰۰۰ توسط فریرا ابداع شد. این روش از تکامل زیست‌شناسی برای ایجاد یک برنامه کامپیوتری به منظور مدل‌سازی بعضی پدیده‌ها تقلید می‌کند [۱۳]. همچنین، نوعی تکنیک برنامه‌ریزی خودکار است که راه‌حل مسئله را با استفاده از برنامه کامپیوتری ارائه می‌کند. بررسی‌هایی که تا کنون در زمینه‌های مختلف به وسیله مدل‌های GEP صورت گرفته، نشان‌دهنده برتری و کارایی بهتر این مدل است. قربانی و صالحی (۱۳۹۰) در پژوهشی تغییرات سطح آب زیرزمینی و تأثیر آن بر برخی ویژگی‌های کیفی آب زیرزمینی را با استفاده از برنامه‌ریزی بیان ژن بررسی کردند. نتایج پژوهش آنها بیانگر همبستگی قابل قبول نوسان‌های سطح آب با کل املاح محلول و غلظت یون کلسیم است [۱۴].



شکل ۱. موقعیت جغرافیایی دشت مشهد

الکتریکی چاه‌های منطقه و کاربرد روش‌های مختلف میان‌یابی، به بررسی تغییرات مکانی شوری آب‌های زیرزمینی دشت مورد نظر پرداخته شده است. برای تبدیل داده‌های نقطه‌ای به نقشه سطح و تولید نقشه شوری آب زیرزمینی، از روش‌های

در پژوهش حاضر از داده‌های ۱۲۲ حلقه چاه در این دشت استفاده شده است. این داده‌ها از تاریخ یکم اردیبهشت‌ماه ۱۳۸۸ تا سی‌ویکم اردیبهشت‌ماه ۱۳۸۸ اندازه‌گیری شده‌اند. جدول ۱، نشان‌دهنده مشخصات آماری مجموعه داده‌های استفاده‌شده در پژوهش است. با استفاده از داده‌های هدایت

همچنین، برای مدل‌سازی و پیش‌بینی این پارامتر از روش برنامه‌ریزی بیان ژن نیز استفاده شده است.

زمین‌آمار کریجینگ، روش معین عکس فاصله، تخمین‌گر عام و تخمین‌گر موضعی در محیط نرم‌افزار GIS استفاده شده است.

جدول ۱. مشخصات آماری مجموعه داده‌ها

ردیف	پارامتر	حداکثر	حداقل	میانگین	انحراف معیار	ضریب تغییرات	چولگی	میانه
۱	X	۷۶۰,۷۵۰	۶۴۳,۳۰۰	۷۰۶,۷۳۹	۲۷۲۶۷/۷۸	۲۵/۹۱۸۴۷	-۰/۲۲۵۶۹	۷۰۶,۱۸۲/۵
۲	Y	۴,۰۹۵,۵۵۰	۳,۹۹۹,۸۲۲	۴,۰۴۰,۱۹۴	۲۴۴۶۶/۹۵	۱۶۵/۱۲۸۷	-۰/۴۸۶۲۳	۴,۰۳۴/۱۱۳
۳	CL	۳۸	۰/۲	۴/۹۰۳۳۰۶	۷/۳۵۵۲۹۲	۰/۶۶۶۳۶	۲/۴۰۶۳۲۳	۱/۳
۴	EC	۴,۷۶۰	۱۶۹	۱۴۲۱/۸۹۴	۱۳۳۳/۲۶۶	۱/۰۶۶۴۷۴	۱/۴۱۴۰۵۵	۸۵۱

فاصله معینی اضافه می‌شود و پس از آن، به حد ثابتی می‌رسد که حد آستانه نامیده می‌شود.

پس از جمع‌آوری داده‌ها در مرحله بعد پس از بررسی و تجزیه و تحلیل داده‌ها، به نرمال‌سازی داده‌ها پرداخته شد و با استفاده از رابطه ۱ تمامی داده‌ها در محدوده ۰ و ۱ نرمال شدند [۱۵].

### روش کریجینگ معمولی

یکی از روش‌های استفاده‌شده در پژوهش حاضر، کریجینگ معمولی است که بهترین تخمین‌گر خطی نارایب (B. L. U. E) نام گرفته است. روش یادشده نوعی روش تخمین زمین‌آمار است که با استفاده از مقادیر معلوم و یک نیم‌تغییرنما، مقادیر مجهول را برآورد می‌کند. در میان روش‌های مختلف، ویژگی روش کریجینگ این است که در عین نارایب بودن، واریانس تخمین نیز حداقل است.

$$X_{norm} = \left[ \frac{X - X_{min}}{X_{max} - X_{min}} \right] \times 0.8 + 0.1 \quad (1)$$

که در آن  $X_{norm}$ : مقدار نرمال‌شده متغیر،  $X$ : مقدار واقعی متغیر،  $X_{min}$ : حداقل مقدار متغیر و  $X_{max}$ : حداکثر مقدار متغیر است.

### تجزیه و تحلیل‌های زمین‌آمار

روش‌های گوناگونی برای پهنه‌بندی تغییرات غلظت، پیش‌بینی و پایش مؤلفه‌های شیمیایی آب‌های زیرزمینی وجود دارد که هر یک از آنها بسته به شرایط منطقه و وجود آمار و اطلاعات، دقت‌های مختلفی دارند.

