

## تعیین پتانسیل آب زیرزمینی با استفاده از روش‌های داده کاوی و آماری در منطقه یاسوج-سی سخت

- ❖ محمد تقی آوند؛ دانشجوی دکتری علوم و مهندسی آبخیزداری، دانشکده منابع طبیعی و علوم دریایی، دانشگاه تربیت مدرس، نور.
- ❖ سعید جانی زاده؛ دانشجوی دکتری علوم و مهندسی آبخیزداری، دانشکده منابع طبیعی و علوم دریایی، دانشگاه تربیت مدرس، نور.
- ❖ محسن فرزین\*؛ استادیار آبخیزداری، دانشکده کشاورزی و منابع طبیعی، دانشگاه یاسوج.

### چکیده

با افزایش جمعیت و توسعه کشاورزی نیاز به منابع آبی به شدت افزایش یافته و منابع آب زیرزمینی، بیش از پیش، به خصوص در مناطق خشک و نیمه خشک مورد توجه بسیاری قرار گرفته است. هدف از این پژوهش تهیه نقشه پتانسیل منابع آب زیرزمینی با استفاده از دو مدل داده کاوی جنگل تصادفی (RF) و آماری رگرسیون خطی تعمیم یافته (GLM) در محدوده یاسوج-سی سخت می باشد. بدین منظور لایه‌های اطلاعاتی شامل درجه شیب، جهت شیب، طول شیب، ارتفاع از سطح دریا، شاخص رطوبت توپوگرافی، فاصله از گسل، فاصله از آبراهه، بارندگی، کاربری اراضی، سنگ شناسی، شاخص موقعیت توپوگرافی و شاخص قدرت جریان به عنوان مهم‌ترین عوامل مؤثر بر پتانسیل آب زیرزمینی تعیین شده و در نرم‌افزار ArcGIS و SAGAGIS رقومی و تهیه شدند. از پراکنش ۳۶۲ چشمه موجود در سطح منطقه، ۷۰ درصد (۲۵۳ چشمه) به عنوان چشمه‌های آموزشی و ۳۰ درصد (۱۰۹ چشمه) به عنوان چشمه‌های آزمایشی استفاده گردید. نتایج نشان داد که سطح طبقات حضور آب زیرزمینی با پتانسیل کم، متوسط، زیاد و خیلی زیاد در نقشه حاصل از روش جنگل تصادفی به ترتیب ۳۷/۷۸، ۲۲/۲۲، ۱۸/۸۹ و ۲۱/۱۱ درصد و در روش رگرسیون خطی تعمیم یافته به ترتیب ۱۴/۴۹، ۳۲/۰۴ و ۳۱/۱۱ و ۲۲/۳۶ درصد می‌باشد. همچنین با حساسیت‌سنجی عوامل مؤثر در هر دو روش، عامل‌های بارندگی، ارتفاع از سطح دریا و فاصله از گسل حساس‌ترین عوامل تعیین شدند. ارزیابی دقت مدل‌های داده کاوی مورد استفاده در این تحقیق نیز با استفاده از منحنی عملکرد نسبی (ROC) مورد سنجش قرار گرفت. سطح زیر منحنی (AUC) برای دو مدل RF و GLM به ترتیب ۹۲٪ و ۶۵٪ درصد را نشان می‌دهد، بنابراین دقت مدل جنگل تصادفی در تهیه نقشه پتانسیل آب زیرزمینی در منطقه مورد مطالعه بیشتر از مدل رگرسیون خطی تعمیم یافته است. مدل‌های نوین داده کاوی و آماری در تلفیق با GIS برای پتانسیل‌یابی منابع آب زیرزمینی می‌تواند برای مدیریت پایدار، مورد توجه طراحان و تصمیم‌گیران طرح‌های توسعه‌ای واقع گردد.

کلید واژگان: پتانسیل‌یابی منابع آب زیرزمینی، داده کاوی، جنگل تصادفی، رگرسیون خطی تعمیم یافته، یاسوج-سی سخت

## ۱. مقدمه

همچنین به عقیده بسیاری از محققان، امروزه با استفاده از روش‌های مختلف مبتنی بر سیستم اطلاعات جغرافیایی، پتانسیل‌یابی منابع آب زیرزمینی امکان‌پذیر بوده و نتایج نیز حاکی از آن است که پتانسیل‌یابی با این روش‌ها دقت قابل قبولی داشته است [۲۵، ۲۷ و ۳۲].

نقشه‌های پتانسیل منابع آب زیرزمینی، هزینه‌های حفاری چاه‌ها را کاهش داده و ابزار مفیدی برای طراحان و مهندسان جهت مکانیابی مناطق مستعد حضور آب زیرزمینی می‌باشد. همچنین این امکان را فراهم می‌سازند که با کمترین هزینه بتوان نواحی محتمل حضور چشمه را شناسایی کرده و در مواقع خشکسالی از آن بهره‌برداری نمود [۹]. تهیه نقشه پتانسیل حضور چشمه آب زیرزمینی برای موفقیت در طرح‌های شهری مفید می‌باشد [۳۴]. بر اساس پیشینه پژوهش، روش‌های مختلفی برای ارزیابی پتانسیل چشمه آب زیرزمینی و پهنه‌بندی آن ارائه شده است. این روش‌ها عبارتند از: نسبت فراوانی<sup>۱</sup> [۱۵]، شواهد وزنی<sup>۲</sup> [۳۹]، فاکتور اطمینان<sup>۳</sup> [۱۷] و رگرسیون لجستیک<sup>۴</sup> [۶] و اخیراً روش‌های یادگیری ماشین به دلیل شناسایی الگوهای پنهان داده‌های آب زیرزمینی و به دست آوردن رابطه غیرخطی بین متغیرها، توجه محققان را به خود جلب کرده است [۴، ۴۵].

تکنیک‌های مختلف داده کاوی از قبیل جنگل تصادفی (RF<sup>۵</sup>) [۳۷]، درخت رگرسیون تقویت شده (BRT<sup>۶</sup>) [۱۷]، حداکثر آنتروپی (ME<sup>۷</sup>) [۳۷]، شبکه عصبی مصنوعی (ANN<sup>۸</sup>) [۷]. در بسیاری از مطالعات پتانسیل‌یابی آب‌های زیرزمینی از داده‌های سنجش از دور و تلفیق آن با سیستم اطلاعات جغرافیایی استفاده شده است [۱۹ و ۴۳]. در همین راستا محققان بسیاری به تهیه و ارزیابی نقشه‌های مناطق دارای پتانسیل آب زیرزمینی به روش‌های مختلف از جمله منطق فازی و بولین پرداخته‌اند [۹، ۱۳].

مطالعه‌ای با هدف ارزیابی پتانسیل آب زیرزمینی با

دسترسی به منابع آب، مهم‌ترین عامل محدود کننده توسعه در مناطق خشک و نیمه‌خشک است. افزایش جمعیت را می‌توان یکی از مهم‌ترین عواملی دانست که اهمیت اکتشاف و تعیین نواحی پتانسیل حضور منابع آب زیرزمینی را بیش از پیش هویدا ساخته است. منابع آب زیرزمینی را می‌توان مخازن مهم دینامیک طبیعی محدود قلمداد کرد که به دلیل نوساناتی که دارد اندازه‌گیری زمانی و فضایی آن مشکل است. اگرچه منابع آب سطحی از طریق منافذ و شکستگی‌های سطح زمین به درون زمین نفوذ می‌کنند، اما برای یافتن آبخوان مطمئن، جمع‌آوری اطلاعات مربوط به سایر متغیرهای مؤثر در پتانسیل‌یابی و تعیین عوامل مؤثر نیز ضروری می‌باشد [۳]. با توجه به شرایط اقلیمی، جغرافیایی و توپوگرافی، بخش وسیعی از کشور ایران جزء مناطق خشک و نیمه خشک به شمار می‌رود. با افزایش تقاضا برای استحصال و مصرف آب از این مناطق، تعیین نواحی با پتانسیل آب زیرزمینی و حفاظت و مدیریت آن اهمیت بسیاری دارد. وجود آب زیرزمینی در هر منطقه، نتیجه برهم‌کنش عوامل اقلیمی، زمین‌شناسی، هیدرولوژیک، فیزیوگرافی و اکولوژیک است [۲۸]. به منظور مدیریت صحیح این منابع با ارزش و با در نظر گرفتن عوامل مؤثر می‌توان به تهیه نقشه‌های پتانسیل و پیش‌بینی نواحی حضور آب زیرزمینی با استفاده از روش‌های مختلف در هر منطقه پرداخت [۲۷]. نتایج کار بسیاری از پژوهشگران مبین این مطلب است که عوامل تعیین کننده مناطق مستعد آب زیرزمینی متفاوت بوده و عواملی همچون خطواره‌ها و شکستگی‌ها، تراکم زهکشی، ژئومورفولوژی، زمین‌شناسی، شیب، کاربری اراضی، شدت بارندگی و بافت خاک از جمله عوامل مؤثر در پتانسیل‌یابی منابع آب زیرزمینی محسوب می‌گردد [۱۲ و ۴۱].

