

## Comparison of Standard Drastic and Nonparametric Models Instance-Based Learning with parameter K (IBK) and the Tree Decision M5 in Determination of Groundwater Pollution Potential (Case study: Kuchesfahan- Astane plain)

### Samira Rahnama

Ph. D Student of Water Resources Engineering, Water Engineering Department, College of Agriculture, University of Birjand, Birjand, Iran.

### Hossein Khozeymehnezhad

\*Associate professor, Water Engineering Department, College of Agriculture, University of Birjand, Birjand, Iran. (Corresponding author):

Email: Hkhozeymeh@birjand.ac.ir.

### Abbas KhasheiSiuki

Associate professor, Water Engineering Department, College of Agriculture, University of Birjand, Birjand, Iran.

Received: 2019/10/13

Accepted: 2020/01/02

**Document Type:** Research article

### ABSTRACT

**Background and Aim:** Due to the increasing demands of the human population to groundwater, protection and prevention of these water resources from pollution are necessary. The purpose of this study was to evaluate the vulnerability of groundwater aquifer in Kuchesfahan- Astane plain located in Gilan province using DRASTIC method and nonparametric models.

**Materials and methods:** In this study, seven layers were prepared for parameters in GIS software, and after weighting and combining standard ranks, the groundwater vulnerability maps for the study area were prepared. Nitrate data were used to validate the model in this region. Subsequently, by using the nonparametric models, Instance-Based Learning with parameter K (IBK) and the Tree Decision M5, the amount of nitrate was estimated. Meanwhile, Gamma test was conducted to find the best combination of input parameters.

**Results:** The results revealed that the vulnerability of groundwater aquifer in this plain has 4 classes including 18.56 % in low vulnerability, 51.29 % in low to medium vulnerability, 28.46% in medium to high vulnerability, and 1.67% in high vulnerability classes. Also, the results showed that both of the nonparametric models have suitable estimates of the nitrate content, but the  $M_5$  decision tree model yielded the best results ( $R^2=0.98$ ).

**Conclusion:** The results showed that nonparametric models are efficient method to estimate the aquifer vulnerability and provide accurate results to estimate the potential of contamination in the study area. This demonstrates the superiority of the  $M_5$  model over other aquatic vulnerability assessment methods.

**Keywords:** Decision tree, Geographic Information System, Groundwater, Population

► **Citation:** Rahnama S, Khozeymehnezhad H, KhasheiSiuki A. Comparison of Standard Drastic and Nonparametric Models Instance-Based Learning with parameter K (IBK) and the Tree Decision M5 in Determination of Groundwater Pollution Potential (Case study: Kuchesfahan- Astane plain). *Iranian Journal of Research in Environmental Health*. Winter 2020;5 (4): 315-329.

## مقایسه روش دراستیک استاندارد و مدل‌های ناپارامتریک یادگیری بر پایه نمونه با پارامتر (IBK) K و درخت تصمیم M5 در مکان‌یابی پتانسیل آلودگی آب‌های زیرزمینی (مطالعه موردی: دشت آستانه - کوچصفهان)

### چکیده

**زمینه و هدف:** با توجه به نیاز روزافزون جوامع بشری به منابع آب زیرزمینی، حفاظت و جلوگیری از آلودگی این منابع امری ضروری تلقی می‌گردد. مطالعه حاضر با هدف ارزیابی آسیب‌پذیری آبخوان آب زیرزمینی آبخوان دشت آستانه - کوچصفهان استان گیلان، با استفاده از روش دراستیک و مدل‌های ناپارامتریک انجام شد.

**مواد و روش‌ها:** در این پژوهش، پارامترها به صورت ۷ لایه در محیط نرم‌افزار سیستم اطلاعات جغرافیایی (GIS) برای دشت تهیه و پس از وزن‌دهی و ترکیب رتبه‌های استاندارد، نقشه آسیب‌پذیری آب‌های زیرزمینی دشت با استفاده از روش دراستیک تعیین گردید. برای صحت‌سنجی مدل، از داده‌های نیترا در منطقه استفاده شد. سپس با کمک مدل‌های ناپارامتریک یادگیری بر پایه نمونه با پارامتر K و درخت تصمیم M5 مقدار نیترا تخمین زده شد. همچنین آزمون گاما برای یافتن بهترین ترکیب پارامترهای ورودی اجرا گردید.