### محاسبه تغییرنما

نیم‌تغییرنما واریانس وابسته به فاصله است که با نماد  $\gamma$  نشان داده می‌شود. اگر واریانس بین نقاطی به فاصله  $h$  کوچک باشد، نشانه وابستگی بیشتر بین نقاط است. نیم‌تغییرنما ( $\gamma$ ) به صورت رابطه ۲ محاسبه می‌شود [۱۵].

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(X_i + h) - Z(X_i)]^2 \quad (2)$$

### روش کوکریجینگ

در این روش می‌توان براساس همبستگی بین متغیرهای اصلی و کمکی به تخمین متغیر اصلی پرداخت. این خصوصیت می‌تواند باعث دقت بیشتر تخمین‌ها و صرفه‌جویی در هزینه‌ها (با نمونه‌برداری کمتر) شود. محل‌هایی که در آنها کمبود نمونه وجود دارد، به کمک متغیرهای ثانوی و با استفاده از همبستگی متقابل بین متغیرهای اصلی و ثانوی تخمین زده می‌شود. تخمین‌گر کوکریجینگ مطابق رابطه ۳ تعریف می‌شود:

$$Z(X_i) = \sum_{i=1}^n \lambda_{1i} Z_1(x_i) + \sum_{j=1}^m \lambda_{2j} Z_2(x_j) \quad (3)$$

در رابطه یادشده  $Z(X_i)$ : مقدار نامعلوم متغیر در نقطه  $x_i$ ،  $Z_1(x_i)$ : متغیر مکانی اصلی،  $Z_2(x_j)$ : متغیر کمکی یا ثانوی،  $m$  و  $n$  به ترتیب برابر با تعداد نقاط نمونه‌برداری متغیرهای اصلی و فرعی و  $\lambda_{1i}$  و  $\lambda_{2j}$  به ترتیب وزن‌های آماری اختصاص داده‌شده به متغیرهای اصلی و کمکی است.

که در آن:  $\gamma(h)$ : مقدار نیم‌تغییرنما در فاصله  $(h)$ ،  $Z(X_i + h)$ : مقدار مشاهده‌شده متغیر مورد نظر که به فاصله  $h$  از  $Z(X_i)$  قرار دارد،  $Z(X_i)$ : مقدار مشاهده‌شده متغیر مورد نظر،  $N(h)$ : تعداد جفت نمونه‌های به‌کاررفته در محاسبه که در فاصله  $h$  از یکدیگر قرار دارند، است. معمولاً هرچه  $h$  افزایش یابد، تعداد جفت‌ها کم می‌شود. با افزایش  $h$ ، مقدار نیم‌تغییرنما تا

## روش درون‌یابی IDW

روش درون‌یابی IDW یکی از روش‌های معمول و پرکاربرد درون‌یابی است که اساس آن بر مبنای این فرضیه است که در یک سطح درون‌یابی، اثر یک پارامتر بر نقاط اطراف یکسان نیست و نقاط نزدیک بیشتر و نقاط دور کمتر تحت تأثیر هستند و هرچه فاصله از مبدأ افزایش یابد، اثر پارامتر کمتر می‌شود (رابطه ۴).

$$Z_{X_0} = \frac{\sum_{i=1}^n z_{x_i} d_i^{-\alpha}}{\sum_{i=1}^n d_i^{-\alpha}} \quad (4)$$

$Z_{X_0}$ : مقدار تخمینی متغیر  $Z$  در نقطه  $X_0, X_1$ : مقدار نمونه در نقطه  $X_1, d_i$ : فاصله نقطه مد نظر تا نقطه مورد تخمین،  $\alpha$ : ضریبی که وزن بر اساس فاصله تعیین می‌کند و مقدار آن غالباً معلوم نیست و با سعی و خطا تعیین می‌شود.

## روش تخمین‌گر عام (LPI)

پیشنهاد می‌شود هنگامی که تغییرات از منطقه‌ای به منطقه دیگر آرام باشد مانند بررسی آلودگی در یک منطقه صنعتی و یا بررسی‌هایی در یک منطقه وسیع و جهانی که بررسی‌ها سطحی و وسیع هستند، از روش تخمین‌گر عام استفاده شود [۱۶].