<sup>5</sup> Random forest

<sup>6</sup> boost Regression tree

<sup>7</sup> Maximum Entropy

<sup>8</sup> Artificial Neural Network

<sup>1</sup> Frequency Ratio

<sup>2</sup> Weighted of Evidence

<sup>3</sup> Confidence Factor

<sup>4</sup> Logistic Regression

استفاده از نقاط چشمه در شهرستان بیرجند انجام شد؛ بدین منظور از روش‌های نسبت فراوانی، وزن واقعه و رگرسیون لجستیک در محیط سامانه اطلاعات جغرافیایی استفاده کردند و نشان دادند که روش نسبت فراوانی با مقدار سطح زیر منحنی ( $AUC^1$ ) برابر  $75/69$  درصد بهترین عملکرد را داشته است [۳۵]. در پژوهشی دیگر، از مدل نسبت فراوانی و آنتروپی شانون به منظور تهیه نقشه پتانسیل آب زیرزمینی در حوزه آبخیز مغان خراسان رضوی استفاده کردند؛ نتایج آن‌ها نشان داد که مدل آنتروپی شانون با سطح زیر منحنی برابر با  $91/21$  درصد کارایی بسیار مناسبی داشته است [۳۰]. از الگوریتم‌های جنگل تصادفی و حداکثر آنتروپی برای مدل‌سازی پتانسیل آب زیرزمینی استفاده شده و نتایجی با دقت خوب و قابل قبول حاصل شده است [۳۷]. علاوه بر این، الگوریتم ژنتیک (GA) برای حل مسائل چند هدفه در نظارت آب زیرزمینی و تعیین مکان‌های توزیع مطلوب توسط محققین مختلف مورد استفاده قرار گرفته است [۲۱].

## ۲.۲. روش تحقیق

با استفاده از روش‌های پتانسیل‌یابی می‌توان در مدت زمان کوتاهی منطقه وسیعی را مورد بررسی قرار داده و مکان‌های مناسب و مستعد حضور آب زیرزمینی را تعیین نمود. بررسی مطالعات قبلی نشان می‌دهد که روش‌های داده‌کاوی در سال‌های اخیر در مدل‌سازی و تعیین پهنه‌های خصوصیات هیدرولوژیکی به فراوانی مورد استفاده قرار گرفته و نتایج قابل قبولی ارائه نمودند. بنابراین هدف از پژوهش حاضر نیز بررسی کارایی روش‌های داده‌کاوی در تعیین مناطق دارای پتانسیل آب زیرزمینی محدوده یاسوج سی‌سخت می‌باشد.

## ۲. روش شناسی

### ۱.۲. معرفی منطقه مورد مطالعه

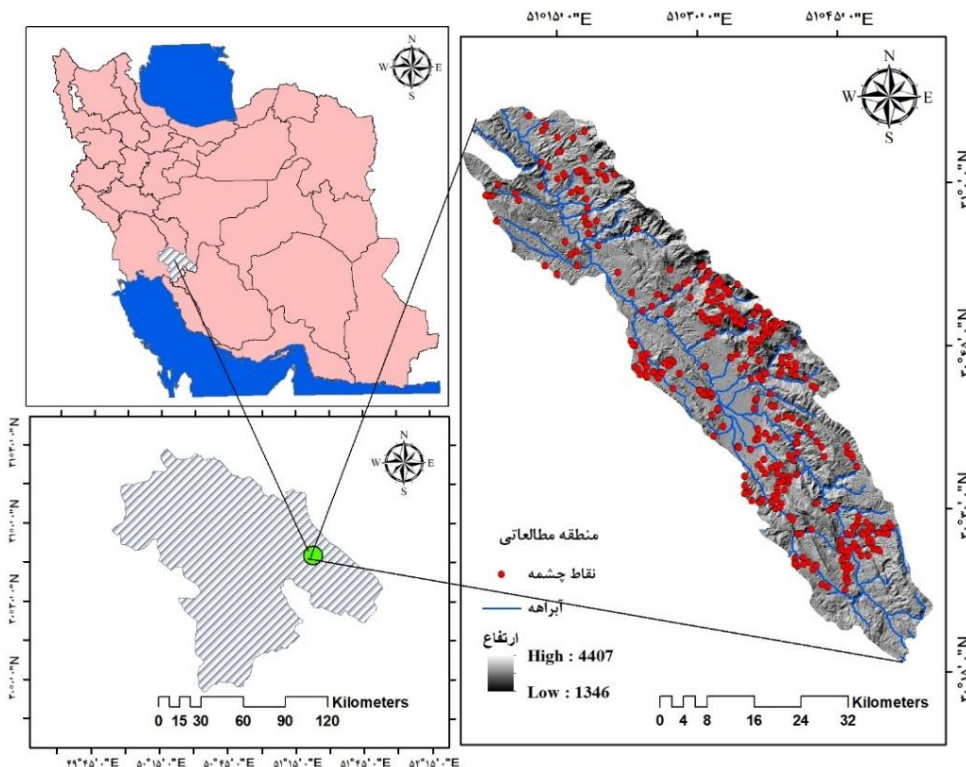
محدوده مورد مطالعه در منطقه یاسوج-سی سخت در استان کهگیلویه و بویراحمد و در موقعیت جغرافیایی

### ۳.۲. عوامل مؤثر بر وجود منابع آب زیرزمینی پتانسیل

در این پژوهش، ابتدا داده‌های مورد نیاز از نهادهای ذیربط و سازمان‌های متولی (شرکت مدیریت منابع آب ایران، شرکت آب منطقه‌ای و اداره کل منابع طبیعی

<sup>1</sup> Area under curve

استان) اخذ گردید.



شکل ۱. موقعیت منطقه مورد مطالعه و پراکنش چشمه‌ها

جدول ۱. سازندهای زمین‌شناسی برگرفته از نقشه زمین‌شناسی ۱:۲۵۰۰۰۰ سازمان زمین‌شناسی کشور

ردیف	علامت	سنگ‌شناسی	سن
۱	E	سنگ آهک توده‌ای تا ضخیم لایه کرم رنگ، سنگ آهک دولومیتی و مارن، دولومیت ضخیم لایه با هوازدگی کم تا متوسط (سازند آسماری-چهرم)	سنوزوئیک- پالئوسن
۲	Jkk	لایه‌های سنگ آهک سیاه، دولومیت سیاه و شیل، سنگ آهک سیا خاکستری توده‌ای با میان لایه‌های شیل (سازند سورمه و فهلیان)	مزوزوئیک
۳	K	سنگ آهک خاکستری اوربیتولینادار ضخیم لایه و سیل، تناوب لایه‌های سنگ آهک خاکستری و مارن کرم، سنگ آهک سیا خاکستری توده‌ای با میان لایه‌های شیل (سازندهای داریان، گدوان، فهلیان)	مزوزوئیک- آپتین
۴	Mgs	انیدریت، نمک و مارل خاکستری و قرمز با میان لایه‌های انیدریت و سنگ آهک همراه با لایه‌های رس (سازند گچساران)	مزوزوئیک- الیگوسن- میوسن
۵	OMa	سنگ آهک ضخیم لایه تا توده‌ای کرم رنگ و مارن (سازند آسماری)	سنوزوئیک- الیگومیوسن
۶	Plb	کنگومرای توده‌ای تا ضخیم لایه با میان لایه ماسه سنگ و مارن (سازند بختیاری)	سنوزوئیک- پلیستوسن
۷	Q	پادگانه و مخروط افکنه‌های آبرفتی قدیمی تا جوان (کواترنری)	سنوزوئیک
۸	Tr	دولومیت خاکستری تا سفید رنگ لایه نازک با میان لایه‌های شیل (سازند خانه کت)	پالئوزوئیک
۹	OE	سنگ آهک ضخیم لایه آسماری به همراه سازند چهرم	سنوزوئیک- پالئوسن- میوسن
۱۰	SpH	سنگ نمک، گچ و بلوک توده ای متخلخل از مواد رسوبی مانند سنگ آهک لمینیت سیاه و سفید، دولومیت قهوه‌ای، ماسه سنگ قرمز و شیل متغیر در ارتباط با سنگ‌های آذرین مانند بازالت و ریولیت	مزوزوئیک

باشد و به وسیله معادله زیر قابل محاسبه می‌باشد [۱۰].

$$TPI = \frac{Epixel}{Esurrounding} \quad \text{رابطه (۱)}$$

که در این معادله، Epixel ارتفاع هر پیکسل را نشان می‌دهد و Esurrounding نشان دهنده ارتفاع پیکسل‌های همسایه است.

