

## GIS-based Groundwater Spring Potential Modelling and Assessment Mapping in the the Omarak Watershed

HAMED KARIMI<sup>1</sup>, FARIBORZ YOSEFVAND<sup>1\*</sup>, SAEID SHABANLOU<sup>1</sup>, AHMAD RAJABI<sup>1</sup>

1. Department of Water Engineering, Kermanshah Branch, Islamic Azad University, Kermanshah, Iran.

(Received: March. 16, 2021- Revised: July. 1, 2021- Accepted: July. 10, 2021)

### ABSTRACT

Nowadays, in most countries, water supply in order to achieve the objectives of sustainable development is one of the most important challenges. Because of this, one of the important tools in the protection, management and exploitation of water resources, is to determine groundwater areas. Therefore, the purpose of this study, is to prepare the potential map of groundwater springs, using a well-known machine-learning model (i.e. random forest) and a statistical model (i.e. frequency ratio model) and comparing the efficiency of these methods in the Omarak watershed, Tehran Province. First, 18 factors influencing the emergence of springs including: lithological formations, the distance from the fault, fault density, elevation classes, slope percentage, slope direction, slope length factor, curvature maps, distance from the stream, stream density, maximum height, wetness index, relative slope position, soil texture, terrain roughness index, flow convergence index and land use cover were selected and their maps were prepared in the ArcGIS10.5 and SAGA systems. After the Multicollinearity test and classification of the effective layers, using the natural fracture method, then, the percentage of groundwater potential frequency in each layer obtained using the overlap of the distribution map of the springs with each of the layers. The relative operating characteristic (ROC) curve was used to evaluate the performance of the mentioned models and the area under the curve (AUC) of the random forest models and the frequency ratio were 88 and 72%, respectively. The results indicated that both methods are suitable to estimate to prepare the groundwater source potential map in the studied area. However, the random forest model with a higher area under the curve was introduced as a better method to identify and zoning the potential of groundwater springs.

**Keywords:** Groundwater Potential, Spring, Frequency Ratio Model, Stochastic Forest Model, Relative Operating Characteristic Curve.

## مدلسازی و ارزیابی پتانسیل چشمه‌های آب زیرزمینی مبتنی بر GIS در حوزه آبخیز عمرک

حامد کریمی<sup>۱</sup>، فریبرز یوسفوند<sup>۱\*</sup>، سعید شعبانلو<sup>۱</sup>، احمد رجبی<sup>۱</sup>

۱. گروه مهندسی آب، واحد کرمانشاه، دانشگاه آزاد اسلامی، کرمانشاه، ایران.

(تاریخ دریافت: ۱۳۹۹/۱۲/۲۶ - تاریخ بازنگری: ۱۴۰۰/۴/۱۰ - تاریخ تصویب: ۱۴۰۰/۴/۱۹)

### چکیده

امروزه تأمین آب به منظور تحقق اهداف توسعه پایدار، یکی از مهم‌ترین چالش‌ها در اکثر کشورهای جهان است. به همین دلیل، تعیین مناطق دارای پتانسیل آب زیرزمینی، از ابزارهای مهم در حفاظت، مدیریت و بهره‌برداری از منابع آب به شمار می‌رود. بنابراین هدف از این پژوهش تهیه نقشه پتانسیل چشمه‌های آب زیرزمینی با استفاده از یک مدل مشهور یادگیری ماشین (جنگل تصادفی) و یک مدل آماری (مدل نسبت فراوانی) و مقایسه کارایی این روشها در حوزه آبخیز عمرک، استان تهران می‌باشد. ابتدا ۱۸ عامل مؤثر در ظهور چشمه‌ها شامل: سازندهای سنگ‌شناسی، فاصله از گسل، تراکم گسل، طبقات ارتفاعی، درصد شیب، جهت شیب، فاکتور طول شیب، نقشه‌های انحنا، فاصله از آبراهه، تراکم آبراهه، حداکثر ارتفاع، شاخص نمناکی، موقعیت شیب نسبی، بافت خاک، شاخص زبری سطح، شاخص همگرایی جریان و پوشش کاربری اراضی انتخاب شدند و نقشه آن‌ها در نرم‌افزار ArcGIS 10.5 و SAGA<sup>۱</sup> تهیه گردید. پس از بررسی هم‌خطی و طبقه‌بندی لایه‌های مؤثر با استفاده از روش شکستن طبیعی<sup>۲</sup>، درصد فراوانی پتانسیل آب زیرزمینی در هر طبقه با استفاده از هم‌پوشانی نقشه پراکنش چشمه‌ها با هر کدام از لایه‌ها به دست آمد. برای ارزیابی عملکرد مدل‌های مذکور از منحنی مشخصه عملکرد نسبی (ROC) استفاده شد و میزان مساحت زیر منحنی (AUC) مدل‌های جنگل تصادفی و نسبت فراوانی به ترتیب ۸۸ و ۷۲ درصد به دست آمد. نتایج نشان داد که هر دو روش تخمین‌گرهای مناسبی برای تهیه نقشه پتانسیل چشمه آب زیرزمینی در منطقه مورد مطالعه هستند. اما مدل جنگل تصادفی با مقدار مساحت زیر منحنی بالاتر روش بهتری برای شناسایی و پهنه‌بندی پتانسیل چشمه‌های آب زیرزمینی معرفی گردید.

**واژه‌های کلیدی:** پتانسیل آب زیرزمینی، چشمه، مدل نسبت فراوانی، مدل جنگل تصادفی، منحنی مشخصه عملکرد نسبی.

### مقدمه

زمینه می‌باشد (Jha et al., 2007). اکثر شهرها و روستاهای ایران برای تأمین آب مورد نیاز شرب، آبیاری، صنعت و... به آبهای زیرزمینی متکی هستند. از سویی دیگر بخش وسیعی از کشور را رسوبات آبرفتی و کواترنری فراگرفته که غالباً توانایی بسیار خوبی در ذخیره آب زیرزمینی دارند. بنابراین شناخت مناطق مستعد استخراج آب زیرزمینی، هم می‌تواند مقدار آب بیشتر و باکیفیت بهتر را در اختیار قرار دهد و هم اینکه هزینه حفاری و استحصال آب زیرزمینی را کاهش دهد (Zandi et al., 2016). روش‌های مختلفی جهت تعیین پتانسیل آب زیرزمینی وجود دارد که مهم‌ترین آن‌ها استفاده از امواج صوتی، حفر چاه گمانه و روش ژئوالکتریک است (Rangzan et al., 2005). با این حال، نکته قابل توجه این است که این روش‌ها با وجود داشتن دقت زیاد، به مهارت، تخصص و صرف هزینه و زمان زیاد نیاز

آب‌های زیرزمینی یکی از منابع مهم تأمین آب شیرین مورد نیاز انسان است. با توجه به نیاز روزافزون کشور ایران به آب و همچنین توزیع نامتعادل زمانی و مکانی منابع آب‌های سطحی و پتانسیل بالای آلودگی منابع آب زیرزمینی، شناخت این منابع به عنوان یک اصل مهم در مدیریت منابع آب، امری ضروری و اجتناب‌ناپذیر تلقی می‌گردد. منابع آب زیرزمینی به دلیل ضریب اطمینان بالاتر و نوسانات کمتر، به عنوان یک گزینه مطمئن از دیرباز مورد استفاده انسان بوده است. با این وجود، این منابع در طی دهه‌های اخیر و در اثر برداشت بیشتر از تغذیه، با کاهش کمی و کیفی روبرو شده است. مدیریت و جلوگیری از تشدید این مشکلات از طریق اکتشاف و بهره‌برداری متناسب با پتانسیل آب زیرزمینی یکی از استراتژی‌های منتخب در این

\* نویسنده مسئول: fariborzoyosefvand@gmail.com

<sup>۱</sup> System for automated geoscientific analyses

<sup>۲</sup> Natural Break

تأثیر بر توان اکولوژی یک سرزمین یک پدیده مهم و مؤثر در توسعه اقتصادی، تنوع اکولوژیکی و سلامت جامعه به حساب می‌آید (Pradhan, 2009). بنابراین تعیین پتانسیل منابع آب زیرزمینی با توجه به نیاز روزافزون کشور ایران به آب، امری ضروری و اجتناب ناپذیر تلقی می‌گردد.

با وجود تحقیقات انجام شده در زمینه پهنه‌بندی پتانسیل چشمه‌های آب زیرزمینی، انجام پژوهش در این زمینه و بررسی مدل‌های نوین داده‌کاوی و روش‌های معروف آماری و احتمالاتی می‌تواند افق روشنی را پیش روی محققین کشورمان در خصوص مطالعه آب‌های زیرزمینی قرار دهد. هدف اصلی از انجام این پژوهش، بررسی فاکتورهای مؤثر در ظهور چشمه‌ها، معرفی مهمترین فاکتورها و تعیین نواحی با پتانسیل بالای ظهور چشمه‌های در حوضه عمرک می‌باشد.