## روش تخمین‌گر موضعی (GPI)

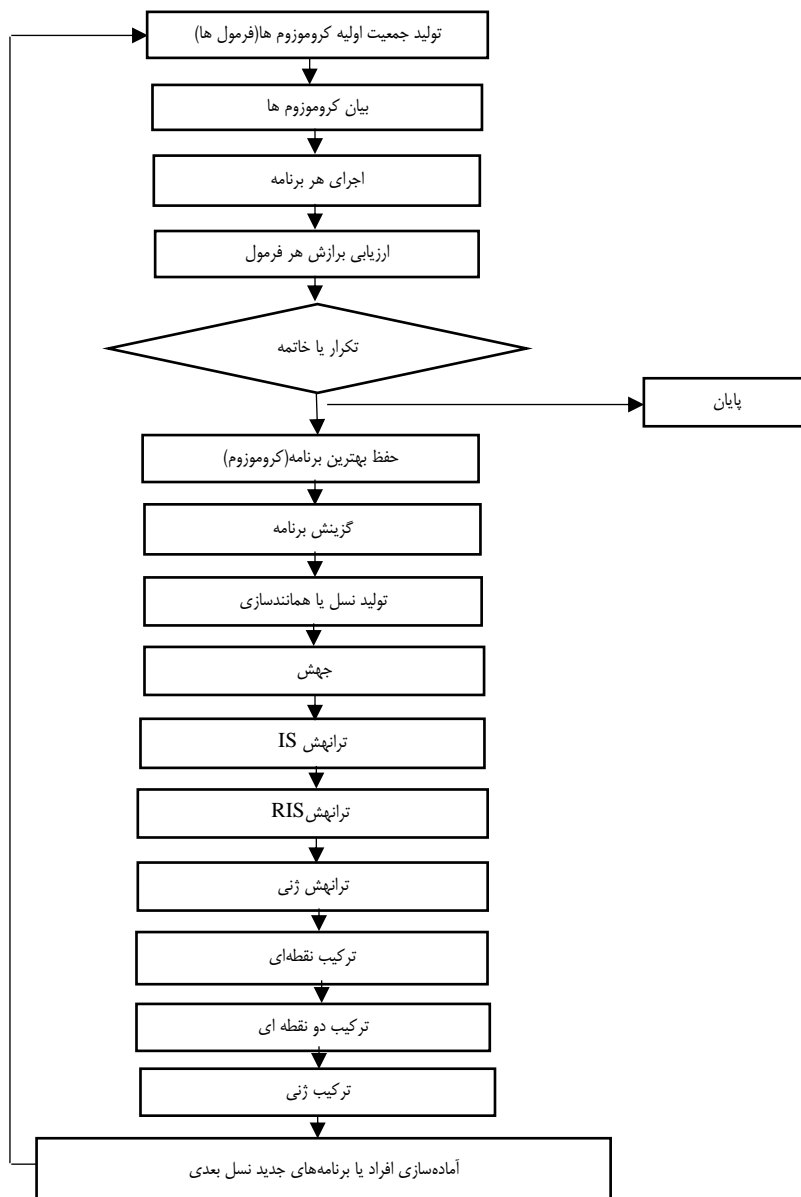
تخمین‌گر موضعی فرایندی است که یک فرمول (غالباً چندجمله‌ای) برای درون‌یابی پیدا می‌کند، روشی سریع با دقتی کمتر و ایجادکننده سطحی صاف است. این روش نسبت به روش کریجینگ خودکارتر عمل می‌کند و انعطاف‌پذیری کمتری در ایجاد نقشه‌ها دارد. پیشنهاد می‌شود از این روش در مناطق کوچک که تغییرات و یا نقاط نمونه در آن کم است، استفاده شود [۱۷].

## مدل برنامه‌ریزی بیان ژن

برنامه‌ریزی بیان ژن روشی است که از تکامل زیست‌شناسی برای ایجاد یک برنامه کامپیوتری به منظور مدل‌سازی بعضی پدیده‌ها، تقلید می‌کند [۱۳]. نوعی تکنیک برنامه‌ریزی خودکار است که راه‌حل مسئله را با استفاده از برنامه کامپیوتری ارائه می‌کند و قادر به بهینه‌سازی ساختار مدل و مؤلفه‌های آن است. مزیت این مدل نسبت به مدل‌های دیگر از جمله شبکه عصبی مصنوعی این است که در برنامه‌ریزی بیان ژن، ابتدا ساختار (متغیرهای ورودی، هدف و مجموعه توابع) تعریف شده و سپس، ساختار بهینه مدل و ضرایب طی فرایند آموزش تعیین می‌شوند. در برنامه‌ریزی بیان ژن، افراد به صورت رشته‌های خطی با طول ثابت کدگذاری شده و سپس، به شکل نهادهای غیرخطی با اندازه‌ها و شکل‌های متفاوت (یعنی نمایش دیاگرام ساده یا بیان درختی (۱) بیان می‌شوند [۱۸].

فرایند گام‌به‌گام حل یک مسئله با استفاده از برنامه‌ریزی بیان ژن شامل ۵ مرحله است. مرحله اول: انتخاب مجموعه ترمینال که همان متغیرهای مستقل مسئله و متغیرهای حالت سامانه است. انتخاب تابع برازش نیز در این مرحله صورت می‌گیرد؛ مرحله دوم: انتخاب مجموعه عملگرها و توابع ریاضی؛ مرحله سوم: شاخص اندازه‌گیری دقت مدل که بر مبنای آن می‌توان مشخص کرد که توانایی مدل در حل یک مسئله خاص تا چه اندازه است؛ مرحله چهارم: تعیین پارامترهای کنترل‌کننده اجرای برنامه و مرحله پنجم: معیار پایان و توقف برنامه که معیاری برای حصول نتیجه و توقف اجرای برنامه است (مثل تعداد تولید جمعیت جدید یا تعیین یک مقدار مشخص برای برازش فرمول‌ها). در شکل ۲ فلوچارت الگوریتم بیان ژن نشان داده شده است.

باید توجه کرد که همه پارامترهای یادشده با اعمال سعی و خطا تا دستیابی به حداکثر ضریب همبستگی مدل توسط کاربر انتخاب می‌شوند.



شکل ۲. فلوجارت الگوریتم بیان ژن

نقطه  $(x_i)$ ، مقدار واقعی متغیر مد نظر در نقطه،  $Z(x_i)$ : مقدار واقعی متغیر مد نظر در نقطه،  $\bar{Z}(x_i)$ : میانگین مقادیر واقعی متغیر و  $n$ : تعداد نقاط مشاهده‌ای.