یکی از مهم‌ترین بخش‌ها در پیش‌بینی مناطق دارای پتانسیل آب‌های زیرزمینی، سنگ‌شناسی است. نوع سنگ شناسی و خصوصیات وابسته به آن نظیر بافت، سن، جنس و درجه خلوص سنگ‌ها نقش مهمی در تخلخل، نفوذپذیری و تمرکز جریان آب زیرزمینی در داخل سنگ‌ها ایفا می‌کند. نقشه زمین شناسی منطقه مورد مطالعه با استفاده از نقشه‌های ۱:۲۵۰۰۰۰ سازمان زمین شناسی کشور تهیه شده و واحدهای سنگ شناسی در منطقه مورد مطالعه به ۱۰ سازند تقسیم‌بندی شدند (جدول ۱ و شکل ۲) [۲۵]. توپوگرافی سطح، جهت رواناب را در سطح زمین تعیین می‌کند و شامل پستی و بلندی‌های زمین می‌باشد. TWI یکی از مهم‌ترین عواملی است که بر تجمع و حرکت رواناب سطحی در سطح زمین اثر دارد. در نقشه‌های TWI عملکرد قابل توجهی برای توپوگرافی در تنوع مکانی شرایط هیدرولوژیک مانند رطوبت خاک و جریان آب زیرزمینی وجود دارد. این فاکتور می‌تواند رطوبت خاک، مناطق اشباع، عمق آب و تجمع جریان را نشان دهد [۲۰]. مقدار TWI با استفاده از معادله زیر محاسبه می‌شود:

$$TWI = \ln \frac{\alpha}{\tan \beta} \quad \text{رابطه (۲)}$$

در این معادله،  $\alpha$  نشان دهنده مساحت ویژه هر پیکسل و  $\beta$  نشان دهنده زاویه شیب هر پیکسل می‌باشد. طول شیب یک عامل مهم در سرعت جریان است. با

با توجه به سابقه پژوهش‌های مرتبط انجام شده در ایران و جهان، لایه‌های اطلاعاتی مختلف شامل درجه شیب، جهت شیب، طول شیب، ارتفاع از سطح دریا، شاخص رطوبت توپوگرافی ( $TWI^1$ )، فاصله از گسل، فاصله از آبراهه، بارندگی، کاربری اراضی، سنگ‌شناسی، شاخص موقعیت توپوگرافی ( $TPI^2$ ) و شاخص قدرت جریان ( $SPI^3$ ) به عنوان عوامل مؤثر بر پتانسیل آب زیرزمینی از روی نقشه‌های پایه (توپوگرافی، زمین‌شناسی و ...) موجود با استفاده از الحاقیه‌های مختلف در محیط سامانه اطلاعات جغرافیایی رقوم و تهیه گردید [۶ و ۳۶]. مدل رقوم ارتفاع منطقه مورد نظر با اندازه پیکسل  $12/5 \times 12/5$  متر اخذ شده و با استفاده از آن لایه‌های درجه شیب، جهت شیب، فاصله از آبراهه و  $SPI$ ،  $TWI$  در محیط نرم افزارهای ArcGIS 10.4 و SAGAGIS تهیه گردید [۳۶ و ۴۰]. نقشه‌های کاربری اراضی، فاصله از گسل، بارندگی و زمین شناسی با استفاده از نرم افزارهای ArcGIS 10.4 و ENVI 5.1 تهیه گردیدند [۳۱].

در بیشتر مواقع گرادیان هیدرولیکی آب زیرزمینی از شیب توپوگرافی زمین تبعیت می‌کند. بنابراین شیب می‌تواند فاکتور مهمی در جریان رواناب و نفوذ در نظر گرفته شود. جهت شیب نیز به دلیل تأثیری که از تابش خورشید می‌گیرد بر روی تبخیر و تعرق و مقدار آب نفوذ یافته به زمین مهم می‌باشد [۸]. نقشه شیب و جهت با استفاده از نرم افزار ArcGIS 10.4 و نقشه مدل رقوم ارتفاع منطقه مورد مطالعه ایجاد شد [۳۶ و ۴۰]. شکستگی‌ها و درزها و گسل‌ها با عبور دادن آب به لایه‌های زیرزمینی و همچنین جلوگیری از فرار آب در لایه‌های پایینی یکی از بهترین مکان‌ها برای شناسایی منابع آب زیرزمینی می‌باشد. لایه فاصله از گسل نیز در نرم افزار ArcGIS 10.4 با دستور تراکم خط ایجاد گردید [۳۱].  $TPI$  نیز نشان دهنده موقعیت مکانی هر پیکسل در یک نقشه توپوگرافی می‌

<sup>3</sup> Stream position index

<sup>1</sup> Topography wetness index

<sup>2</sup> Topography position index

متغیر وابسته و چندین متغیر مستقل است مدل رگرسیون لجستیک می‌تواند به ساده‌ترین شکل به صورت زیر بیان شود:

$$P = 1/(1 + e^2) \quad \text{رابطه (۴)}$$

که،  $P$  احتمال برآورد یک واقعه در حال وقوع است. از آنجا که  $Z$  می‌تواند از  $-\infty$  تا  $+\infty$  متغیر باشد، احتمال از ۰ تا ۱ به عنوان یک منحنی  $S$  شکل متغیر است. پارامتر  $Z$  به شرح زیر است:

$$Z = B_0 + B_1X_1 + B_2X_2 + \dots + B_nX_n \quad \text{رابطه (۵)}$$

که،  $B_0$  حائل است و  $n$  تعداد متغیرهای مستقل است. مقدار  $B_i$  ( $i = 0, 1, 2, \dots, n$ ) شیب متغیر است و  $X_i$  ( $i = 0, 1, 2, \dots, n$ ) متغیر مستقل است. بر اساس رابطه ۴ و ۵ رگرسیون لجستیک می‌تواند به صورت زیر ارائه شود [۲۸]:

رابطه (۶)

$$\text{Logit}(P) = 1/(1 + e^{-B_0+B_1X_1+B_2X_2+\dots+B_nX_n})$$

## ۵.۲. مدل جنگل تصادفی (RF)

این الگوریتم را نخستین بار لئوپرایمن و آدی کاتلر ایجاد کردند و توسعه دادند. الگوریتم جنگل تصادفی مبتنی بر دسته‌ای از درخت‌های تصمیم است و در حال حاضر یکی از بهترین الگوریتم‌های یادگیری است. این الگوریتم یکی از ابزارهای کارآمد مورد استفاده در مسائل مربوط به تخمین متغیرهای هدف و یا طبقه‌بندی الگوها، درخت تصمیم است. در ساخت هر درخت تصمیم، یک استراتژی پارتیشن‌بندی بازگشتی بالا به پایین استفاده می‌شود [۵].

مدل پیش‌بینی کننده (RF) بر اساس میانگین‌گیری از نتایج حاصل از تمامی درخت‌های تصمیم مربوطه استوار است و برای بسیاری از مجموعه داده‌ها، طبقه‌بندی با صحت بالایی را انجام می‌دهند. درختان تصادفی با در

طول شیب بزرگتر، آب بیشتری نیز در پایین دست حوضه تجمع می‌یابد [۲۸]. طول شیب با استفاده از معادله زیر محاسبه می‌شود.

$$LS = \left(\frac{As}{22.13}\right)^{0.6} \left(\frac{\sin\alpha}{0.0896}\right)^{1.3} \quad \text{رابطه (۳)}$$

در فرمول بالا،  $AS$  مساحت ویژه حوزه آبخیز و  $\alpha$  تخمین درجه شیب می‌باشد.

مقدار بارش تعیین کننده مقدار آبی خواهد بود که به سیستم‌های آب زیرزمینی به عنوان منبع اصلی تغذیه، نفوذ می‌کند. بارش دارای اثر قابل توجهی بر روی پتانسیل آب‌های زیرزمینی و بهره‌وری از آن‌ها است [۱].