## مواد و روش‌ها

### منطقه مطالعاتی

حوضه عمرک یکی از سرشاخه‌های حوضه حبله رود است که با مساحت ۱۱۳۶۵ هکتار بین  $22^{\circ} 00'$ ،  $52^{\circ}$  تا  $30'$ ،  $32^{\circ}$ ،  $52'$  طول شرقی و  $47^{\circ} 00'$  تا  $35^{\circ}$ ،  $56'$  عرض شمالی در فاصله ۳۰ کیلومتری شمال غرب شهرستان فیروزکوه قرار داشته و از شمال و شمال غرب با حوضه رود هراز، از شمال شرق تا جنوب با حوزه فرح رود محدود می‌شود. آبراهه اصلی حوضه با طول ۱۹/۱۴ کیلومتر از ارتفاعات کوه چنار، کیوان چال سرچشمه گرفته و در جنوب منطقه به فرح رود می‌ریزد. ارتفاعات مهم حوضه عبارتند از کوه چنار، میان‌چنار، بوم و کیوان چال در شمال، کوه وردل در غرب و کوه خشکه رودبار، کینزور، کردپیله در شرق. حوزه آبخیز عمرک به ۸ واحد هیدرولوژیکی اصلی و دو واحد جانبی تقسیم شده است. از نظر زمین‌شناسی این حوضه در زون زمین‌شناسی البرز مرکزی در زیر زون جنوبی مرکزی که در آن رسوبات قبل از دوران سنوزوئیک توسط آذرآواری ائوسن پوشیده شده است و سازندهای دارای رخنمون در گستره منطقه به دوران مزوزوئیک، سنوزوئیک و کواترنری تعلق دارد. گستره منطقه از نظر تکتونیکی در زون ساختاری البرز در شمال گسل مشاء در پهنه البرز شمالی واقع شده است. این پهنه از شمال به وسیله گسل خزر و از جنوب به وسیله گسل‌های مشاء فیروزکوه بسته می‌شود. گسلها، درزه‌ها و پین‌ها، مهم‌ترین عناصر ساختاری منطقه می‌باشند. حوضه آبخیز عمرک به طور کلی از گردش

دارند. در حالی که در مدل پیش‌بینی پتانسیل آب زیرزمینی مهارت زیادی نیاز نبوده و در زمان و هزینه صرفه‌جویی خواهد شد. بنابراین استفاده از روش‌های مدل‌سازی مکانی و پارامترهای محیطی (مقدار بارندگی، سنگ‌شناسی، و...) برای تعیین پتانسیل ذخیره و استحصال آب زیرزمینی گامی روبه جلو در زمینه مدیریت منابع آب زیرزمینی می‌باشد. مدل‌های متعددی در پتانسیل‌یابی منابع آب زیرزمینی استفاده شده‌اند که از آن جمله می‌توان مدل‌های وزن‌واژه<sup>۱</sup>، رگرسیون لجستیک<sup>۲</sup>، آنتروپی شانون<sup>۳</sup> و نسبت فراوانی<sup>۴</sup> را ذکر کرد (Oh *et al.*, 2011; Moghaddam *et al.*, Pourtaghi and Pourghasemi, 2014; Pourtaghi and Pourghasemi, 2015; Naghibi *et al.*, 2015). (2014) به ارزیابی پتانسیل آب زیرزمینی با استفاده از موقعیت چشمه‌های موجود در شهرستان بیرجند پرداختند. بدین منظور ایشان از مدل‌های وزن‌واژه، رگرسیون لجستیک و نسبت فراوانی استفاده کردند و نشان دادند روش نسبت فراوانی با مقدار مساحت زیرمنحنی ۷۵ درصد بهترین عملکرد را داشته است. (Zabihi *et al.*, 2015) به تهیه نقشه پتانسیل آب زیرزمینی با استفاده از مدل‌های آنتروپی شانون و جنگل تصادفی در دشت بجنورد پرداختند. نتایج ارزیابی دقت روش‌های پهنه‌بندی پتانسیل آب زیرزمینی با استفاده از منحنی تشخیص عملکرد نسبی بیانگر دقت بالای هر دو مدل بود. همچنین براساس مدل آنتروپی شانون؛ لایه‌های کاربری اراضی، لیتولوژی، فاصله از آبراهه و طبقه ارتفاعی بیشترین تأثیر را بر پتانسیل آب زیرزمینی در منطقه مطالعاتی داشته‌اند. (Nguyen *et al.*, 2020)، به تهیه نقشه پتانسیل آب زیرزمینی با استفاده از مدل رگرسیون لجستیک در بخشی از منطقه ویتنام پرداختند. در پژوهش مذکور از ۲۰ عامل محیطی از جمله، شیب، جهت شیب، ارتفاع، انحنای دامنه، شاخص انتقال رسوب، شاخص رطوبت توپوگرافی، جهت جریان، بارش، تراکم زهکشی، خاکشناسی، کاربری اراضی و سنگ‌شناسی به‌عنوان پیش‌بینی‌کننده‌ها استفاده شد. معیار ارزیابی AUC و همچنین سایر معیارهای ارزیابی از جمله حساسیت، صحت سنجی، تشخیص و کاپا، توانایی بالای این مدل را در پیش‌بینی نقشه پتانسیل آب زیرزمینی نشان داد. آب زیرزمینی از یک سو به دلیل شیرین بودن، ترکیبات ثابت شیمیایی، دمای ثابت، ضریب آلودگی کمتر و سطح اطمینان بیشتر یک منبع قابل اتکا به ویژه در مناطق خشک و نیمه خشک محسوب شده و از سوی دیگر با

<sup>۳</sup> Shannon's Entropy

<sup>۴</sup> Frequency Ratio

<sup>۱</sup> Weights-of-Evidence

<sup>۲</sup> Logistic Regression

۷۰ درصد به عنوان نمونه‌های آموزشی مورد استفاده قرار گرفت (Ozdemir, 2011؛ Ghosh and Carranza, 2010). موقعیت چشمه‌های موجود در منطقه در شکل یک نشان داده شده است. در تحلیل‌ها از چشمه‌های آموزشی استفاده گردید و اعتبار سنجی نقشه نهایی با استفاده از چشمه‌های آزمایشی انجام شد.

#### نقشه عوامل مؤثر بر ظهور چشمه‌ها

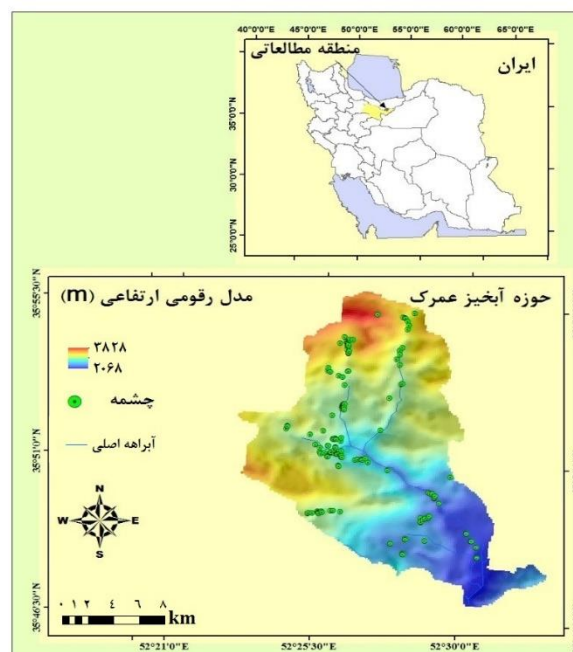
جهت شناسایی و ارزیابی نواحی با پتانسیل چشمه‌های آب زیرزمینی، باید چندین عامل مرتبط را با نقشه موجودی چشمه‌ها مورد ارزیابی قرار داد (Ozdemir, 2011). با مرور منابع و پژوهش‌های سابق انجام شده در جهان و ایران (Manap *et al.*, 2015; Zabihi *et al.*, 2014; Ozdemir, 2011)، در این پژوهش ۱۸ عامل مؤثر در ظهور چشمه‌ها شامل: سازندهای سنگ‌شناسی، فاصله از گسل، تراکم گسل، طبقات ارتفاعی، درصد شیب، جهت شیب، فاکتور طول شیب، نقشه‌های انحنا، فاصله از آبراهه، تراکم آبراهه، حداکثر ارتفاع، شاخص نمناکی، موقعیت شیب نسبی، بافت خاک، شاخص زبری سطح، شاخص همگرایی جریان و پوشش کاربری اراضی انتخاب شدند و نقشه آن‌ها در نرم‌افزار ArcGIS 10.5 و SAGA<sup>۲</sup> تهیه گردید.

سپس با استفاده از نقشه توپوگرافی در مقیاس ۱:۲۵۰۰۰، مدل رقومی ارتفاعی تهیه و عوامل درصد شیب، جهت شیب، طول شیب و نقشه‌های انحنا از آن استخراج شد. لایه کاربری اراضی و بافت خاک از اداره کل منابع طبیعی و آب‌خیزداری استان تهران اخذ شد. سپس نقشه کاربری اراضی براساس تصاویر گوگل ارث اصلاح شد. نقشه زمین‌شناسی منطقه مطالعاتی با مقیاس ۱:۱۰۰۰۰۰ از سازمان زمین‌شناسی کشور تهیه و لایه‌های فاصله از گسل و تراکم گسل از آن استخراج شد. لایه‌های مربوط به آبراهه نیز از نقشه توپوگرافی تهیه شد. نقشه شاخص رطوبت توپوگرافی (TWI<sup>۳</sup>) بر اساس مدل رقومی ارتفاع زمین در محیط GIS بدست آمد. TWI در روشن ساختن اثر توپوگرافی بر محل و اندازه منطقه اشباع شده و تولید رواناب کاربرد زیادی دارد. برای محاسبه شاخص TWI از معادله زیر استفاده شد (Zabihi *et al.*, 2015):

$$TWI = \ln \frac{AS}{\tan \beta} \quad (\text{رابطه ۱})$$

در اینجا AS مساحت منطقه مورد نظر و  $\beta$  گرادبان شیب است.