در پژوهش حاضر، برای مدل‌سازی مقادیر شوری با استفاده از مدل برنامه‌ریزی بیان ژن از الگویی با سه ورودی  $X$ ،  $Y$  و پارامتر  $CL$  استفاده شد.

### نتایج و بحث

پس از نرمال‌سازی داده‌ها، اقدام به برازش واریوگرام مناسب به ساختار مکانی داده‌ها و تعیین مشخصات واریوگرام‌های رسم‌شده با استفاده از نرم‌افزار GS+ شد. با استفاده از داده‌های چاه‌های بررسی‌شده، پس از تحلیل سمیواریوگرام‌ها بهترین مدل برازش‌شده بر داده‌های کیفی مدل کروی است (شکل ۳) که این مدل استحکام ساختار فضایی قوی و همچنین، دامنه تأثیر بالایی برای تحلیل داده‌های کیفی دارد. در جدول ۲ مشخصات واریوگرام‌های برازش‌شده مشاهده به‌طور خلاصه آورده شده است که با توجه به آن مشاهده می‌شود که در تمامی مدل‌ها معیار  $C/C+CO$  دارای مقدار بیشتر از ۰/۵ است و در نتیجه، در تمامی مدل‌ها ساختار مکانی داده‌ها دارای وضعیت خوب است. مقایسه نتایج جدول ۲، نشان‌دهنده دقت بیشتر مدل کروی نسبت به سایر مدل‌ها است. کرمانی و نجف‌آبادی (۱۳۹۶)، یزدانی‌مقدم و همکاران (۱۳۹۱) نیز در پژوهشی مدل کروی را به عنوان مناسب‌ترین مدل برای استفاده و برازش واریوگرام معرفی کردند [۲۰ و ۲۱].

### معیارهای ارزیابی روش‌ها

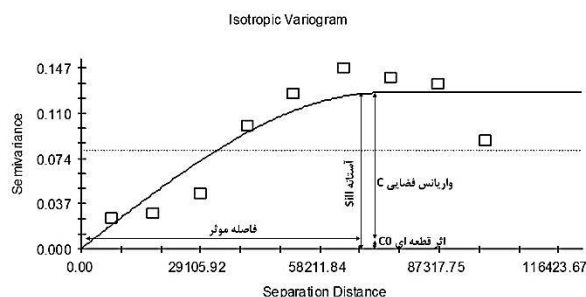
در پژوهش حاضر برای ارزیابی روش‌ها، از روش اعتبارسنجی متقابل استفاده شده است. این روش بر این اساس است که در هر مرحله یک نقطه مشاهده‌ای حذف شده و با استفاده از بقیه نقاط، آن نقطه برآورد می‌شود. سپس، مقدار واقعی به محل قبلی برگردانده شده و برای تمامی نقاط شبکه این عمل تکرار می‌شود. ارزیابی روش‌های به‌کاررفته با استفاده از معیارهای ارزیابی مجذور میانگین مربعات خطا (RMSE) و میانگین قدرمطلق خطا (MAE) و معیار نش-سانتکیف (NSE) بین نقاط تخمینی و واقعی انجام می‌شود که روابط ۵-۷ نشان‌دهنده شیوه محاسبه این معیارها است.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^n \{Z^*(x_i) - Z(x_i)\}^2} \quad (5)$$

$$MAE = \frac{\sum_{i=1}^n |Z^*(x_i) - Z(x_i)|}{n} \quad (6)$$

$$NSE = 1 - \frac{\sum_{i=1}^n (Z^*(x_i) - Z(x_i))^2}{\sum_{i=1}^n (Z(x_i) - \bar{Z}(x_i))^2} \quad (7)$$

که در روابط یادشده:  $Z^*$ : مقدار تخمینی متغیر مد نظر در



Spherical model ( $C_0 = 0.00010$ ;  $C_0 + C = 0.12720$ ;  $A_0 = 73200.00$ ;  $r_2 = 0.825$ ;  
RSS = 3.519E-03)

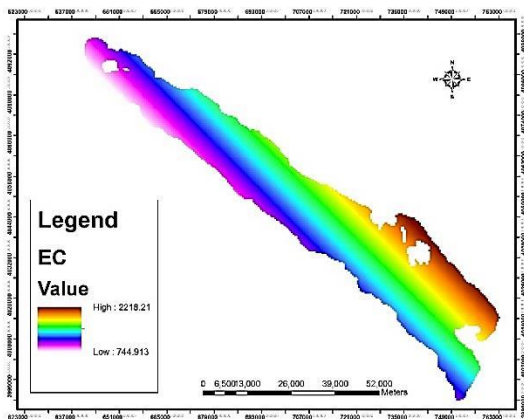
شکل ۳. واریوگرام برازش‌یافته برای پارامتر هدایت الکتریکی