SPI اندازه‌گیری قدرت فرسایش رودخانه‌ها یا قدرت مخرب جریان آب بر اساس این فرض است که تخلیه متناسب با سطح ویژه حوضه آبخیز است و به عنوان یک عامل مؤثر برای پایداری شیب در نظر گرفته شده و به طور معکوس متناسب با ثبات شیب است [۳۸]. انواع استفاده از زمین یکی از عوامل اصلی کنترل کننده تغذیه آب‌های زیرزمینی است. وجود پوشش گیاهی در هر منطقه سرعت جریان را کاهش داده و سبب نفوذ بیشتر آب به داخل خاک می‌گردد [۲۳]. نقشه کاربری اراضی منطقه مورد مطالعه نیز با استفاده از تصاویر ماهواره لندست ۸ در تاریخ ۲۴ شهریورماه ۱۳۹۲ و با استفاده از الگوریتم درست‌نمایی تهیه شد [۵] که شامل ۶ کلاس مختلف شامل: کاربری شهری، باغ، کشاورزی، مرتع، بایر و جنگل می‌باشد.

## ۴.۲. مدل خطی تعمیم یافته (GLM)

مدل خطی تعمیم یافته (رگرسیون لجستیک) یکی از متداول‌ترین روش‌های آماری است که برای مدل سازی مکانی متغیرهای محیطی مورد استفاده قرار می‌گیرد. رگرسیون لجستیک رابطه خطی بین یک تابع لینک (یعنی هویت، لجستیک) متغیر وابسته و متغیرهای مستقل مختلف را فرض می‌کند [۶ و ۳۴]. هدف اصلی این مدل یافتن بهترین مدل برای نشان دادن رابطه بین

انجام شد. متغیر اندازه گره (که نشان دهنده تعداد برگ‌ها در هر شاخه است) با آزمون و خطا تعیین گردید.

## ۶.۲. ارزیابی مدل‌های تهیه نقشه پتانسیل آب

### زیرزمینی

به منظور ارزیابی دقت مدل‌های مورد استفاده، از منحنی تشخیص عملکرد نسبی (ROC) استفاده شد. منحنی ROC روش کمی و گرافیکی مناسبی جهت تعیین میزان دقت مدل‌های پیش‌بینی مکانی و روش‌های داده‌کاوی محسوب می‌گردد. در این روش برای بررسی دقت نقشه نهایی از ۳۰ درصد چشمه‌ها که در پتانسیل‌یابی استفاده نشده برای بررسی دقت نقشه نهایی به کار برده می‌شود [۴۲]. در این منحنی میزان مثبت واقعی (بر روی محور  $X$ ) و میزان مثبت کاذب (بر روی محور  $Y$ ) در آستانه‌های مختلف نشان داده می‌شود. هنگامی که حساسیت افزایش پیدا کند، میزان مثبت کاذب نیز افزایش می‌یابد. بنابراین منحنی ROC به ما اجازه می‌دهد تا میزان مثبت واقعی و مثبت کاذب را در هر نقطه بر روی منحنی بررسی و مقایسه کنیم [۲]. مساحت زیر این منحنی (AUC) بیانگر میزان دقت نقشه نهایی به صورت کمی و توانایی مدل در پیش‌بینی درست می‌باشد. مساحت مذکور توسط نرم افزارهای آماری همچون نرم افزار آماری  $R$  و با استفاده از اطلاعات استفاده نشده‌ای که برای اعتبارسنجی به صورت تصادفی در نظر گرفته می‌شود، محاسبه می‌گردد [۲۲]. در یک حالت ایده‌آل مقدار زیر منحنی برابر یک خواهد بود و مقادیر کمتر از آن، دقت کمتر مدل را نشان می‌دهد. این شاخص برای ارزیابی صحت مدل، شاخص مناسبی می‌باشد.

### ۳. نتایج

مدل‌سازی پتانسیل آب زیرزمینی با استفاده از دو مدل جنگل تصادفی و مدل خطی تعمیم یافته بر اساس

نظر گرفتن بردار ورودی، آن را با هر درخت در جنگل طبقه‌بندی کرده و خروجی، برچسب‌های کلاسی هستند که از اکثریت آرا دریافت شده است [۴۸].

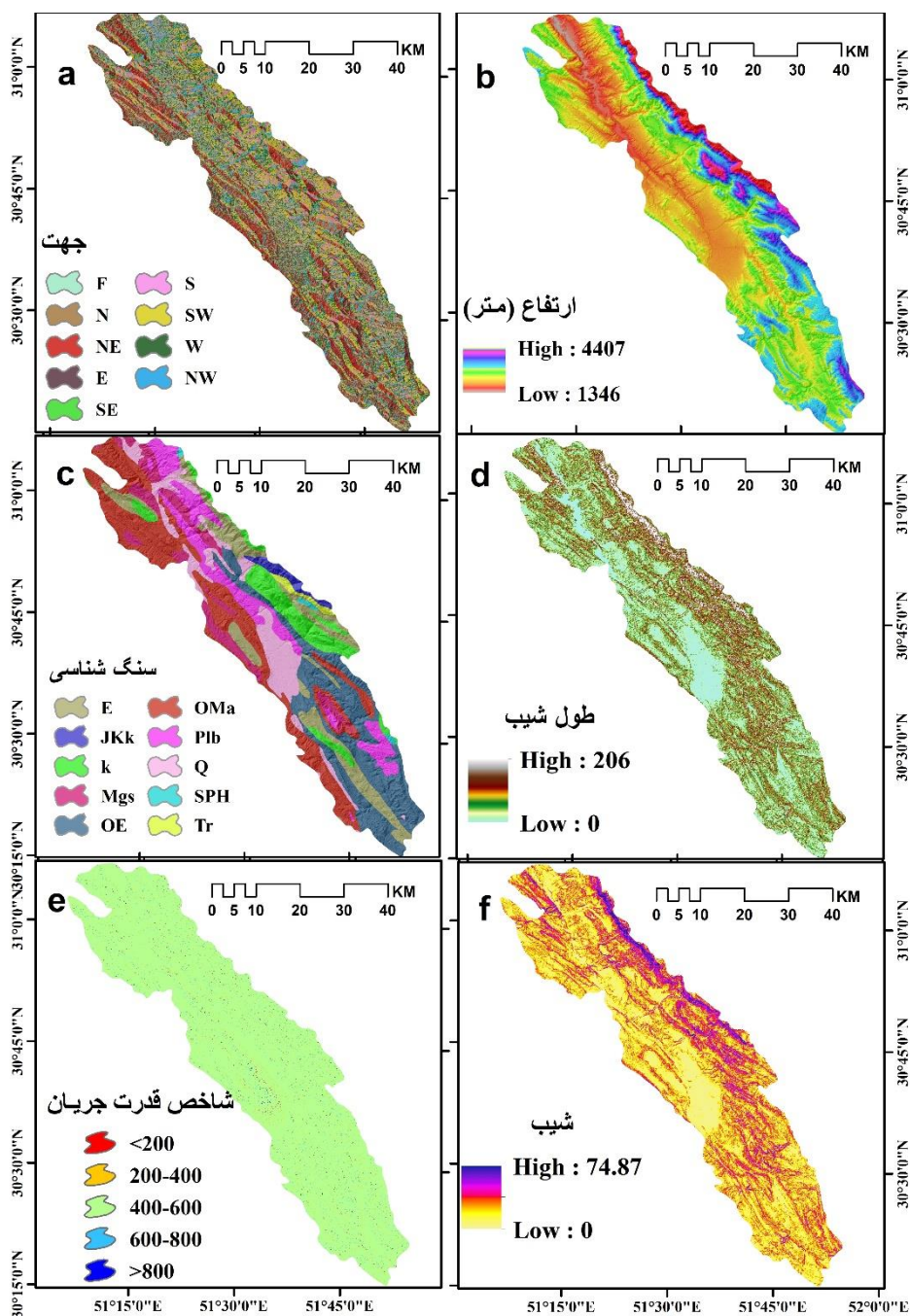
جنگل تصادفی با استفاده از مجموعه‌ای از درخت‌ها با در نظر گرفتن  $n$  داده مشاهده‌ای مستقل ساخته می‌شود [۳۵].