**یافته‌ها:** بر اساس نتایج این پژوهش، آسیب‌پذیری آبخوان دشت کوچصفهان در ۱۸/۵۶٪ دارای آسیب‌پذیری اندک، ۵۱/۲۹٪ دارای آسیب‌پذیری اندک تا متوسط، ۲۸/۴۶٪ دارای آسیب‌پذیری متوسط تا زیاد و ۱/۶۷٪ دارای آسیب‌پذیری زیاد می‌باشد. همچنین هر دو مدل ناپارامتریک به کار گرفته شده تخمین مناسبی از مقدار نیترا می‌دهند، اما مدل M5 بهترین نتایج را دربرداشت ( $R^2 = 0.98$ ).

**نتیجه‌گیری:** مدل‌های ناپارامتریک، روشی کارا در تخمین آسیب‌پذیری آبخوان محسوب می‌شوند و نتایج دقیقی از برآورد پتانسیل آلودگی در منطقه می‌دهند. این نکته برتری مدل M5 نسبت به سایر روش‌های مورد بررسی در آسیب‌پذیری آبخوان را نشان می‌دهد.

**کلید واژه‌ها:** آب‌زیرزمینی، آلودگی، درخت تصمیم، سیستم اطلاعات جغرافیایی.

سمیرا رهنما

دانشجوی دکتری منابع آب، گروه علوم و مهندسی آب، دانشکده کشاورزی، دانشگاه بیرجند، بیرجند، ایران.

حسین خزیمه‌نژاد

\* دانشیار، گروه علوم و مهندسی آب، دانشکده کشاورزی، دانشگاه بیرجند، بیرجند، ایران.  
(نویسنده مسئول): پست الکترونیک

Hkhozeymeh@birjand.ac.ir

عباس خاشعی سیوکی

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

تاریخ دریافت: ۱۳۹۸/۰۷/۲۱

تاریخ پذیرش: ۱۳۹۸/۱۰/۱۳

نوع مقاله: مقاله پژوهشی

◀ **استناد:** رهنما س، خزیمه‌نژاد ح، خاشعی سیوکی ع. مقایسه روش دراستیک استاندارد و مدل‌های ناپارامتریک یادگیری بر پایه نمونه با پارامتر (IBK) K و درخت تصمیم M5 در مکان‌یابی پتانسیل آلودگی آب‌های زیرزمینی (مطالعه موردی: دشت آستانه - کوچصفهان). *فصلنامه پژوهش در بهداشت محیط*. زمستان ۱۳۹۸؛ ۵(۴): ۳۱۵-۳۲۹.

افزایش جمعیت، توسعه صنعت و کشاورزی در دهه‌های اخیر باعث استفاده روزافزون از مواد شیمیایی، کودها و افزایش ضایعات صنعتی در بخش‌های مختلف و در نتیجه آلودگی منابع آب خصوصاً آب‌های زیرزمینی شده است. اعمال یک مدیریت مناسب بر منابع آب با تعیین مناطق آسیب‌پذیر به عنوان اولین راهکار می‌تواند مفید باشد. مفهوم آسیب‌پذیری برای اولین بار در اواخر سال ۱۹۶۰ میلادی در فرانسه برای آگاهی در مورد آلودگی آب زیرزمینی ارائه شده است (۱). تاکنون تعاریف زیادی در مورد آسیب‌پذیری آبخوان و مفهوم آن در هیدروژئولوژی ارائه شده است که از آن جمله می‌توان به تعریف کمیته ملی علوم آمریکا در سال ۱۹۹۳ اشاره کرد. این کمیته، آسیب‌پذیری آب‌های زیرزمینی نسبت به آلودگی را، تمایل یا احتمال رسیدن آلاینده‌ها به یک مکان مشخص در سیستم آب زیرزمینی بعد از به وجود آمدن آن‌ها در برخی محل‌ها در بالای سطح آبخوان می‌داند. روش‌های مختلفی برای بررسی و ارزیابی پتانسیل آلودگی در یک آبخوان وجود دارد که از آن جمله می‌توان به روش‌های مبتنی بر شبیه‌سازی، روش‌های آماری و روش‌های شاخص و همپوشان اشاره کرد (۲). دراستیک<sup>۱</sup> یک روش شاخص و همپوشان است که در آن اطلاعات حاصل از پارامترهای مختلف به صورت تلفیقی و به‌طور موازی مورد تجزیه و تحلیل قرار می‌گیرند. اصولاً روش‌های شاخص و همپوشان از ترکیب پارامترهای هیدروژئولوژیک مؤثر در انتقال آلودگی به آب‌های زیرزمینی تشکیل می‌شوند. در این روش‌ها برای تعیین اهمیت نسبی، هر کدام از پارامترها نسبت به سایرین ارزیابی می‌شوند. اگرچه این روش‌ها فرآیندهای مؤثر در آلوده شدن منابع آب زیرزمینی را کاملاً شرح نمی‌دهند، اما به دلیل ساده بودن روش و در دسترس بودن داده‌های مورد نیاز، از اهمیت ویژه‌ای برخوردارند (۳).