عمومی جو حاکم بر البرز مرکزی پیروی می‌کند؛ بنابراین در تابستان تحت تاثیر سیستم پرفشار جنب حاره قرار می‌گیرد، که بدین دلیل میزان بارش کم و ناچیز است. بارندگی‌های تابستانه اکثراً به خاطر عبور سیستم کم فشار دینامیکی می‌باشد، که از نواحی شمالی ایران عبور می‌کند و باعث بارش‌هایی در قسمت شمالی می‌شود. در برخی موارد جریان‌های گرم و مرطوب موسمی اقیانوس هند در لایه زیر جو منطقه تاثیر می‌گذارد و چنانچه با امواج سیستم‌های کم فشار عرض‌های میانی تداخل نماید، ضمن ایجاد ناپایداری بارش‌های شدیدتری را سبب می‌شود. متوسط بارندگی کل حوضه ۵۶۳/۰۵ میلی‌متر می‌باشد. اقلیم منطقه براساس روش دومارتن نیمه‌خشک می‌باشد (Seyed Ali *et al.*, 2016). شکل (۱) موقعیت منطقه مطالعاتی و موقعیت چشمه‌های استفاده شده در پژوهش به منظور تهیه نقشه پتانسیل آب زیرزمینی را نشان می‌دهد.



شکل ۱- موقعیت منطقه مطالعاتی و موقعیت چشمه‌های استفاده شده در پژوهش

#### نقشه‌ی پراکنش موقعیت چشمه‌های آب زیرزمینی

برای تهیه نقشه پتانسیل چشمه‌های آب زیرزمینی با استفاده از روش‌های نسبت فراوانی و جنگل تصادفی، ابتدا براساس آمار جمع‌آوری شده از اداره‌ی مدیریت منابع آب ایران (IDWRM) در سال ۱۳۹۸، بازدید صحرایی و تهیه پرسشنامه موقعیت چشمه‌های موجود در منطقه شناسایی گردیدند. از تعداد ۱۵۰ چشمه موجود در منطقه مورد مطالعه، ۳۰ درصد از موقعیت چشمه‌های موجود به صورت تصادفی جهت اعتبار سنجی مدل و

رگرسیون انجام می‌شود. برای تعیین تعداد بهینه درختان، ابتدا یک تعداد درخت اولیه جهت تولید گراف تغییرات میانگین مربعات خطا (MSE) در مقابل تعداد درختان نمونه آموزشی و ارزیابی است، مورد استفاده قرار می‌گیرد. این یک ابزار تحلیلی بسیار قدرتمند جهت جستجوی اطلاعات و اصلاح تعداد بهینه درختان در مدل جنگل تصادفی است. تعداد درختان بهینه به گونه‌ای انتخاب می‌شود که اولاً کمترین مقدار میانگین مربعات خطا در مدل‌سازی به دست آید و ثانیاً تعداد درختان نباید آن قدر زیاد شود که تجزیه و تحلیل پتانسیل وقوع نیازمند زمان و محاسبات کامپیوتری زیادی باشد. یکی از پارامترهای اصلی در اجرای مدل جنگل تصادفی، پیش‌بینی‌کننده  $K$  (متغیر مستقل) در هر گره جهت پیش‌بینی مقدار وابسته (پاسخ) است. ساده‌ترین راه جهت تعیین مقدار  $K$ ، محاسبه جذر تعداد کل متغیرهای مستقل می‌باشد (Shataee et al., 2011). کار ساخت جنگل با استفاده از درخت‌ها اغلب به روش کیسه‌گذاری انجام می‌گیرد. ایده‌ی اصلی روش کیسه‌گذاری آن است که ترکیبی از مدل‌های یادگیری، نتایج و عملکرد کلی مدل را افزایش می‌دهد. به بیان ساده، جنگل تصادفی، چندین درخت تصمیم می‌سازد و آنها را با یکدیگر ادغام می‌کند تا پیش‌بینی‌های صحیح و پایدارتری حاصل شوند. در این روش بردار تصادفی  $\Theta_k$  که مستقل از بردارهای تصادفی  $\Theta_1, \dots, \Theta_{k-1}$  بوده، برای درخت  $K$  ام، تولید می‌شود. همچنین همه‌ی بردارها دارای توزیع یکسانی می‌باشند. درخت رگرسیونی با استفاده از مجموعه داده‌های آموزش و  $\Theta_k$  رشد می‌کند و نتیجه‌ی مجموعه درخت‌های  $K$  برابر است با:

$$K = \{h_1(x), h_2(x) \dots h_k(x)\} \quad \text{(رابطه ۳)}$$

که در اینجا

$$X = \{x_1, x_2, \dots, x_p\} \quad \text{(رابطه ۴)}$$

$$h_k(x) = h(x, \theta_k) \quad \text{(رابطه ۵)}$$

این بردارها یک بردار ورودی  $P$  بعدی هستند که یک جنگل را تشکیل می‌دهند. خروجی‌های  $k$  تولیدشده گروهی مربوط به هر درخت برابر است با:

$$\widehat{y}_1 = h_1(x), \widehat{y}_2 = h_2(x), \dots, \widehat{y}_k = h_k(x) \quad \text{(رابطه ۶)}$$

که  $y_k$  خروجی درخت  $K$ ام می‌باشد. برای بدست آوردن خروجی نهایی، متوسط همه‌ی پیش‌بینی‌های درخت‌ها محاسبه می‌شود (Chen et al., 2020). در پژوهش حاضر به منظور به کارگیری این مدل در مدل‌سازی از نرم‌افزار R و پکیج

شاخص توان جریان (SPI) بر اساس مدل رقومی ارتفاع زمین در محیط GIS است. برای محاسبه این شاخص از معادله زیر استفاده شد:

$$SPI = AS \times \tan \beta \quad \text{(رابطه ۲)}$$

#### بررسی هم‌خطی<sup>۱</sup> بین عوامل مؤثر

با توجه به این که هم‌خطی یکی از دلایل افزایش خطا و کاهش کارایی مدل می‌باشد و ممکن است به پیش‌بینی‌های خارج از دامنه مورد انتظار منجر شود، از این رو پیش از فرایند مدل‌سازی این موضوع می‌باید بررسی شود (Zare Chahouki et al., 2014). از دلایل ایجاد هم‌خطی می‌توان به کاربرد نامناسب متغیرهای ساختگی، استفاده از یک متغیر که از طریق متغیرهای دیگر محاسبه شده و به کارگیری دوباره یک متغیر اشاره کرد (Kouhi Kamali and Rajabi, 2010). بعد از آماده‌سازی لایه‌ها، برای آنالیز این داده‌ها و بررسی هم‌خطی آنها از نرم‌افزار SPSS استفاده شد. با کنترل کردن شاخص‌های آستانه تحمل<sup>۲</sup> و عامل تورم واریانس<sup>۳</sup> (VIF)، می‌توان قضاوت درستی در مورد هم‌خطی داشت (Greene, 2000). در حقیقت VIF عکس آستانه تحمل است (Daoud, 2017). شرط هم‌خطی بودن لایه‌ها این است که VIF بیشتر از ۵ و آستانه تحمل کمتر از ۰/۱ باشد (Ozdemir, 2011) که این موضوع فرایند مدل‌سازی را دچار مشکل می‌کند. با برقرار بودن این شرط لایه‌های دارای همبستگی حذف و دوباره مدل بودن آن لایه اجرا می‌شود. بررسی همبستگی تا زمانی ادامه پیدا می‌کند که این مقادیر به حالت نرمال برسند.

#### مدل جنگل تصادفی<sup>۴</sup> (RF)

مدل جنگل تصادفی یکی از روش‌های مدل‌سازی درخت تصمیم است (Jafarian et al., 2019) به گونه‌ای که از تلفیق نتایج به‌دست‌آمده از درخت‌های مختلف، نتیجه یا برآورد نهایی را راجع به ارتباط میان پدیده مورد بررسی و عوامل محیطی ارائه می‌دهد. در مسائل مربوط به رگرسیون، جنگل تصادفی تعداد اختیاری از درختان ناآزموده (زیرمجموعه‌ای از متغیرهای مستقل) است که نتایج واکنش این درختان به صورت ترکیبی جهت تخمین صحیح پتانسیل وقوع (متغیر وابسته) مورد استفاده قرار می‌گیرد (Servati et al., 2014). نمونه‌گیری تصادفی از اطلاعات عوامل محیطی (متغیرهای مستقل)، به صورت خودکار و تکرارشونده جهت تولید انبوهی از درختان

۴ Random forest

۱ Multi-collinearity test

۲ Tolerance

۳ Variance Inflation Factor (VIF)

S نشان‌دهنده‌ی چشمه‌هاست.