$$(Y, X), i=1, \dots, n \quad (1)$$

این روش ترکیبی از چندین درخت تصمیم است که در ساخت آن چندین نمونه بوت استرپ از داده‌ها شرکت دارند و در ساخت هر درخت به طور تصادفی تعدادی از متغیرهای ورودی شرکت می‌کنند. با استفاده از روش بوت استرپ به تعداد زیاد (به عنوان مثال ۲۰۰۰ مرتبه) نمونه‌هایی  $n$  تایی از مجموعه داده‌های مشاهداتی اولیه، نمونه برداری و جای گذاری می‌شوند. در طی فرآیند نمونه‌گیری حدود یک سوم از داده‌ها نمونه‌گیری نشده و به عنوان نمونه خارج از کیسه (از این داده‌ها برای تعیین متغیرهای مهم و همچنین برآورد نا اریب خطا استفاده می‌شود) در نظر گرفته می‌شوند. سپس بر روی هر نمونه بوت استرپ یک درخت گسترش داده می‌شود. در طی فرآیند ساخت درخت در هر شاخه، از بین تمام  $M$  متغیر مستقل به صورت تصادفی  $m$  متغیر برای تقسیم شدن انتخاب می‌شود. برای حالت رگرسیون نسبت  $m/M$  برابر با یک سوم است و برای کلاسه‌بندی برابر با  $m = \sqrt{M}$  پیشنهاد شده است [۳۶ و ۳۷]. پس از ساخت تمام درخت، داده‌های تست به درخت معرفی شده و به تعداد درخت‌ها برای بردار ورودی یک خروجی به دست می‌آید. با میانگین‌گیری این خروجی‌ها، خروجی نهایی مدل و با در نظر گرفتن توزیع تجربی خروجی‌ها مقادیر صدک‌ها و دامنه عدم قطعیت محاسبه می‌شود. روش درخت رگرسیون جنگل تصادفی به ویژه هنگامی که تعداد مشاهدات در مقایسه با تعداد پیش‌بینی کننده‌ها نسبتاً کم باشد یک روش پیش‌بینی کارآمد است [۴۴]. در این پژوهش محاسبات مدل جنگل تصادفی در محیط نرم افزار  $R$

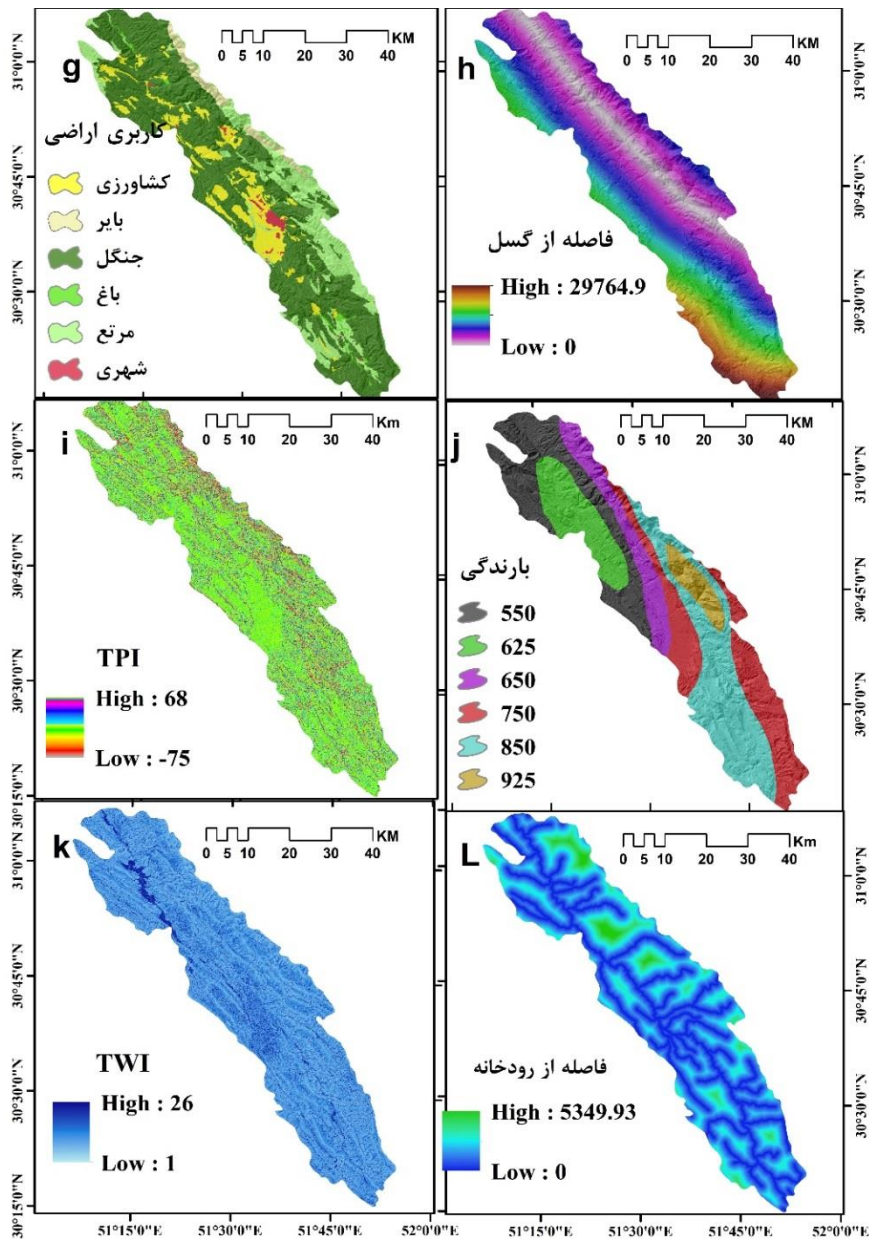
مهم‌ترین عوامل مؤثر بر آب زیرزمینی به ترتیب شامل ارتفاع، بارندگی، فاصله از گسل می‌باشد و در مدل رگرسیون خطی تعمیم یافته نیز مهمترین عامل‌ها بارندگی، فاصله از گسل و ارتفاع است که نشان می‌دهد هر دو مدل شاخص‌های یکسانی را به عنوان مؤثرترین عامل معرفی کرده است.

پکیج biomod2 در محیط نرم افزار R انجام شد. شکل های ۲ و ۳ شاخص‌های مورد بررسی در مدل‌های مد نظر را نشان می‌دهد. نتایج حاصل از تعیین بهترین پارامترهای مدل جنگل تصادفی در جدول ۲ نشان داده شده است. پس از تجزیه و تحلیل‌های صورت گرفته بین عوامل مورد بررسی، اهمیت هر عامل مشخص گردید (شکل ۴)؛ بر اساس نتایج این نمودار، در مدل جنگل تصادفی (RF)





شکل ۲. نقشه شاخص‌های جهت، ارتفاع از سطح دریا، سنگ‌شناسی، طول شیب، شاخص قدرت جریان و درجه شیب



شکل ۳. نقشه شاخص‌های فاصله از گسل، کاربری اراضی، بارندگی، فاصله از رودخانه، TPI و TWI

جدول ۲. تعیین بهترین پارامترهای مدل جنگل تصادفی

تعداد گره	آزمون و خطا	تعداد درخت	بهترین درخت
۵	۵	۵۰۰	۱۸۹

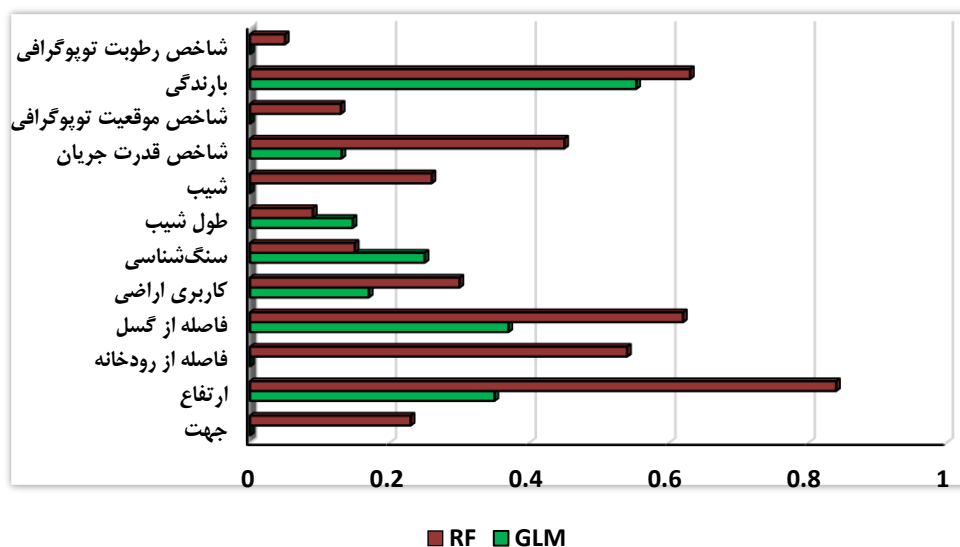
طبقه کم، متوسط، زیاد و خیلی زیاد تهیه گردید (شکل ۵). نتایج حاصل از پهنه‌بندی پتانسیل آب زیرزمینی نشان

برای تهیه نقشه پتانسیل آب زیرزمینی در منطقه مورد مطالعه با استفاده از ۱۳ شاخص مؤثر، نقشه نهایی در چهار

و در روش رگرسیون خطی تعمیم یافته به ترتیب ۱۴/۴۹، ۳۲/۰۴، ۳۱/۱۱ و ۲۲/۳۶ درصد می‌باشد (جدول ۲).