جدول ۲. مشخصات واریوگرام‌های رسم‌شده با استفاده از نرم‌افزار GS+

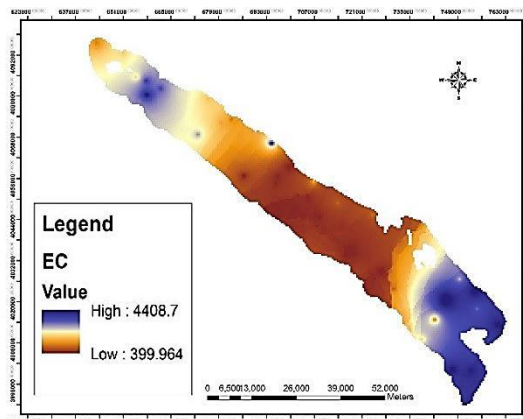
سمیواریوگرام				مدل
RSS	R2	C/C+CO	C0	
۰/۰۰۱۴	۰/۸۳	۰/۹۹۹	۰/۰۰۰۱	کروی
۰/۰۰۲۱	۰/۷۳	۰/۹۹۹	۰/۰۰۰۱	نمایی
۰/۰۰۱۳	۰/۸۳	۰/۹۴۷	۰/۰۰۴۴	گوسی
۰/۰۲۱	۰/۵۶	۰/۷۸۴	۰/۰۲	خطی

نشان‌دهنده پهنه‌بندی شوری آب زیرزمینی منطقه مطالعه‌شده با استفاده از ۵ روش مختلف درون‌یابی است. در ادامه در جدول ۳ نیز نتایج ارزیابی روش‌های استفاده‌شده بیان شده است.

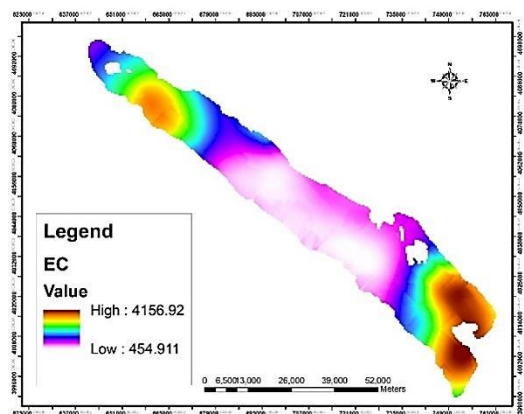
در روش درون‌یابی کوکریجینگ از پارامتر CI برای تخمین مقادیر هدایت الکتریکی استفاده شد. دلیل استفاده از متغیر کمکی این است که به این وسیله مقدار خطا کم می‌شود و تخمین مکانی بهتری به وجود خواهد آمد. شکل ۴



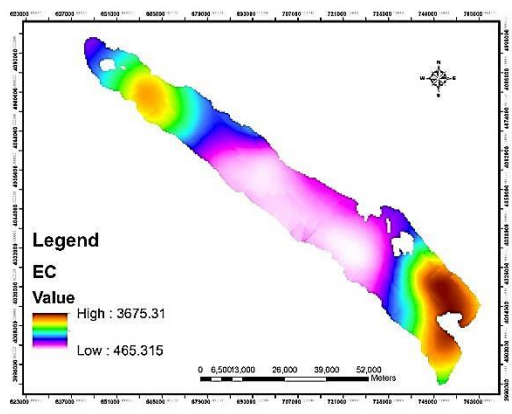
GPI (ب)



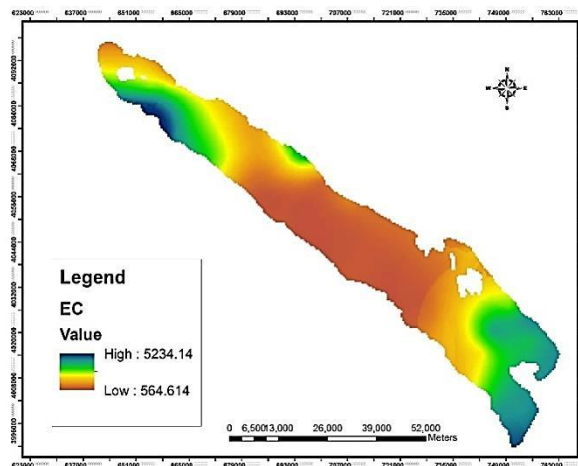
IDW (الف)



CoKriging (د)



Kriging (ج)



LPI (ه)

شکل ۴. پهنه‌بندی پارامتر هدایت الکتریکی ( $\mu\text{mos/cm}$ ) با استفاده از روش‌های درون‌یابی



جدول ۳. ارزیابی روش‌های درون‌یابی پارامتر هدایت الکتریکی

روش	RMSE ( $\mu\text{mos/cm}$ )	MAE ( $\mu\text{mos/cm}$ )	NSE
IDW	۸۷۶/۷۵	۵۱۹/۳۲	۰/۳۵
GPI	۹۹۶/۱۱	۷۵۵/۵۶	۰/۱۶
LPI	۶۲۳/۴۰	۴۲۴/۴۸	۰/۶۶
Kriging	۶۰۷/۹۸	۳۳۴/۴۹	۰/۶۹
CoKriging	۵۷۳/۵۹	۳۱۹/۷۳	۰/۷۲

با توجه به نتایج جدول یادشده مشاهده می‌شود که در مورد پارامتر هدایت الکتریکی در بین روش‌های IDW, GPI, LPI, Kriging و CoKriging، روش کوکریجینگ بر اساس مجموع معیارهای ارزیابی متقابل به‌عنوان مناسب‌ترین روش برای این پارامتر در منطقه مطالعاتی انتخاب می‌شود. کریجینگ پس از روش کوکریجینگ، دقت بیشتری نسبت به بقیه روش‌ها دارد. همچنین، نتایج نشان داد از میان روش‌های بررسی‌شده، روش GPI دارای کمترین دقت و بدترین نتایج بود.