#### ارزیابی نقشه

منحنی ROC یک روش کمی و گرافیکی جهت ارزیابی نقشه‌های احتمالاتی است. مساحت زیر این منحنی (AUC) شاخصی مناسب جهت تعیین دقت مدل‌های پیش‌بینی مکانی و تکنیک‌های داده‌کاوی می‌باشد. در این پژوهش برای ارزیابی نقشه‌های پتانسیل آب زیرزمینی بدست آمده، از دو روش نسبت فراوانی و جنگل تصادفی از شاخص ROC استفاده گردید (Moghaddam *et al.*, 2015, 2014). این روش توسط محققین بسیاری برای ارزیابی دقت نقشه‌های پتانسیل آب زیرزمینی، استعداد و ریسک زمین لغزش‌ها و استعداد وقوع فرسایش خندقی به کار گرفته شده است.

#### نتایج و بحث

نتایج مربوط به تهیه نقشه پتانسیل چشمه‌های آب زیرزمینی با استفاده از روش‌های نسبت فراوانی و جنگل تصادفی در قالب نقشه‌ها، نمودارها و جداول در این بخش ارائه شده است.

#### نقشه عوامل مؤثر بر ظهور چشمه‌ها

تصاویر مربوط به لایه‌های مرتبط با پیش‌بینی پتانسیل آب زیرزمینی از جمله سازندهای سنگ‌شناسی، فاصله از گسل، تراکم گسل، طبقات ارتفاعی، درصد شیب، جهت شیب، فاکتور طول شیب، نقشه‌های انحنای دامنه، فاصله از آبراهه، تراکم آبراهه، حداکثر ارتفاع، شاخص نمناکی، موقعیت شیب نسبی، بافت خاک، شاخص زبری سطح، شاخص همگرایی جریان و پوشش کاربری اراضی در شکل ۲ ارائه شده است.

#### نتایج بررسی هم‌خطی<sup>۲</sup> بین عوامل مؤثر بر ظهور چشمه در منطقه مورد مطالعه

در این بررسی، لایه‌های داری همبستگی حذف شده و مجدداً اجرای مدل انجام می‌شود. بررسی همبستگی تا زمانی ادامه پیدا می‌کند که این مقادیر به حالت نرمال برسند. در پژوهش حاضر با انجام این آزمون، لایه‌ی ارتفاع حداکثر از بین متغیرهای پیش‌بینی‌کننده حذف شد.

#### نقشه پتانسیل چشمه‌های آب زیرزمینی

برای تهیه نقشه پتانسیل آب زیرزمینی در روش نسبت فراوانی، ابتدا مطابق با جدول (۱) مقدار شاخص پتانسیل آب زیرزمینی، برای هر سلول محاسبه گردید و سپس نقشه پتانسیل آب زیرزمینی در روش نسبت فراوانی مطابق شکل (۳) تهیه گردید.

randomForest استفاده شده است. مدل جنگل تصادفی این قابلیت را دارد که میزان خطای به وجود آمده در پیش‌بینی استعداد وقوع خطرات را برحسب شاخص OOB<sup>۱</sup> نشان دهد که بر اساس آن می‌توان تعداد درختان ساخته‌شده برای آنالیز داده‌های وارد شده به نرم‌افزار را بهینه‌سازی کرد.

#### مدل نسبت فراوانی<sup>۲</sup> (FR)

مدل نسبت فراوانی به‌عنوان یک مدل آماری دو متغیره می‌تواند به‌عنوان یک ابزار ساده‌ی مکانی برای محاسبه رابطه‌ی احتمالی بین متغیرهای مستقل و وابسته استفاده شود که شامل چندین نقشه طبقه‌بندی شده است. این مدل، ارتباط کمی میان رخداد هر پدیده و متغیرهای مختلف تأثیرگذار بر آن را مشخص می‌نماید. در تعیین نرخ نسبت فراوانی، نسبت رخداد پدیده در هر کلاس از فاکتورهای تأثیرگذار، نسبت به کل خطر موردنظر به دست می‌آید و نسبت سطح هر کلاس نسبت به کل مساحت منطقه نیز محاسبه می‌شود. در نهایت، با تقسیم نرخ رخداد هر پدیده در هر کلاس بر نرخ مساحت هر کلاس نسبت به کل منطقه مورد مطالعه، نسبت فراوانی کلاس‌های هر فاکتور محاسبه می‌شود. مراحل محاسبه نسبت فراوانی برای هر کلاس از فاکتورهای مؤثر در ظهور چشمه‌های آب زیرزمینی در رابطه ۳ ارائه شده است (Erfanian *et al.*, 2016).

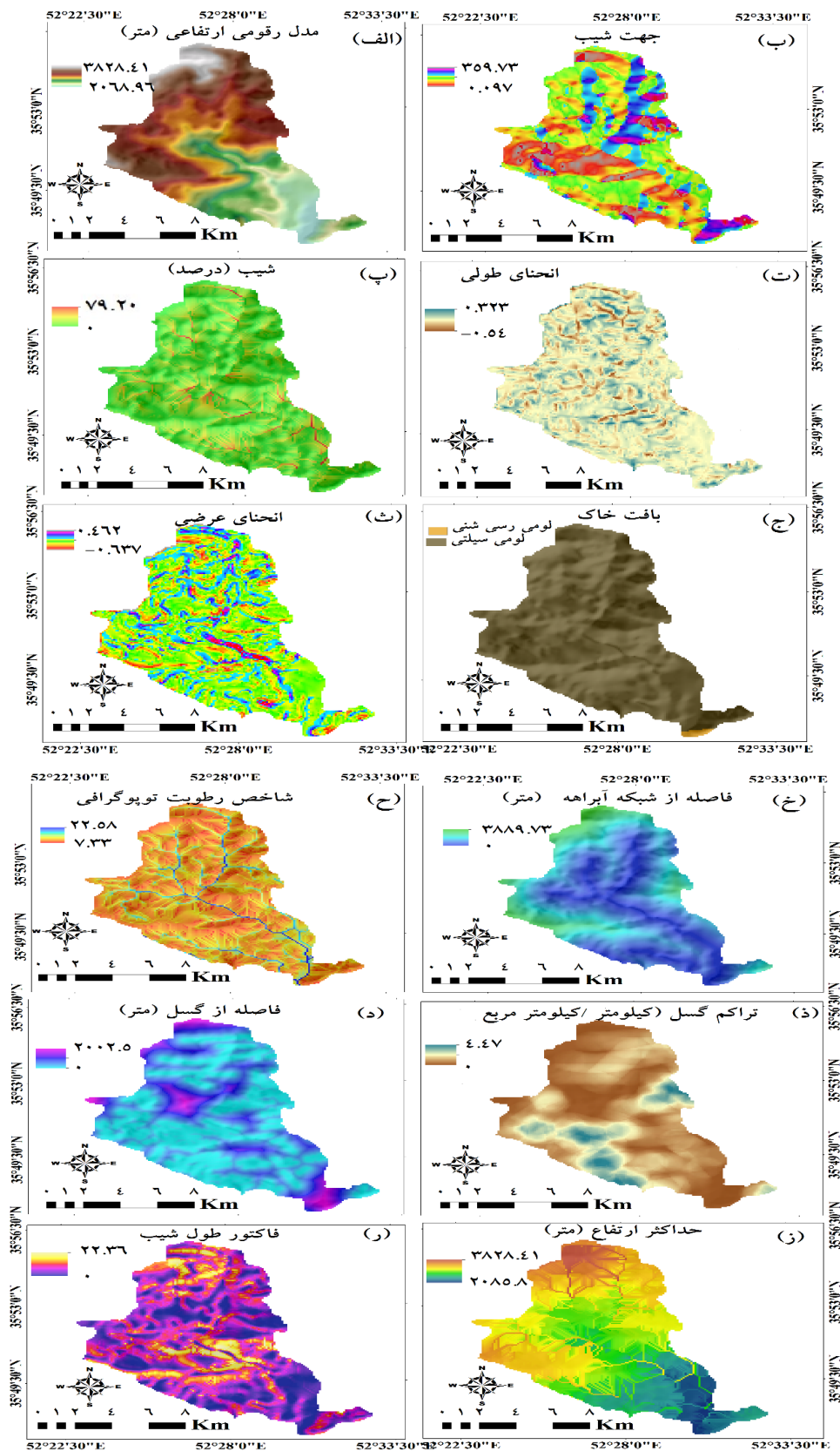
$$FR = \frac{A}{B} / \frac{C}{D} \quad (\text{رابطه ۷})$$

که در آن: A تعداد پیکسل‌های چشمه‌ها در کلاس موردنظر، B مجموع پیکسل‌های هر چشمه در کل حوضه، C تعداد پیکسل‌های کلاس موردنظر و D تعداد کل پیکسل‌های مربوط به فاکتور مربوطه (مثلاً شیب) می‌باشد. اگرچه روش نسبت فراوانی از مفاهیم ساده و قابل فهم استفاده می‌کند و همچنین قادر به انجام آنالیزهای آماری دو متغیره و طبقه‌های هر معیار بر وجود پتانسیل آب زیرزمینی را بررسی می‌کند، ضعف روش یاد شده در این است که اغلب رابطه‌ی بین متغیرها در آن نادیده گرفته می‌شود (Lee and Pradhan, 2007). برای به‌دست آوردن شاخص پتانسیل چشمه‌های آب زیرزمینی، نتایج فاکتورها در محیط GIS با یکدیگر جمع می‌گردد. مثلاً برای چشمه‌های آب زیرزمینی این رابطه به شکل زیر است (Lee and Pradhan, 2007):