داد که سطح طبقات حضور آب زیرزمینی با پتانسیل کم، متوسط، زیاد و خیلی زیاد در نقشه حاصل از روش جنگل تصادفی به ترتیب ۳۷/۷۸، ۲۲/۲۲، ۱۸/۸۹ و ۲۱/۱۱ درصد

اهمیت فاکتورها



شکل ۴. اهمیت هر کدام از عوامل ۷ تری بر پتانسیل یابی منابع آب زیرزمینی

جدول ۳. درصد مساحت مربوط به هر کلاس نقشه پتانسیل آب زیرزمینی در دو مدل RF و GLM

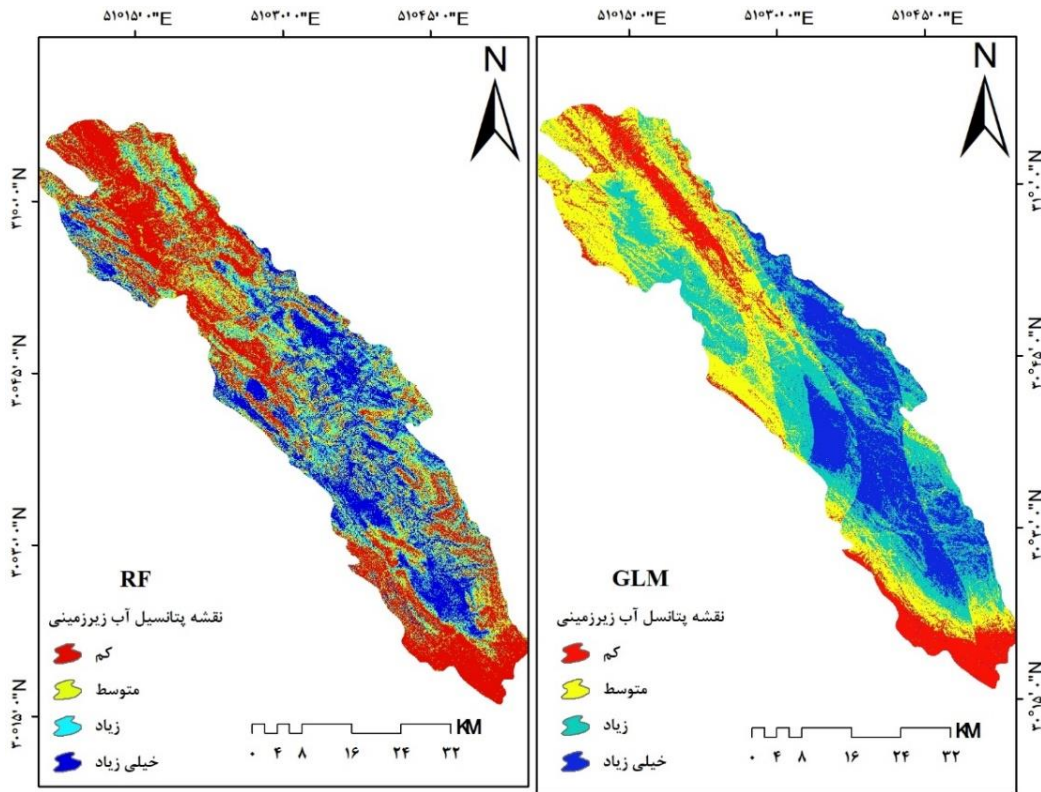
کلاس	مدل RF درصد مساحت هر طبقه	مدل GLM درصد مساحت هر طبقه
کم	۳۷/۷۸	۱۴/۴۹
متوسط	۲۲/۲۲	۳۲/۰۴
زیاد	۱۸/۸۹	۳۱/۱۱
خیلی زیاد	۲۱/۱۱	۲۲/۳۶

۳۶۲ نقطه تصادفی از منطقه مورد مطالعه (نقاط عدم وجود چشمه) قرار گرفت. قابلیت پیش‌بینی مدل به وسیله سطح زیرمنحنی مشخص می‌گردد. نتایج ارزیابی مدل های جنگل تصادفی و رگرسیون خطی تعمیم یافته با استفاده از منحنی ROC نشان داد که مدل جنگل تصادفی با مقدار AUC برابر با ۹۲ درصد در مقایسه با مدل رگرسیون خطی تعمیم یافته با مقدار AUC برابر با

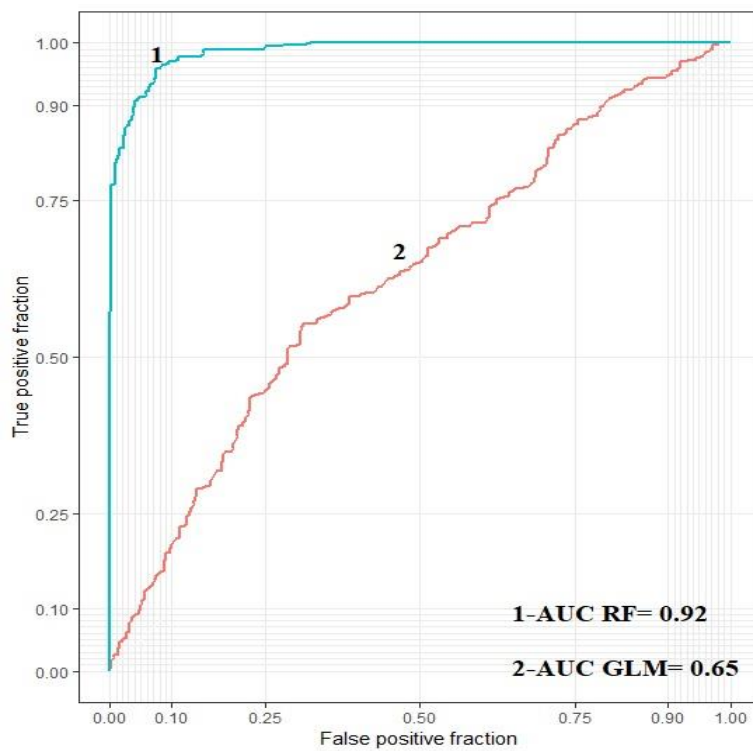
یکی از روش های تعیین دقت و ارزیابی مدل ها، استفاده از منحنی تشخیص عملکرد نسبی (ROC) است. مقادیر این منحنی بین ۰.۵ تا ۱ می‌باشد و برای ارزیابی دقت مدل به کار می‌رود. هر چقدر این مقدار به ۱ نزدیک تر باشد، دقت مدل بیشتر است [۲۶ و ۳۳]. جهت به کارگیری این روش موقعیت مکانی مجموعه چشمه های مشاهداتی منطقه مورد مطالعه (۳۶۲ چشمه) در برابر

طبقه خیلی خوب قرار می‌گیرد (شکل ۶).