نیکبخت و دلبری (۲۰۱۴) نیز در پژوهشی با بررسی روش‌های مختلف درون‌یابی برای برآورد سطح آب زیرزمینی در دشت زاهدان بر اساس روش‌های آماری روش کوکریجینگ را به‌عنوان بهترین روش درون‌یابی معرفی کردند [۷].

پس از انتخاب روش کوکریجینگ با توجه به شکل پهنه‌بندی آن برای پارامتر هدایت الکتریکی، ملاحظه می‌شود که حداقل و حداکثر مقدار هدایت الکتریکی محدوده آبخوان به‌ترتیب

#### نتایج به‌دست‌آمده از مدل برنامه‌ریزی بیان ژن

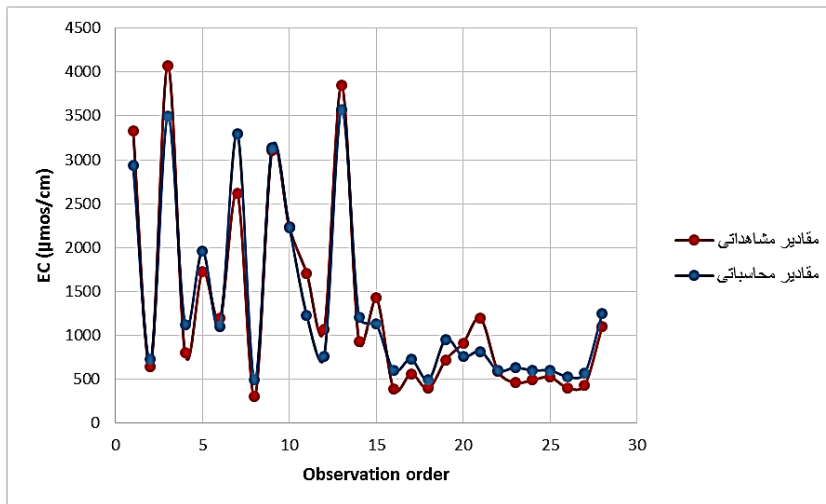
به منظور انتخاب ساختار بهینه مدل برنامه‌ریزی بیان ژن، پس از اعمال سعی و خطای فراوان، سرانجام مدل مناسب انتخاب شد که جدول ۴ نشان‌دهنده پارامترهای بهینه مدل پس از اعمال سعی و خطا است. همچنین، جدول ۵ نیز نشان‌دهنده نتایج ارزیابی این مدل در دو بخش آموزش و آزمون است. در شکل ۵ نمودار مقادیر پیش‌بینی‌شده توسط مدل و مقادیر مشاهداتی در بخش آزمون رسم شده است.

جدول ۴. پارامترهای مطلوب استفاده‌شده نرم‌افزار GenXproTools5.0

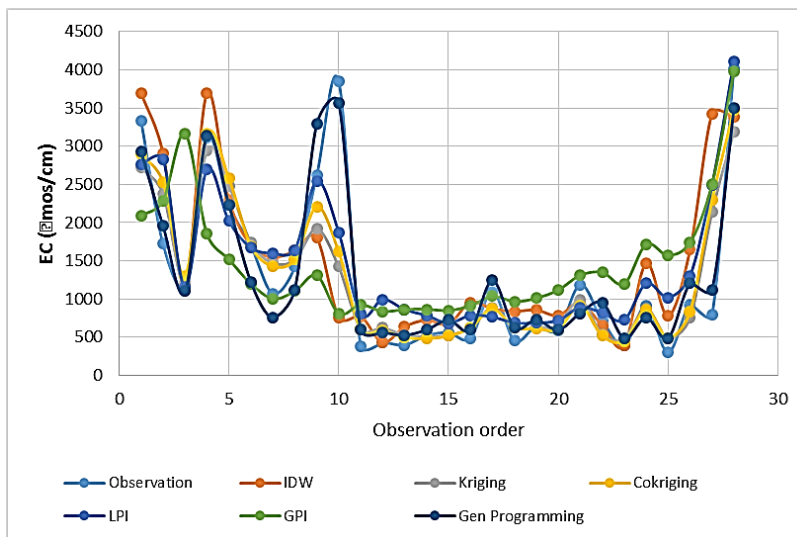
مقدار	پارامتر
۹	اندازه سر (Head Size)
۳۰	تعداد کروموزوم‌ها (Number of Chromosomes)
۵	تعداد ژن‌ها در هر کروموزوم (Number of Genes)
۰/۰۱	میزان جهش (Mutation Rate)
۰/۰۱	میزان وارونه‌سازی (Inversion Rate)
۰/۳	میزان ترکیب تک‌نقطه‌ای (One-Point Recombination Rate)
۰/۲	میزان ترکیب دو نقطه‌ای (Two-Point Recombination Rate)
۰/۲	میزان ترکیب ژن (Gene Recombination Rate)
۰/۰۰۵	میزان ترانشه درجه متوالی (IS Transposition Rate)
۰/۰۰۵	میزان ترانشه ریشه درجه متوالی (RIS Transposition Rate)
۰/۱	میزان ترانشه ژن (Gene Transposition Rate)
RMSE	معیار خطای تابع (Fitness Function Error Type)
تابع جمع "+"	تابع پیوند (Linking Function)
+, *, /, -	مجموعه توابع (Function Set)
۲	تعداد اجرا (Number of Runs)
۴۰۰۰	تعداد جمعیت یا نسل (Number of Generation)
۰/۰۰۲	میزان عملگر جهش در ثوابت تصادفی (R.C Mutation Rate)
۰/۰۰۲	میزان عملگر ترانشه در دامنه DC