در پژوهشی از مدل دراستیک جهت ارزیابی آسیب‌پذیری آبخوان حوضه گوموشاسیکوی<sup>۲</sup> استفاده کردند. نتایج آن‌ها نشان

1. DRASTIC
2. Gumushacikoy

داد که منطقه مورد مطالعه در سه محدوده پتانسیل آسیب‌پذیری پایین، متوسط و زیاد قرار می‌گیرد (۴). ارزیابی ریسک آب زیرزمینی دشت شیراز بر اساس بهینه‌سازی روش دراستیک انجام شد. نتایج نشان داد که استفاده از آزمون ویلکاکسون و بهینه‌سازی الگوریتم ژنتیک (GA)<sup>۳</sup> بهتر از دیگر روش‌ها عمل می‌کند و ضریب تبیین نسبت به روش دراستیک به‌طور قابل توجهی افزایش یافت (از ۰/۵۷ به ۰/۸۲) (۵). ارزیابی آسیب‌پذیری آب‌های زیرزمینی دشت قزوین با استفاده از روش دراستیک و به کمک سیستم اطلاعات جغرافیایی (GIS)<sup>۴</sup> نشان داد که مقدار شاخص دراستیک عمومی از ۳۵ تا ۱۰۸ در کل منطقه می‌باشد (۶). در مطالعه امیراحمدی و همکاران که به منظور بررسی آسیب‌پذیری دشت نیشابور با استفاده از روش دراستیک از محیط GIS استفاده کردند، بیشترین درصد پتانسیل آسیب‌پذیری مربوط به طبقه زیاد بود و مناطق جنوبی و غربی دشت از پتانسیل آسیب‌پذیری زیادی برخوردار بوده و نواحی شمالی کمترین پتانسیل را داشتند (۷). در ارزیابی آسیب‌پذیری دشت قائن با استفاده از مدل دراستیک، حدود ۲۹/۶٪ از آبخوان در محدوده پتانسیل آسیب‌پذیری کم، ۳۴/۹۴٪ در محدوده آسیب‌پذیری متوسط و ۳۶/۲٪ واقع در شمال شرقی آبخوان در محدوده آسیب‌پذیری زیاد قرار داشتند (۸). یوسف‌دوست و خاشعی‌سیوکی در پژوهشی نیز به تعیین آسیب‌پذیری آبخوان کوچصفهان با استفاده از روش دراستیک استاندارد و روش‌های داده‌مبنا پرداختند. نتایج آن‌ها برتری روش‌های هوش مصنوعی در بررسی آسیب‌پذیری آبخوان را در مقایسه با روش دراستیک نشان داد (۹). در پژوهشی با توجه به اهمیت دشت نیشابور در تأمین آب کشاورزی، شرب و صنعت منطقه، از الگوریتم‌های تنبیل<sup>۵</sup>، K\*<sup>۶</sup>، KNN<sup>۷</sup> و LWL<sup>۸</sup> و مدل درختی M5<sup>۹</sup> نرم‌افزار وکا<sup>۹</sup>

3. Genetic Algorithm
4. Geographic Information System (GIS)
5. K-Nearest Neighbors Algorithm
6. K-Star Algorithm
7. Locally Weighted Learning
8. Decision M5 Tree
9. Weka (Waikato Environment for Knowledge)

موجود در منطقه (نیترات) می باشد. علاوه بر این، به منظور دستیابی به روشی مناسب و مؤثر برای حفاظت منابع آب زیرزمینی از آلودگی‌هایی که در آینده آن‌ها تهدید می‌کند، آسیب‌پذیری دشت آستانه- کوچصفهان با استفاده از روش دراستیک استاندارد و مدل‌های ناپارامتریک یادگیری بر پایه نمونه  $IBK^2$  و درخت تصمیم M5 نرم‌افزار WEKA تخمین زده شد. در ادامه از میان این روش‌ها، آن مورد که آسیب‌پذیری آبخوان را با دقت بیشتری تخمین می‌زند، در مقام روش برتر انتخاب و با کاهش پارامترهای مدل، مقدار آسیب‌پذیری آبخوان تخمین زده شد.