۶۵ درصد عملکرد بهتری داشته است که این مقدار در



شکل ۵. نقشه نهایی پتانسیل آب زیرزمینی با استفاده از دو مدل RF و GLM



شکل ۶. نمودار ROC مربوط به نقشه پتانسیل آب زیرزمینی برای ارزیابی دقت دو مدل RF و GLM

#### ۴. بحث و نتیجه گیری

نقشه پتانسیل آب زیرزمینی یک تلاش سیستماتیک است که به کمک عوامل کلیدی و کنترل کننده به دست می‌آید. با توجه به هزینه‌بر بودن آزمایش‌های اکتشافی منابع آب زیرزمینی، تهیه نقشه پتانسیل آب زیرزمینی با استفاده از روش‌های نوین با هدف حفاظت، مدیریت و بهره‌برداری صحیح از منابع آب امری حائز اهمیت است. هدف از این پژوهش پتانسیل‌یابی آب زیرزمینی در منطقه یا سوچ-سی سخت با استفاده از دو روش داده‌کاوی جنگل تصادفی و رگرسیون خطی تعمیم یافته بود. بدین منظور از ۱۳ عامل تأثیرگذار بر پتانسیل منابع آب زیرزمینی استفاده گردید. که سه عامل ارتفاع، فاصله از گسل و بارندگی مهم ترین شاخص مؤثر بر حضور آب زیرزمینی تشخیص داده شدند و سایر عوامل با اهمیت کمتر در رتبه‌های بعدی قرار گرفتند. با توجه به کوهستانی بودن منطقه مورد مطالعه و وجود ارتفاعات زیاد در قسمت‌های شمال و شمال غربی اغلب بارش‌های موجود در این ارتفاعات به شکل برف می باشد و به دلیل وجود سازندهای آهکی آسماری در این مناطق عامل افزایش ارتفاع باعث تقویت انحلال در این سازندها شده و در نتیجه نفوذپذیری در این مناطق گسترش می‌یابد بنابراین پتانسیل وجود آب زیرزمینی در این مناطق بالا می‌باشد. یکی دیگر از عوامل مهم تأثیرگذار در پتانسیل آب زیرزمینی وجود گسل در منطقه است. بررسی نقشه فاصله از گسل نشان داد که هر چه فاصله از گسل کمتر باشد به دلیل وجود درز و شکاف‌های موجود نفوذپذیری افزایش پیدا کرده و پتانسیل آب زیرزمینی بیشتر می‌شود به طور کلی، تراکم چشمه‌ها و در نتیجه پتانسیل آب زیرزمینی در بخش شمالی محدوده مورد مطالعه بیشتر است، دلیل آن را می‌توان وجود ارتفاعات با سازندهای کارستی و نیز وجود گسل سراسری دنا دانست. در بخش جنوبی منطقه نیز به دلیل تراکم گسل و وجود آبراه‌های متعدد به ویژه رودخانه بشار، پتانسیل آب زیرزمینی بالا می‌باشد که با نتایج سایر محققان مطابقت دارد [۲۹ و ۴۸].

طبق نتایج حاصل از تهیه نقشه‌های پتانسیل آب زیرزمینی بر اساس مدل‌های جنگل تصادفی و مدل رگرسیون خطی تعمیم یافته مشخص می‌شود که نتایج دو مدل مذکور با یکدیگر تفاوت کمی دارند. به طوری که در مدل جنگل تصادفی، مناطق با پتانسیل بالا اغلب در قسمت شمالی و جنوب شرقی قرار دارند در حالی که در مدل رگرسیون خطی تعمیم یافته مناطق با پتانسیل بالا بیشتر به صورت پراکنده در مناطق مختلف منطقه مورد مطالعه می‌باشد. دلیل تفاوت این دو مدل را می‌توان در ساختار مدل‌های استفاده شده و عدم قطعیت‌های موجود در مدلسازی عنوان کرد [۲۸]. مهم‌ترین عوامل در هر دو مدل یکسان می‌باشد ولی ترتیب آن‌ها متفاوت است به طوری که در مدل جنگل تصادفی مهم‌ترین عامل ارتفاع می‌باشد در حالی که در مدل رگرسیون خطی تعمیم یافته بارندگی مهم‌ترین عامل بوده است. پهنه‌بندی پتانسیل‌یابی آب زیرزمینی با مدل‌های جنگل تصادفی و رگرسیون خطی تعمیم یافته و ارزیابی دقت آن‌ها با استفاده از منحنی ROC، بیانگر دقت خیلی خوب مدل جنگل تصادفی (۹۲ درصد) در مقایسه با مدل رگرسیون خطی تعمیم یافته (۶۵ درصد) است که با نتایج پژوهشگران دیگر مبنی بر دقت خوب مدل جنگل تصادفی در پتانسیل‌یابی آب زیرزمینی تطابق دارد [۲۵، ۳۷ و ۴۸]. مدل جنگل تصادفی به دلیل توانایی شناسایی روابط غیرخطی پیچیده بین عوامل تأثیرگذار در آب‌های زیرزمینی و پتانسیل آب‌های زیرزمینی، مناسب است. همچنین می‌تواند به طور خودکار تعامل بین عوامل مؤثر در آب‌های زیرزمینی را در نظر بگیرد [۱۴ و ۳۷].

پیشنهاد می‌شود به منظور صرفه جویی در زمان و هزینه، مدل‌های نوین داده‌کاوی و یادگیری ماشین به منظور بررسی اولیه کارهای اجرایی مدنظر قرار گیرد. همچنین نتایج حاصل از این تحقیق می‌تواند برای اتخاذ تصمیم‌های دقیق مدیریتی به متولیان و نهادهای اجرایی بخش مدیریت منابع آب استان کهگیلویه و بویراحمد کمک نماید. در نهایت توصیه می‌شود به منظور درک بهتر

تر، بررسی تفصیلی در منطقه مورد مطالعه انجام شود.

پتانسیل آب زیرزمینی هر منطقه و همچنین ارزیابی دقیق

## References

- [1] Adiat, K.A.N., Nawawi, M.N.M. and Abdullah, K. (2012). Assessing the accuracy of GIS-based elementary multi criteria decision analysis as a spatial prediction tool--a case of predicting potential zones of sustainable groundwater resources. *Journal of Hydrology*, 440, 75–89.
- [2] Akobeng, A.K. (2007). Understanding diagnostic tests 3: receiver operating characteristic curves. *Acta paediatrica*, 96(5), 7-644.
- [3] Alai Taleghani, M. and Saeidikiaya, M. (2013). The role of geomorphologic components in the formation and feeding of groundwater table (Case study: Zahab Plain). *Geographical Research Quarterl*, 28(10), 171-186.
- [4] Barzegar, R., Moghaddam, A.A., Deo, R., Fijani, E. and Tziritis, E. (2018). Mapping groundwater contamination risk of multiple aquifers using multi-model ensemble of machine learning algorithms. *Science of the Total Environment*, 621, 697–712.
- [5] Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.
- [6] Bui, D.T., Bui, Q.T., Nguyen, Q.P., Pradhan, B., Nampak, H. and Trinh, P.T. (2017). A hybrid artificial intelligence approach using GIS-based neural-fuzzy inference system and particle swarm optimization for forest fire susceptibility modeling at a tropical area. *Agricultural and Forest Meteorology*, 233, 32–44.
- [7] Chen, W., Li, H., Hou, E., Wang, S., Wang, G., Panahi, M., Li, T., Peng, T., Guo, C., Niu, C., others.. (2018). GIS-based groundwater potential analysis using novel ensemble weights-of-evidence with logistic regression and functional tree models *Science of the Total Environment*, 634, 853–867.
- [8] Corsini, A., Cervi, F. and Ronchetti, F. (2009). Weight of evidence and artificial neural networks for potential groundwater spring mapping: an application to the Mt. Modino area (Northern Apennines, Italy). *Geomorphology*, 111, 79–87.
- [9] Dai, F.C., Lee, C.F., Li, J. and Xu, Z.W. (2001). Assessment of landslide susceptibility on the natural terrain of Lantau Island, Hong Kong. *Environmental Geology*, 40, 381–391.
- [10] Dar, I.A., Sankar, K. and Dar, M.A. (2010). Remote sensing technology and geographic information system modeling: an integrated approach towards the mapping of groundwater potential zones in Hardrock terrain, Mamundiyar basin. *Journal of Hydrology*, 394, 285–295.
- [11] De Reu, J., Bourgeois, J., Bats, M., Zwervaegher, A., Gelorini, V., De Smedt, P., Chu, W., Antrop, M., De Maeyer, P., Finke, P., others. (2013). Application of the topographic position index to heterogeneous landscapes. *Geomorphology*, 186, 39–49.
- [12] Farhoodi, G. and Eghbalpoor, A. (2006). Structural analysis of the northwest of Dana mountains. *Geological Journal Educational Monthly*, 3, 112-118.
- [13] Ganapuram, S., Kumar, G.T.V.V., Krishna, I.V.M.M., Kahya, E. and Demirel, M.C. (2009). Mapping of groundwater potential zones in the Musi basin using remote sensing data and GIS. *Advances in Engineering Software*, 40, 506–518.
- [14] Gaur, S., Chahar, B.R. and Graillet, D. (2011). Combined use of groundwater modeling and potential zone analysis for management of groundwater. *International Journal of Applied Earth Observation*, 13, 127–139.
- [15] Golkarian, A., Naghibi, S.A., Kalantar, B. and Pradhan, B. (2018). Groundwater potential mapping using C5.0, random forest, and multivariate adaptive regression spline models in GIS. *Environmental Monitoring and Assessment*, 190, 149-155.
- [16] Guru, B., Seshan, K. and Bera, S. (2017). Frequency ratio model for groundwater potential mapping and its sustainable management in cold desert, India. *Journal of King Saud University*, 29, 333–347.
- [17] Hou, E., Wang, J. and Chen, W. (2018). A comparative study on groundwater spring potential analysis based on statistical index, index of entropy and certainty factors models. *Geocarto International*, 33, 754–769.