جدول ۵. نتایج ارزیابی مدل مناسب برنامه‌ریزی بیان ژن

داده	RMSE	MAE	NSE
آموزش	۲۴۵/۷۲	۱۸۵/۳۸	۰/۹۶
آزمون	۲۷۵/۵۴	۲۲۳/۱۵	۰/۹۴



شکل ۵. نمودار مقادیر پارامتر هدایت الکتریکی (µmos/cm) با استفاده از روش برنامه‌ریزی بیان ژن



شکل ۶. نمودار مقایسه مقادیر پارامتر هدایت الکتریکی (µmos/cm) در روش‌های مختلف

به‌کار گرفته شده دارد. نتایج روش برنامه‌ریزی بیان ژن در پژوهش حاضر نشان داد در منطقه مطالعه شده می‌توان با استفاده از رابطه ۸ مقادیر شوری آب زیرزمینی را پیش‌بینی کرد.

$$\begin{aligned}
 EC = & ([\tan^{-1}(cl) - \tan^{-1}(y) + \cos(cl)] \times cl^2) \\
 & + (((\sin((cl)/-0.525) \\
 & + (\tan^{-1} y^2))^2) \times cl) \\
 & + (((\tan^{-1}((cl^2) \times cl)) \\
 & - \log((cl/y)))^2) \times cl) \\
 & + 449.12 + (\tan^{-1}(cl) + (8.55 \\
 & \times (-5.249/cl)))
 \end{aligned}
 \tag{۸}$$

با توجه به نتایج جدول‌های ۳ و ۵ ملاحظه می‌شود که مدل برنامه‌ریزی بیان ژن در بین تمام روش‌های بررسی شده، دارای بهترین دقت و مناسب‌ترین نتایج است و پس از آن، روش کوکریجینگ با دقت بیشتر و میزان خطای کمتر مدل‌سازی شوری آب زیرزمینی را در منطقه مطالعاتی انجام می‌دهد. همچنین، با توجه به شکل ۶ نیز مشاهده می‌شود که مقادیر به‌دست‌آمده از مدل برنامه‌ریزی بیان ژن بیشترین مطابقت را با مقادیر مشاهداتی دارند و پس از آن، مقادیر به‌دست‌آمده از روش کوکریجینگ بیشترین مطابقت را در بین روش‌های

Salm A. L., Roelofs J. G. M. High groundwater nitrate concentrations inhibit eutrophication of sulphate-rich freshwater wetlands. *Biogeochemistry*. 2004; (2): 249-267.

[2]. Ayotte J. D., Belaval M., Olson S. A., Burow K. R., Flanagan S. M., Hinkle S. R., Lindsey B. D. Factors affecting temporal variability of arsenic in groundwater used for drinking water supply in the United States. *Science of the Total Environment*. 2015; (505): 1370-1379.

[3]. Ghorbani K. Evaluation of the Empirical Bayesian Kriging method in ground water level zoning. *Journal of water and soil conservation*. 2018; 25(1): 165-182.

[4]. Goovaerts P. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *Journal of hydrology*. 2000; (1-2): 113-129.

[5]. Einax J. W., Soldt U. Geostatistical and multivariate statistical methods for the assessment of polluted soils—merits and limitations. *Chemometrics and Intelligent Laboratory Systems*. 1999; (1): 79-91.

[6]. Hassani Pak A. geostatistics. 3. Tehran: University of Tehran; 2010. [Persian]

[7]. Nikbakht S., Delbari M. Estimation of Groundwater Levels using Geostatistical Methods. *Journal of Water and Sustainable Development*. 2014; 1(1): 49-56.

[8]. Shabani M. Evaluation of Geostatistical Methods in Preparing Groundwater Quality Maps and Their Zoning (Case Study: Neyriz Plain, Fars Province). *Physical Geography Quarterly*. 2011; (13): 83-96. [Persian]

[9]. Delbari M., Afrasiab P., Salari M. Mapping Water Salinity and Sodicity Using Selected Geostatistical Methods, Case Study: Kerman Plain. *Journal of Water Resources Engineering*. 2013; (16): 11-24. [Persian]

[10]. Valiollahi J. Evaluating groundwater level and water-quality variation in Oshnaveh–Naqadeh Plain, Urmia Lake basin, northwestern Iran. *International Journal of Energy and Water Resources*. 2020; 4(1): 27-35.

[11]. Al Kuisi M., Al-Qinna M., Margane A., Aljazzar T. Spatial assessment of salinity and nitrate pollution in Amman Zarqa Basin: a case study. *Environmental Earth Sciences*. 2009; (1): 117-129.

[12]. Arslan H. Spatial and temporal mapping of groundwater salinity using ordinary kriging and indicator kriging: The case of Bafra Plain, Turkey. *Agricultural water management*. 2012; 57-63.