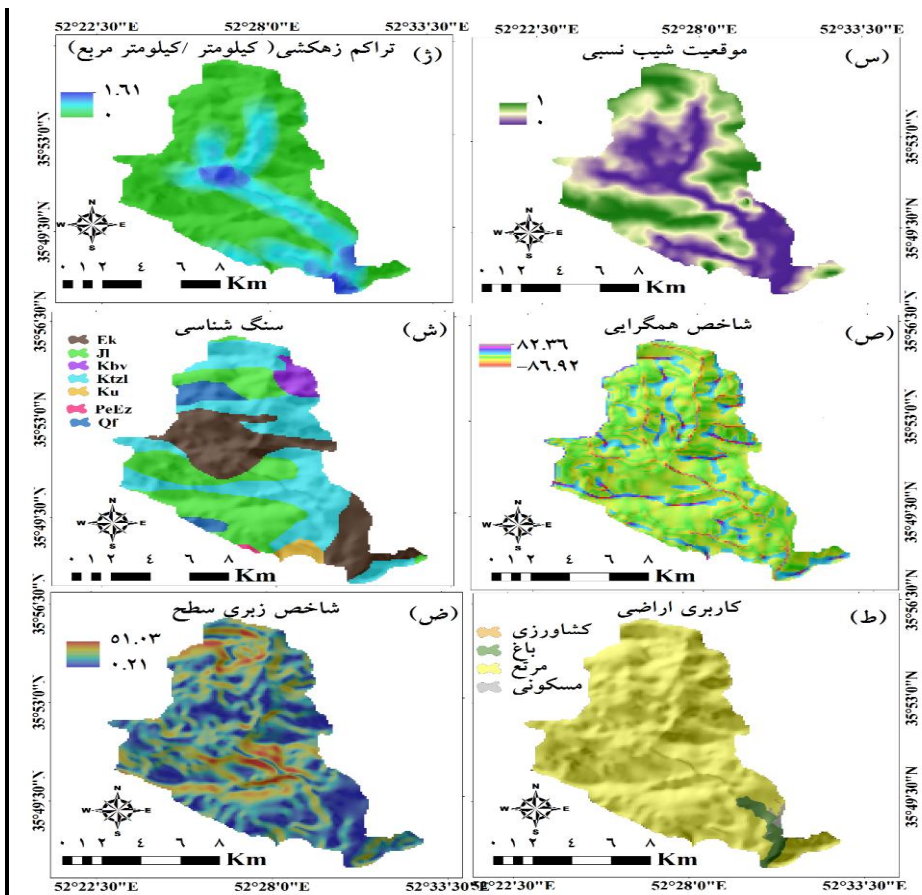
(رابطه ۸)

$$SI = \sum (FR)_i \quad (i=1,2,\dots,n)$$

براساس نتایج به‌دست آمده، طبقه‌ی ارتفاعی (۲۷۲۴-۳۳۶۶) بیشترین تعداد چشمه را در خود جای داده است و در طبقات







شکل ۲- نقشه عوامل مؤثر بر ظهور چشمه‌ها: الف) مدل رقمی ارتفاعی؛ ب) جهت شیب؛ پ) درصد شیب؛ ت) انحنای طولی دامنه؛ ث) انحنای عرضی دامنه؛ ج) بافت خاک؛ د) فاصله از شبکه آبراهه؛ ه) فاصله از گسل؛ و) تراکم گسل؛ ز) فاکتور طول شیب؛ ح) حداکثر ارتفاع؛ ز) تراکم زهکشی؛ س) موقعیت شیب نسبی؛ ش) سنگ‌شناسی؛ ص) شاخص همگرایی؛ ض) شاخص زبری سطح؛ ط) کاربری اراضی

عوامل دیگر نشان می‌دهد که بیشترین تعداد ظهور چشمه به ترتیب مربوط به کلاس‌های درصد شیب کمتر از ۲۲ درصد، طول شیب کمتر از ۶/۱۳ متر، تراکم زهکشی ۰/۸۹-۰/۵۲، رطوبت توپوگرافی ۱۱/۵۸-۹/۷۳، انحنای طولی و عرضی مسطح، کاربری اراضی مرتع و کشاورزی، فاصله از گسل کمتر از ۲۱۹ متر و تراکم گسل کمتر از ۰/۷ است. وجود رابطه‌ی معکوس بین فاصله از گسل و تعداد چشمه نیز از دیگر نتایج پژوهش حاضر است. این نتیجه بدان معناست که هرچه فاصله از گسل کمتر باشد، تعداد چشمه بیشتر است. دلیل این امر، امکان ورود آب زیرزمینی به سطح، در مناطق نزدیک به گسل است. همچنین بین فاصله از آبراهه، تراکم گسل و طول شیب با وجود پتانسیل آب زیرزمینی رابطه معکوس مشاهده می‌شود. نفوذ آب به اعماق و عدم ظهور بر سطح زمین در منطقه با تراکم زیاد گسل به دلیل ساختار آن، می‌تواند دلیلی بر کم بودن پتانسیل آب زیرزمینی باشد. نتایج پژوهش حاضر در خصوص ارتباط پتانسیل آب زیرزمینی با تراکم زهکشی نشان داد که بیشترین تعداد چشمه در طبقه ۰/۸۹-۰/۵۲ کیلومتر بر کیلومتر مربع از تراکم زهکشی قرار دارد. نتایج به دست آمده با مطالعه *Zabihi et al. (2015)* که به تهیه نقشه پتانسیل آب زیرزمینی در دشت بجنورد پرداختند همخوانی دارد.

دیگر تفاوت‌های چندانی از لحاظ تعداد چشمه وجود ندارد؛ که این امر احتمالاً به دلیل کم بودن تراکم زهکشی در این طبقه و در نتیجه نفوذ بیشتر و امکان ورود آب زیرزمینی از طریق درز و شکاف‌ها به سطح طبقه ارتفاعی متوسط است. براساس نتایج ارائه شده، بیشترین تعداد چشمه در کلاس جهت شیب شمال شرق و جنوب شرق است که این امر نشان‌دهنده‌ی زیاد بودن پتانسیل آب زیرزمینی در جهت مذکور است و کمترین آن در مناطق مسطح رخ داده است. شاید دلیل آن وجود گسل‌های منطقه‌ی مطالعاتی می‌باشد که به نحوی در ایجاد درزه و چشمه موثرند. نتایج همچنین نشان داد که فاصله صفر تا ۵۳۳ متری از آبراهه نیز بیشترین تعداد چشمه را در خود جای داده است. در مورد سنگ‌شناسی، کلاس EK و Ktzi، که به ترتیب مربوط به سازند کرج و تیزکوه است بیشترین تعداد ظهور چشمه را دارد. سازند کرج دارای لیتولوژی‌های متنوع آذر-آواری و گدازه‌های آندزیتی، بازالتی همراه با شیل و توفیت‌های مارنی به سن ائوسن زیرین-میانی بوده و سازند تیزکوه متشکل از سنگ‌های آذرین بازالتی قدیمی در قاعده و آهک‌های اوریتولین دار و سنگ‌های آهکی بیوژنیک چرت دار و سنگ‌های آهکی فسیل دار از کرتاسه آغازین تا پسین است. بررسی

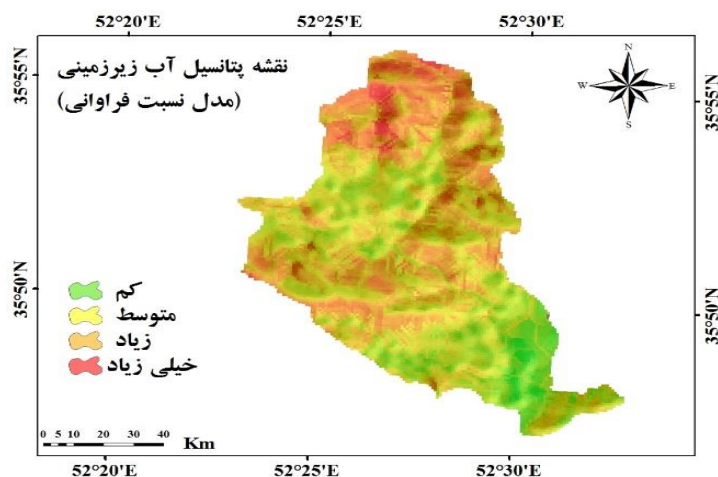


جدول ۱- رابطه مکانی نسبت فراوانی بین پتانسیل آب زیرزمینی و عوامل مؤثر بر ظهور چشمه