- [18] Iranian Department of Water Resources Management (IDWRM) 2017. A brief report regarding water resources of Iran. p 118. <http://www.wrm.ir/> (accessed in May 2017)
- [19] Jafari, A., Rahimi, M. and tahmasebi, P. (2017). Diversity variation and plant species composition due to grazing of livestock in white taffy ranges. *Environmental Research*, 8, 131–142.
- [20] Kalantar, B., Pradhan, B., Naghibi, S.A., Motevalli, A. and Mansor, S. (2018). Assessment of the effects of training data selection on the landslide susceptibility mapping: a comparison between support vector machine (SVM), logistic regression (LR) and artificial neural networks (ANN). *Geomatics, Natural Hazards and Risk*, 9, 49–69.
- [21] Khalaf, S. and Gad, M.I. (2014). Optimal well locations using genetic algorithm for Tushki Project, Western Desert, Egypt *International Journal of Engineering Research & Science*, 3(7), 1-11.
- [22] Kumar, R. and Indrayan, A. (2011). Receiver operating characteristic (ROC) curve for medical researchers. *Indian pediatrics*, 48(4), 87-277.
- [23] Lee, S. and Oh, H.J. (2012). Ensemble-based landslide susceptibility maps in Jinbu area, Korea. *Terrigenous Mass Movements*, Springer, Berlin, Heidelberg, 193–220.
- [24] Manap, M.A., Nampak, H., Pradhan, B., Lee, S., Sulaiman, W.N.A. and Ramli, M.F. (2014). Application of probabilistic-based frequency ratio model in groundwater potential mapping using remote sensing data and GIS. *Arabian Journal of Geosciences*, 7, 711–724.
- [25] Manap, M.A., Sulaiman, W.N.A., Ramli, M.F., Pradhan, B. and Surip, N. (2013). A knowledge-driven GIS modeling technique for groundwater potential mapping at the Upper Langat Basin, Malaysia. *Arabian Journal of Geosciences*, 6, 1621–1637.
- [26] Mathew, J., Jha, V.K. and Rawat, G.S. (2007). Application of binary logistic regression analysis and its validation for landslide susceptibility mapping in part of Garhwal Himalaya, India. *International Journal of Remote Sensing*, 28, 2257-2275.
- [27] Moghaddam, D.D., Rezaei, M., Pourghasemi, H.R., Pourtaghie, Z.S. and Pradhan, B. (2015). Groundwater spring potential mapping using bivariate statistical model and GIS in the Taleghan watershed, Iran. *Arabian Journal of Geosciences*, 8, 913–929.
- [28] Naghibi, S.A. and Moradi Dashtpajardi, M. (2016). Evaluation of four supervised learning methods for groundwater spring potential mapping in Khalkhal region (Iran) using GIS-based features. *Journal of Hydrology*, 25, 169–189.
- [29] Naghibi, S.A. and Pourghasemi, H.R. (2015). A comparative assessment between three machine learning models and their performance comparison by bivariate and multivariate statistical methods in groundwater potential mapping. *Water resources management*, 29(14), 5217-5236.
- [30] Naghibi, S.A., Pourghasemi, H.R. and Dixon, B. (2016). GIS-based groundwater potential mapping using boosted regression tree, classification and regression tree, and random forest machine learning models in Iran. *Environmental Monitoring and Assessment*, 188, 44.
- [31] Naghibi, S.A., Pourghasemi H.R., Pourtaghi, Z.S. and Rezaei, A. (2015). Groundwater qanat potential mapping using frequency ratio and Shannon's entropy models in the Moghan watershed, Iran. *Earth Science Informatics*, 8, 171–186.
- [32] Nampak, H., Pradhan, B. and Manap, M.A. (2014). Application of GIS based data driven evidential belief function model to predict groundwater potential zonation. *Journal of Hydrology*, 513, 283–300.
- [33] Nobre, R.C.M., Rotunno Filho, OC., Mansur, W.J., Nobre, M.M/M. and Cosenza, C.A.N. (2007). Groundwater vulnerability and risk mapping using GIS, modeling and a fuzzy logic tool. *Journal of Contaminant Hydrology*, 94, 277–292.
- [34] Oh, H.J. and Pradhan, B. (2011). Application of a neuro-fuzzy model to landslide susceptibility mapping for shallow landslides in a tropical hilly area. *Computers & Geoscience*, 37, 1264-1276.
- [35] Ozdemir, A. (2011). Using a binary logistic regression method and GIS for evaluating and mapping the groundwater spring potential in the Sultan Mountains (Aksehir, Turkey). *Journal of Hydrology*, 405, 123–136.
- [36] Pourghasemi, H.R. and Rossi, M. (2017). Landslide susceptibility modeling in a landslide prone area in Mazandarn Province, north of Iran: a comparison between GLM, GAM, MARS, and M-AHP methods. *Theoretical and Applied Climatology*, 130(1-2), 609-633.

- [37] Pourtaghi, Z.S. and Pourghasemi, H.R. (2014). GIS-based groundwater spring potential assessment and mapping in the Birjand Township, southern Khorasan Province, Iran. *Hydrogeology Journal*, 22(3), 643-662.
- [38] Rahmati, O., Naghibi, S.A., Shahabi, H., Bui, D.T., Pradhan, B., Azareh, A., Rafiei-Sardooi, E., Nazari Samani, A. and Melesse, A.M. (2018). Groundwater spring potential modelling: Comprising the capability and robustness of three different modeling approaches. *Journal of Hydrology*, 565, 248–261.
- [39] Rahmati, O., Pourghasemi H.R. and Melesse, A.M. (2016). Application of GIS-based data driven random forest and maximum entropy models for groundwater potential mapping: A case study at Mehran Region, Iran. *Catena*, 137, 360–372.
- [40] Rahmati, O., Tahmasebipour, N., Haghizadeh, A., Reza, H. and Feizizadeh, B. (2017). Geomorphology Evaluation of different machine learning models for predicting and mapping the susceptibility of gully erosion. *Geomorphology*, 298, 118–137.
- [41] Sahoo, S., Russo, T.A., Elliott, J. and Foster, I. (2017). Machine learning algorithms for modeling groundwater level changes in agricultural regions of the US. *Water Resources Research*, 53, 3878–3895.
- [42] Sarkar, B.C., Deota, B.S., Raju, P.L.N. and Jugran, D.K. (2001). A geographic information system approach to evaluation of groundwater potentiality of shamri micro-watershed in the Shimla Taluk, Himachal Pradesh. *Journal of the Indian Society of Remote Sensing*, 29, 151–164.
- [43] Sener, E., Davraz, A. and Ozcelik, M. (2005). An integration of GIS and remote sensing in groundwater investigations: A case study in Burdur, Turkey. *Journal of Hydrology*, 13, 826–834.
- [44] Shahbazi, A., Tahmasebipoor, N., Zeynivand, H. and Haghizade, A. (2015). Groundwater potential potential using the EBF definite evidence function model. National Conference on Sustainable Agriculture, Environment and Rural Development, Koohdasht, Kuhdasht Agricultural Jihad Department, New Line Fencing Co.
- [45] Shahid, S., Nath, Sk. and Roy, J. (2000). Groundwater potential modelling in a soft rock area using a GIS. *International Journal of Remote Sensing*, 21, 1919–1924.
- [46] Svetnik, V., Liaw, A., Tong, C., Culberson, J.C., Sheridan, R.P. and Feuston, B.P. (2003). Random forest: a classification and regression tool for compound classification and QSAR modeling. *The Journal for Chemical Information and Computer scientists*, 43, 1947–1958.
- [47] Tongal, H. and Booij, M.J. (2018). Simulation and forecasting of streamflows using machine learning models coupled with base flow separation. *Journal of Hydrology*, 564, 266–282.
- [48] Wang, X., Liu, T., Zheng, X., Peng, H., Xin, J. and Zhang, B. (2018). Short-term prediction of groundwater level using improved random forest regression with a combination of random features. *Applied Water Science*, 8(5), 125.
- [49] Zabihi, M., Pourghasemi, H.R. and Behzadfar, M. (2015). Preparation of groundwater potential map using shannon entropy models and random forest in Bojnourd plain, *Ecohydrology*, 2(2), 221-232. (In persian)