[13]. Roshangar K., Mirheidarian Sh. Use of evolutionary method of gene expression programming in estimating scour of bridge piers in non-adhesive substrates based on laboratory and field data. 8th National Congress of Civil Engineering of Babol. 2014. [Persian]

[14]. Ghorbani H., Salehi A. Using gene expression programming to study changes in groundwater quality data with water level fluctuations in Isfahan Borkhar

در رابطه یادشده  $y$  عرض جغرافیایی در واحد UTM و  $cl$  مقادیر پارامتر کلر است.

### نتیجه‌گیری

با استفاده از روش‌های درون‌یابی موجود در زمین‌آمار، می‌توان تحلیل‌های آماری و پیش‌بینی را در مکان‌های مختلف بر اساس موقعیت مکانی و جغرافیایی پدیده‌ها تحلیل کرد. با توجه به سیر تکاملی مدل‌های هوشمند و اهمیت آنها در مطالعات مربوط به آب، هدف از انجام پژوهش حاضر، انتخاب بهترین روش درون‌یابی به‌منظور بررسی و تحلیل مکانی تغییرات شوری آب زیرزمینی بود که جمع‌بندی نتایج به‌دست‌آمده از تجزیه و تحلیل نیم‌تغییرنما و ارزیابی روش‌های زمین‌آمار و برنامه‌ریزی بیان ژن در این پژوهش نشان داد که نیم‌تغییرنمای هدایت الکتریکی، همبستگی مکانی خوبی را در منطقه نشان می‌دهد. همچنین، نتایج مقایسه روش‌های مختلف برآورد هدایت الکتریکی نشان داد مدل برنامه‌ریزی بیان ژن با میزان خطای  $RMSE=275/54 \mu\text{mos/cm}$  و  $MAE=223/15 \mu\text{mos/cm}$  و  $NSE=0/94$  و پس از آن، روش کوکریجینگ با  $RMSE=573/59 \mu\text{mos/cm}$  و  $MAE=319/73 \mu\text{mos/cm}$  و  $NSE=0/72$  دارای مناسب‌ترین نتایج برای درون‌یابی و پیش‌بینی این پارامتر در دشت مشهد است و می‌توان با استفاده از رابطه پیشنهادی مدل برنامه‌ریزی بیان ژن، مقادیر هدایت الکتریکی را در منطقه مطالعاتی با میزان خطای کمتر نسبت به روش‌های دیگر بررسی‌شده تخمین زد. حاجی‌هاشمی و همکاران (۲۰۱۴) نیز در پژوهشی پس از بررسی روش‌های مختلف درون‌یابی برای برآورد مکانی کیفی آب‌های زیرزمینی در دشت گلپایگان، روش کوکریجینگ را به‌عنوان بهترین روش معرفی کردند [۲۲]. همچنین، نتایج پژوهش حاضر نشان داد روش تخمین‌گر موضعی (GPI) با خطای  $996/11 \mu\text{mos/cm}$  و  $RMSE=755/56 \mu\text{mos/cm}$  و  $MAE=755/56 \mu\text{mos/cm}$  و  $NSE=0/16$  از کمترین دقت در این زمینه بهره‌مند بوده است. نتایج پهنه‌بندی نیز نشان داد در قسمت مرکزی آبخوان دشت مشهد بهترین کیفیت از نظر هدایت الکتریکی ( $EC=459.9 \mu\text{mos/cm}$ ) و در قسمت جنوب شرق و شمال غرب وضعیت بحرانی‌تر ( $EC=4156.9 \mu\text{mos/cm}$ ) نسبت به سایر نقاط منطقه وجود دارد.

### منابع

[1]. Lucassen E. C. H. E. T., Smolders, A. J., van der

(Case study: Zarringol plain). Water and Soil Conservation. 2014; (1): 163-180. [Persian]

[19]. Nekooamal Kermani M, Mirabbasi R. Assessment of Interpolation Methods in Estimation of Groundwater Level (Case study: Sarkhon Plain). Hydrogeology. 2017; (2): 84-95. [Persian]

[20]. Yazdani Y, Vali A, Ghazavi R. Investigation of geostatistical methods in qualitative zoning of groundwater resources in Kashan plain. Geography and environmental planning. 2014; (3): 171-184. [Persian]

[21]. Hajihashemijazi M. R, Atashgahi M, Hamidian A. H. Spatial estimation of groundwater quality factors using geostatistical methods (case study: Golpayegan plain). Journal of Natural Environment. 2011; 63(4): 347-357.

plain. 6th Semnan National Civil Engineering Congress. 2011. [Persian]

[15]. Khashei-Siuki A, Sarbazi M. Evaluation of ANFIS, ANN, and geostatistical models to spatial distribution of groundwater quality (case study: Mashhad plain in Iran). Arabian Journal of Geosciences. 2015; 8(2): 903-912.

[16]. Yan X, Su X. Linear regression analysis: theory and computing. World Scientific. 2009.

[17]. Fan J, Gibels I. Local Polynomial Modelling and Its Applications, Chapman & Hall. London. Water Resources Bulletin. 1996; (87): 998-1004.

[18]. Abareshi F, Meftah Halghi M, Sanikhani H, Dehghani A.A. Comparison of three intelligence techniques for predicting water table depth fluctuations