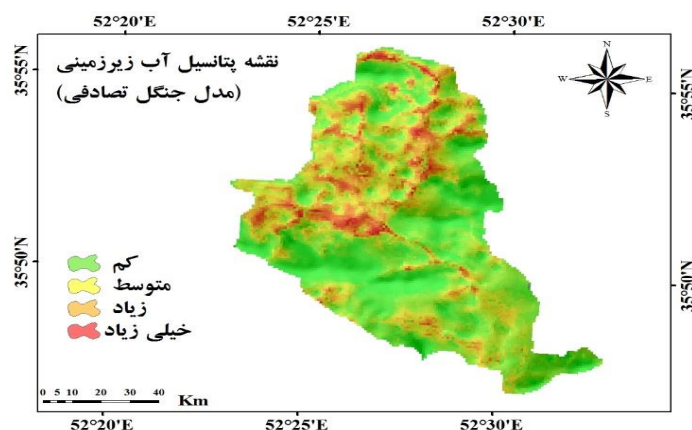
عامل	طبقه	تعداد پیکسل	درصد پیکسل	تعداد چشمه	درصد چشمه	نسبت فراوانی
جهت شیب	مسطح	۷۶۸۷	۱۴/۴۷	۱۴	۹/۳۳	۰/۶۴
	شمال	۸۱۹۹	۱۵/۴۴	۲۸	۱۸/۶۷	۱/۲۱
	شمال شرق	۵۵۷۹	۱۰/۵۰	۲۱	۱۴/۰۰	۱/۳۳
	شرق	۸۱۴۴	۱۵/۳۳	۲۵	۱۶/۶۷	۱/۰۹
	جنوب شرق	۱۰۷۹۳	۲۰/۳۲	۱۷	۱۱/۳۳	۰/۵۶
	جنوب	۵۵۲۹	۱۰/۴۱	۱۴	۹/۳۳	۰/۹۰
شاخص همگرایی	جنوب غرب	۳۲۹۱	۶/۲۰	۱۶	۱۰/۶۷	۱/۷۲
	غرب	۳۸۹۳	۷/۳۳	۱۵	۱۰/۰۰	۱/۳۶
	شمال غرب	۳۶۸۷	۶/۹۴	۲۶	۱۷/۳۳	۲/۵۰
	-۱۲/۵۶ -۸۶/۹۲	۴۵۸۶۹	۸۶/۳۶	۱۲۲	۸۱/۳۳	۰/۹۴
	-۱۱/۹۹ -۱۲/۵۶	۳۵۵۹	۶/۷۰	۲	۱/۳۳	۰/۲۰
	۱۱/۹۹-۸۲/۳۶	۱۰۲۵۵	۱۹/۳۱	۲۳	۱۵/۳۳	۰/۷۹
طبقات ارتفاعی	۲۳۷۲/۵۵-۲۳۷۲/۴۴	۱۲۳۴۸	۲۳/۲۵	۲۵	۱۶/۶۷	۰/۷۲
	۲۷۲۴/۴۴-۳۰۵۵/۶۳	۱۱۸۰۸	۲۲/۲۳	۴۸	۳۲/۰۰	۱/۴۴
	۳۰۵۵/۶۳-۳۳۶۶/۱۲	۱۲۳۸۳	۲۳/۳۱	۴۸	۳۲/۰۰	۱/۳۷
	۳۳۶۶/۱۲-۳۸۲۴/۴۱	۶۳۲۱	۱۱/۹۰	۶	۴/۰۰	۰/۳۴
	۳۸۲۴/۴۱-۴۲۹۰/۸۸	۲۹۲۰۲	۵۴/۹۸	۷۶	۵۰/۶۷	۰/۹۲
	۴۲۹۰/۸۸-۵۰۲/۵۸	۱۳۹۲۹	۲۶/۲۲	۵۷	۳۸/۰۰	۱/۴۵
فاصله از گسل	۵۰۲/۵۸-۸۲۴/۵۵	۶۴۲۳	۱۲/۰۹	۸	۵/۳۳	۰/۴۴
	۸۲۴/۵۵-۱۲۱۷/۲۰	۲۰۵۰	۳/۸۶	۵	۳/۳۳	۰/۸۶
	۱۲۱۷/۲۰-۲۰۰۲/۴۹	۱۵۱۱	۲/۸۴	۴	۲/۶۷	۰/۹۴
	۲۰۰۲/۴۹-۱۵۵۷/۳	۱۵۵۷۳	۲۹/۳۲	۸۴	۵۶/۰۰	۱/۹۱
	۱۵۵۷/۳-۱۱۱۳/۵۳	۱۳۹۰۲	۲۶/۱۷	۲۶	۱۷/۳۳	۰/۶۶
	۱۱۱۳/۵۳-۱۷۳۸/۹۳	۱۱۰۰۲	۲۰/۷۱	۲۴	۱۶/۰۰	۰/۷۷
فاصله از رودخانه	۱۷۳۸/۹۳-۲۴۸۶/۳۷	۷۶۹۰	۱۴/۴۸	۱۰	۶/۶۷	۰/۴۶
	۲۴۸۶/۳۷-۳۸۸۹/۷۲	۴۹۴۸	۹/۳۲	۶	۴/۰۰	۰/۴۳
	۳۸۸۹/۷۲-۴۳۲۴/۵	۲۳۲۴۵	۴۳/۷۶	۶۱	۴۰/۶۷	۰/۹۳
	۴۳۲۴/۵-۱۴۴۲/۸	۱۴۴۲۸	۲۷/۱۶	۵۴	۳۶/۰۰	۱/۳۳
	۱۴۴۲/۸-۱۰۰۹/۷	۱۰۰۹۷	۱۹/۰۱	۲۸	۱۸/۶۷	۰/۹۸
	۱۰۰۹/۷-۵۳۴۵/۷	۵۳۴۵	۱۰/۰۶	۷	۴/۶۷	۰/۴۶
تراکم گسل	۵۳۴۵/۷-۱۸۵۲۳	۱۸۵۲۳	۳۴/۸۷	۶۴	۴۲/۶۷	۱/۲۲
	۱۸۵۲۳-۱۶۶۰۵	۱۶۶۰۵	۳۱/۲۶	۴۸	۳۲/۰۰	۱/۰۲
	۱۶۶۰۵-۱۲۶۰۸	۱۲۶۰۸	۲۳/۷۴	۳۳	۲۲/۰۰	۰/۹۳
	۱۲۶۰۸-۵۳۷۹	۵۳۷۹	۱۰/۱۳	۵	۳/۳۳	۰/۳۳
	۵۳۷۹-۳۵۳۷	۳۵۳۷	۶/۶۶	۱۷	۱۱/۳۳	۱/۰۰
	۳۵۳۷-۴۵۴۹۹	۴۵۴۹۹	۸۵/۶۶	۱۳۰	۸۵/۶۷	۱/۰۱
انحنای طولی	محدب	۴۰۷۹	۷/۶۸	۳	۲/۰۰	۰/۲۶
	مسطح	۵۸۲۷	۱۰/۹۷	۴	۲/۶۷	۰/۲۴
	مقعر	۴۱۴۸۵	۷۸/۱۰	۱۱۶	۷۷/۳۳	۰/۹۹
	محدب	۵۸۰۳	۱۰/۹۳	۳۰	۲۰/۰۰	۱/۸۳
	مسطح	۳۵۳۴۴	۶۶/۵۴	۴۴	۲۹/۳۳	۰/۶۸
	مقعر	۵۹۱۵	۱۱/۱۴	۲۴	۱۶/۰۰	۱/۴۴
تراکم زهکشی	۰/۱۸-۰/۵۲	۸۹۴۰	۱۶/۸۳	۶۸	۴۵/۳۳	۱/۷۴
	۰/۵۲-۰/۸۹	۲۹۱۶	۵/۴۹	۱۴	۹/۳۳	۱/۷۰
	۰/۸۹-۱/۶۱	۱۷۸۵۲	۳۳/۶۱	۷۵	۵۰/۰۰	۱/۴۹
	۰/۶۱-۰/۹۲	۱۱۹۴۲	۲۲/۴۸	۲۶	۱۷/۳۳	۰/۷۷
	۰/۹۲-۰/۷۶	۱۱۱۰۸	۲۰/۹۱	۴۲	۲۸/۰۰	۱/۳۴
	۰/۷۶-۱/۰۰	۱۲۲۱۳	۲۲/۹۹	۷	۴/۶۷	۰/۲۰
درصد شیب	۰/۰۶-۱/۵۵	۱۲۸۱۳	۲۴/۱۳	۴۸	۳۲/۰۰	۱/۳۳
	۱/۵۵-۲۲/۷۲	۱۳۳۹۷	۲۵/۲۲	۴۶	۳۰/۶۷	۱/۲۲
	۲۲/۷۲-۳۳/۸۹	۱۳۱۶۵	۲۴/۷۹	۴۲	۲۸/۰۰	۱/۱۳
	۳۳/۸۹-۴۶/۶۲	۹۴۹۰	۱۷/۸۷	۱۲	۸/۰۰	۰/۴۵
	۴۶/۶۲-۷۹/۲۰	۴۲۴۵	۷/۹۹	۲	۱/۳۳	۰/۱۷
	۷۹/۲۰-۹/۱۸	۱۵۱۹۸	۲۸/۶۱	۵۸	۳۸/۶۷	۱/۳۵
ضریب زبری سطح	۹/۱۸-۱۷/۷۵	۱۶۴۶۰	۳۰/۹۹	۵۳	۳۵/۳۳	۱/۱۴
	۱۷/۷۵-۲۷/۱۲	۱۴۵۹۳	۲۷/۴۷	۳۵	۲۳/۳۳	۰/۸۵
	۲۷/۱۲-۵۱/۰۳	۶۸۶۴	۱۲/۹۲	۴	۲/۶۷	۰/۲۱
	۵۱/۰۳-۹/۷۳	۲۴۸۴۱	۴۶/۷۷	۳۸	۲۵/۳۳	۰/۵۴
	۹/۷۳-۱۱/۵۸	۱۹۴۳۱	۳۶/۵۸	۶۴	۴۲/۶۷	۱/۱۷
	۱۱/۵۸-۱۴/۸۷	۶۸۴۲	۱۲/۸۸	۳۳	۲۲/۰۰	۱/۷۱
شاخص رطوبت توپوگرافی	۱۴/۸۷-۲۲/۵۸	۲۰۰۱	۳/۷۷	۱۵	۱۰/۰۰	۲/۶۵
	۲۰۰۱-۲۹۴۲	۲۹۴۲	۵/۵۴	۱۰	۶/۶۷	۱/۲۰
	۲۹۴۲-۴۹۰۳۳	۴۹۰۳۳	۹۲/۳۱	۱۴۰	۹۳/۳۳	۱/۰۱
	۴۹۰۳۳-۲۰۹۵۱	۲۰۹۵۱	۳۹/۴۴	۵۸	۳۸/۶۷	۰/۹۸
	۲۰۹۵۱-۹۳۹۳	۹۳۹۳	۱۷/۶۸	۲۳	۱۵/۳۳	۰/۸۷
	۹۳۹۳-۲۲۰۵	۲۲۰۵	۴/۱۵	۱۸	۱۲/۰۰	۲/۸۹
سنگ شناسی	۲۲۰۵-۱۴۱۸۸	۱۴۱۸۸	۲۶/۷۱	۳۶	۲۴/۰۰	۰/۹۰
	۱۴۱۸۸-۱۳۶۴	۱۳۶۴	۲/۵۷	۴	۲/۶۷	۱/۰۴
	۱۳۶۴-۶۴۶	۶۴۶	۱/۲۲	۰	۰/۰۰	۰/۰۰
	۶۴۶-۱۹۹۸	۱۹۹۸	۳/۷۶	۱۱	۷/۳۳	۱/۹۵
	۱۹۹۸-۴۰۴۰۲	۴۰۴۰۲	۷۶/۰۷	۱۱۷	۷۸	۱/۰۳
	۴۰۴۰۲-۱۲۷۱۳	۱۲۷۱۳	۲۳/۹۳	۳۳	۲۲	۰/۹۲
خاک	لومی رسی سیلتی					
	لومی سیلتی					

جدول ۲ فراوانی نسبی مربوط به هر کلاس را در مدل‌های جنگل تصادفی و نسبت فراوانی نشان می‌دهد. همانطوری که مشاهده می‌شود نتایج دو مدل مذکور با یکدیگر تفاوت‌هایی دارند. به گونه‌ای که براساس مدل نسبت فراوانی، ۱۴ درصد از منطقه مطالعاتی در کلاس پتانسیل کم و براساس مدل جنگل تصادفی ۴۰ درصد از منطقه در این طبقه قرار گرفته است. همچنین براساس مدل‌های نسبت فراوانی و جنگل تصادفی به ترتیب حدود ۵۰/۲ و ۳۴/۵۸ درصد از منطقه مطالعاتی در طبقه پتانسیل زیاد و خیلی زیاد قرار گرفته‌اند. دلیل این امر را می‌توان در ساختار مدل‌های استفاده شده و عدم قطعیت‌های موجود در مدلسازی عنوان کرد. ویژگی‌های آماری پیش‌بینی احتمالی هر دو نقشه در جدول ۳ ارائه شده است.

پس از محاسبه مقدار نسبت فراوانی برای هر طبقه از عوامل موثر، در نهایت نقشه نهایی پتانسیل آب زیرزمینی منطقه مطالعاتی در محیط سیستم اطلاعات جغرافیایی تهیه شد. در این روش هرچه ضریب نسبت فراوانی به‌دست آمده برای هر یک از طبقات یک عامل بیشتر از یک باشد، نشانه تاثیر بیشتر آن طبقه در وقوع آب زیرزمینی و پتانسیل ایجاد چشمه در نقشه‌ی پیش‌بینی شده است (Shirzadi *et al.*, 2012). همانطوری که شکل ۳ نشان می‌دهد نقشه‌ی پیش‌بینی شده در کلاس‌های پتانسیل ضعیف، متوسط، خوب و خیلی خوب طبقه‌بندی گردید. نقشه پتانسیل آب زیرزمینی به‌دست آمده با استفاده از مدل جنگل تصادفی در محیط نرم‌افزار R در شکل ۴ ارائه شده است.



شکل ۳- نقشه پتانسیل آب زیرزمینی منطقه مورد مطالعه بر اساس مدل نسبت فراوانی



شکل ۴- نقشه پتانسیل آب زیرزمینی منطقه مورد مطالعه بر اساس مدل جنگل تصادفی

جدول ۲- فراوانی نسبی کلاس‌های پتانسیل آب زیرزمینی

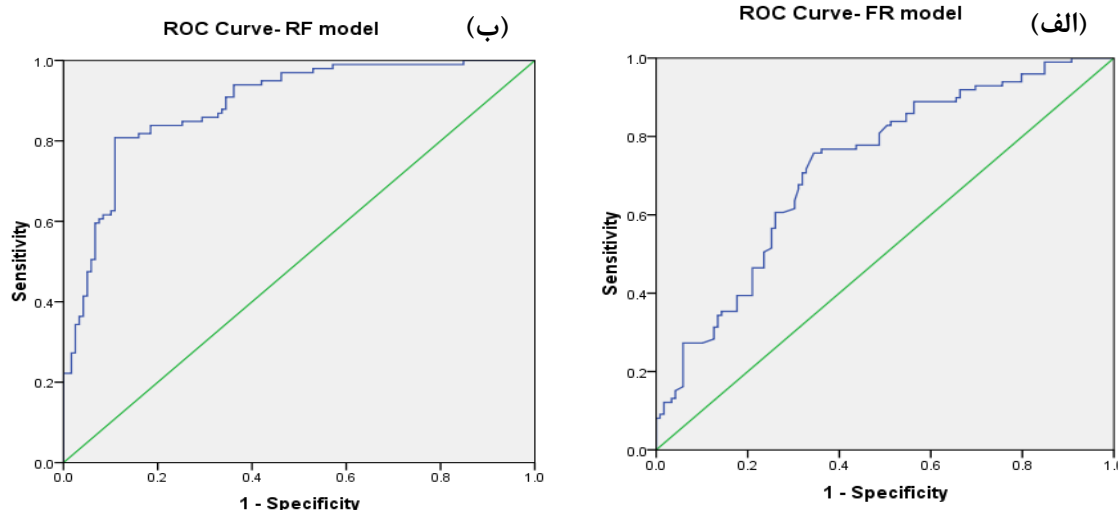
فراوانی نسبی کلاس‌ها (درصد)				
	کم	متوسط	زیاد	خیلی زیاد
مدل جنگل تصادفی	۴۰/۵۳	۲۴/۸۸	۲۱/۲۱	۱۳/۳۷
مدل نسبت فراوانی	۱۴/۰۷	۳۳/۱۲	۳۲/۲۴	۲۰/۵۶

جدول ۳- ویژگی‌های آماری پیش‌بینی احتمالی

مدل‌ها	متغیرهای آماری	
	میانگین	انحراف استاندارد St.dev
RF	۰/۲۴۷	۰/۲۱۵
FR	۰/۵۱۵	۰/۱۶۱

خوب، ۰/۷-۰/۶ متوسط و ۰/۵-۰/۶ ضعیف، را برای همبستگی کیفی و کمی سطح زیر منحنی و ارزیابی تخمین ارائه کرده‌اند (Devkota *et al.*, 2013) دارند. نتایج نشان داد که مدل جنگل تصادفی با دقت ۰/۸۸ دارای کارایی بسیار خوب و مدل نسبت فراوانی با دقت ۰/۷۲ دارای کارایی خوبی برای تهیه نقشه‌ی پتانسیل آب زیرزمینی دارند. که با نتایج Pourtaghi and Pourghasemi (2014) و Zabihi *et al.* (2015) همخوانی دارد. بررسی دقت مدل‌های مذکور در شکل ۵ و جدول ۴ ارائه شده است.

ارزیابی مدل‌های نسبت فراوانی (FR) و جنگل تصادفی (RF) برای بررسی دقت مدل‌های استفاده شده، از منحنی ROC استفاده شد. قابلیت پیش‌بینی مدل، توسط سطح زیر منحنی (AUC) مشخص گردید (Felicísimo *et al.*, 2013). در پژوهش حاضر ۳۰ درصد از چشمه‌ها برای مرحله‌ی ارزیابی استفاده شد. بر اساس منابع علمی در زمینه تحلیل روش منحنی ROC و اعتبارسنجی مدل‌های پیش‌بینی مکانی می‌توان بیان کرد به‌طور کلی، تقسیم ۱-۰/۹ عالی، ۰/۸-۰/۹ بسیار خوب، ۰/۷-۰/۸



شکل ۵- منحنی ROC و سطح زیرمنحنی مربوط به نقشه پتانسیل آب زیرزمینی با استفاده از: الف) مدل جنگل تصادفی و ب) مدل نسبت فراوانی

جدول ۴- خطای استاندارد و سطح زیرمنحنی نقشه‌های پتانسیل آب زیرزمینی

مدل مورد استفاده	سطح زیرمنحنی	خطای استاندارد
نسبت فراوانی	۰/۷۲۴	۰/۰۳۴
جنگل تصادفی	۰/۸۸۹	۰/۰۲۲

## نتیجه‌گیری

با توجه به هزینه‌بر بودن آزمایش‌های اکتشافی منابع آب زیرزمینی، تهیه نقشه‌ی پتانسیل چشمه‌های آب زیرزمینی با استفاده از روش‌های نوین با هدف حفاظت، مدیریت و بهره‌برداری صحیح از منابع آب امری حائز اهمیت است. در پژوهش حاضر نقشه پتانسیل چشمه‌های آب زیرزمینی با استفاده از یک مدل مشهور یادگیری ماشین (جنگل تصادفی) و یک مدل آماری (مدل نسبت فراوانی) تهیه شد، و کارایی این روشها در حوزه آبخیز

عمرک، استان تهران مقایسه شد. بعد از شناسایی و تهیه‌ی عوامل مؤثر در ظهور چشمه‌ها، با استفاده از روش شکستن طبیعی، درصد و فراوانی پتانسیل آب زیرزمینی در هر طبقه با استفاده از هم‌پوشانی نقشه پراکنش چشمه‌ها با هر کدام از لایه‌ها به دست آمد. برای ارزیابی عملکرد مدل‌های مذکور از منحنی مشخصه عملکرد نسبی (ROC) استفاده شد و میزان مساحت زیر منحنی (AUC) مدل‌های جنگل تصادفی و نسبت فراوانی به ترتیب ۸۸ و ۷۲ درصد به دست آمد. نتایج نشان داد که هر دو

از آنها در منطقه مطالعاتی را می‌طلبد. در نهایت اجرای مطالعات بیشتر در منطقه عمرک و حتی مناطق دیگر با استفاده از سایر مدل‌های آماری و داده‌کاوی توصیه می‌شود.

### سپاس‌گزاری

بدین‌وسیله نویسندگان این مقاله مراتب قدردانی و تشکر خود را از معاونت محترم پژوهشی دانشگاه آزاد اسلامی واحد کرمانشاه برای همکاری و حمایت از انجام این تحقیق مستقل بیان می‌دارند.  
"هیچ‌گونه تعارض منافع بین نویسندگان وجود ندارد"

### REFERENCES

- Chen, W., Li, Y., Xue, W., Shahabi, H., Li, S., Hong, H., & Ahmad, B.B. (2020). Modeling flood susceptibility using data-driven approaches of naïve bayes tree, alternating decision tree, and random forest methods. *Science of The Total Environment*, 701, 134-979.
- Daoud, J.I. (2017). *Multicollinearity and regression analysis*. J. Phy, Conference Series, 949(1), pp. 012009). IOP Publishing.
- Devkota, K.C., Regmi, A.D., Pourghasemi, H.R., Yoshida, K., Pradhan, B., and Ryu, I.C. (2013). Landslide susceptibility mapping using certainty factor, index of entropy and logistic regression models in GIS and their comparison at Mugling-Narayanghat road section in Nepal Himalaya. *Natural Hazards*, 65(1), 135-165.
- Feliciísimo, Á., Cuartero, A., Remondo, J., and Quirós, E. (2013). Mapping landslide susceptibility with logistic regression, multiple adaptive regression splines, classification and regression trees, and maximum entropy methods: a comparative study. *Landslides*, 10, 175-189.
- Erfanian, M., Farajollahi, H., Souri, M., and Shirzadi, A. (2016). Comparing the Efficiency of Weight of Evidence, Logistic Regression and Frequency Ratio Methods for Mapping Groundwater Spring Potential in Ghelgazi Watershed, Kordestan Province of Iran. *Journal of Water and Soil Science*, 20 (75), 59-72. (In Farsi)
- Ghosh, S., and Carranza, E.J.M. (2010). Spatial analysis of mutual fault/fracture and slope controls on rock sliding in Darjeeling Himalaya, India. *Geomorphology* 122, 1- 24.
- Greene, W.H. (2000). *Econometric Analysis* (4th Edition), Upper Saddle River, NJ: Prentice Hall.
- Jha, M. K., Chowdhury, A., Chowdary, V. M., and Peiffer, S. (2007). Groundwater management and development by integrated remote sensing and geographic information systems: prospects and constraints. *Water Resources Management*, 21(2), 427-467.
- Jafarian, Z., Kargar, M., and Bahreini, Z. (2019). Which spatial distribution model best predicts the occurrence of dominant species in semi-arid rangeland of northern Iran. *Ecological informatics*, 50, 33-42.
- Kouhi Kamali, M., and Rajabi, M. (2010). *Determining the impact of urban green space on the value of artificial use*. *Journal of Application of GIS and RS in Planning*, 1(1), 31-23. (in Farsi).
- Lee, S., and Pradhan, B. (2007). Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models. *Landslides*, 4, 33-41.
- Manap, M.A., Nampak, H., Pradhan, B., Lee, S., Soleiman, 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.
- 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(2), 913-929.
- Naghbi, 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(1), 171-186.
- Nguyen, P.T., Ha, D.H., Avand, M., Jaafari, A., Nguyen, H.D., Al-Ansari, N., Van Phong, T., Sharma, R., Kumar, R., Le, H.V., Ho, L.S., Prakash, I., Pham, B.T. (2020). *Soft Computing Ensemble Models Based on Logistic Regression for Groundwater Potential Mapping*. *Applied Sciences*, 10(7), 2469.
- Oh, H.J., Kim, Y.S., Choi, J.K., Park, E., and Lee, S. (2011). *GIS mapping of regional probabilistic groundwater potential in the area of Pohang City, Korea*. *Journal of Hydrology*, 399(3-4), 158-172.
- Ozdemir, A. (2011). GIS-based groundwater spring potential mapping in the Sultan Mountains (Konya, Turkey) using frequency ratio, weights of evidence and logistic regression methods and their comparison. *Journal of Hydrology*, 411, 290-308.
- 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.

روش دارای دقت قابل قبولی برای تهیه نقشه پتانسیل چشمه آب زیرزمینی در منطقه مورد مطالعه هستند. اما به عنوان پیشنهاد می‌توان بر نتایج مدل جنگل تصادفی با توجه به مقدار مساحت زیر منحنی بیشتر برای شناسایی و پهنه‌بندی پتانسیل چشمه‌های آب زیرزمینی تأکید کرد. از آنجایی که تاکنون در این منطقه مطالعه‌ای مشابه پژوهش حاضر انجام نشده، نتایج این پژوهش می‌تواند بیانگر وضعیت پتانسیل منابع آب زیرزمینی حوزه آبخیز عمرک در تهران باشد. نتایج حاصل، توجه مدیران و مسئولان امر در اتخاذ تصمیمات آتی مدیریت منابع آب و همچنین بهره‌برداری

- Pradhan, B. (2009). Groundwater potential zonation for basaltic watersheds using satellite remote sensing data and GIS techniques, *Central European Journal of Geosciences*, 1(1), 120-129.
- Rangzan, K., Charchi, E., and Amiri, F. (2005). Comparison of the effect of structural-georphological factors on the emergence of springs in the area of Qala-e-Razeh-Bidrobeh with the Pabdeh anticline in the northwest and northeast of Khuzestan. *Proceedings of the Ninth Conference of the Geological Society of Iran*, Tehran Teacher Training University, 539-526. (in Farsi).
- Seyed ali, S., Rahimi, M., Dastourani, j., and Khosroshahi, M.(2016). Trend Analysis of Hydroclimatological Parameters and Detection of Manageral Changes in Water Resources Conditions of Hablerood Watershed. *Iranian Journal of Range and Desert Research*, 23, 555-556 (in Farsi).
- Servati, M.R., Ghahrodi, T.M., Golkarami, A., and Njafi, E. (2014). *Geomorphological thresholds for gully erosion in Kchick watershed, NE Golestan Province*. *Applied researches in geographical sciences*, 32, 231-249. (in Farsi).
- Shataee, S., Weinaker, H., and Babanejad, M. (2011). Plot- level Forest Volume Estimation Using Airborne Laser Scanner and TM Data, Comparison of Boosting and Random Forest Regression Algorithms. *Procedia Environmental Sciences*, 7: 68-73.
- Shirzadi, A., Saro, L., Hyun-Joo, O.H., and Chapi, K. (2012). A GIS-based logistic regression model in rock fall susceptibility mapping along a mountainous road: Salavat Abad case study, Kurdistan, Iran. *Natural Hazard*. 64, 1639-1656.
- Zabihi, M., Pourghasemi, H., and Behzadfar, M. (2015). *Groundwater Potential Mapping using Shannon's Entropy and Random Forest Models in the Bojnourd Township*. *Iranian journal of Ecohydrology*, 2(2), 221-232. (In Farsi)
- Zandi, S., Soleimani, K., and Zandi, J. (2016). Mapping Groundwater Spring Occurrence Potential using Logistic Regression Statistical Method (Case Study: Mireddeh Mountainous Watershed, Kurdistan). *Journal of Watershed Management Research*, 6 (12), 75-87. (In Farsi)
- Zare Chahouki, M., Khalsi Ahvazi, L., and Azarnivand, H. (2014). *Plant Species Distribution Modeling Using Logistic Regression Models in the North East of Semnan*. *Journal of Range and Watershed Management (Iranian Journal of Natural Resources)*, 67(1), 45-59. (In Farsi)