## روش کار

### منطقه مورد مطالعه

دشت آستانه- کوچصفهان در استان گیلان و منطبق بر مخروط افکنه سفیدرود می‌باشد؛ یعنی محدوده‌ای که تمامی آب‌های سطحی وارد رودخانه سفیدرود یا نهرها و کانال‌های منشعب از آن شده و در نهایت با شیبی کمتر از ۱٪ به دریای خزر ریخته و یا مزارع و شالیزارها را سیراب می‌نماید. منطقه مورد مطالعه در این پژوهش، از جنوب به امامزاده هاشم، از شمال به دلتای سفیدرود و دریای خزر، از شرق به رودخانه شمروود و از غرب به رودخانه سیاه‌رود (زرچوب) منتهی می‌شود. این محدوده بین ۴۰ درجه و ۱۲ دقیقه تا ۵۰ درجه و ۵ دقیقه طول شرقی و ۳۷ درجه و ۷ دقیقه تا ۳۷ درجه و ۲۵ دقیقه عرض شمالی قرار دارد و دارای مساحتی حدود ۱۱۰۰ کیلومتر مربع می‌باشد. شکل ۱ محدوده مورد مطالعه را نشان می‌دهد. بر اساس آمار بارندگی ایستگاه‌های واقع در محدوده مطالعاتی بین سال‌های ۱۳۴۰-۷۸، میزان بارندگی در ایستگاه‌های آستانه، شهر بیجار، سد سنگر و لاهیجان بین ۱۰۰۰-۲۵۰۰ میلی‌متر در سال متغیر است که تغذیه مناسب آبخوان را سبب می‌شود. جنس مواد سازنده آبخوان از رسوبات دلتایی و آبرفتی رودخانه سفیدرود می‌باشد که در بخش‌های جنوبی‌تر، ذرات دانه درشت‌تر هستند و از لحاظ کمی و کیفی، شرایط بهتری را به وجود می‌آورند (۱۵).

تحت هفت سناریوی مختلف، برای برآورد تراز سطح ایستابی این آبخوان استفاده شد. نتایج نشان داد که توانایی مدل KNN با داشتن  $R^2 = 0/96$  از سایر مدل‌ها بیشتر می‌باشد. همچنین مدل LWL مدل مناسبی برای پیش‌بینی تراز سطح ایستابی نمی‌باشد (۱۰). پژوهشی بر روی رودخانه آبی‌چای واقع در دشت تبریز به منظور شناسایی عوامل مؤثر بر آلودگی رودخانه و آبخوان دشت تبریز با بهره‌گیری از روش‌های شبکه عصبی مصنوعی، درخت تصمیم M5 و زمین آمار در طی دوره ۱۳۸۰-۹۲ صورت گرفت. برای شبیه‌سازی مدل درخت تصمیم M5 از نرم‌افزار WEKA استفاده شد. نتایج تحقیق حاکی از آن بود که روش درخت تصمیم M5 روابط ساده خطی، قابل فهم و کاربردی‌تری را برای پیش‌بینی ارائه می‌کند (۱۱). ارزیابی مناطق آسیب‌پذیر در آب‌های زیرزمینی با استفاده از مدل دراستیک و منطق فازی در شهر هرات افغانستان نشان داد که ۵۱٪ از آب‌های زیرزمینی در معرض آلودگی قرار دارد و همچنین مدل پیشنهادی می‌تواند ابزاری مؤثر در مدیریت منابع آب زیرزمینی منطقه واقع شود (۱۲). آسیب‌پذیری ذاتی مخازن آب زیرزمینی دشت ملایر در جنوب همدان، با استفاده از مدل دراستیک نسبت به نیترات انجام شد. نتایج این مطالعه نشان داد که نتایج دراستیک در ارزیابی بهتر از گادز<sup>۱</sup> می‌باشد. همچنین پهنه‌بندی آسیب‌پذیری بین خیلی کم تا بسیار زیاد تعیین شد که نشان‌دهنده افزایش شاخص دراستیک با غلظت نیترات است (۱۳). در مطالعه قنبریان و احمدی‌ندوشن که با هدف ارزیابی آسیب‌پذیری دشت لردگان با استفاده از مدل‌های دراستیک، گادرز و ای وی آی انجام شد، به ترتیب ۱۴٪، ۷۶٪ و ۱۰٪ از مساحت منطقه دارای پتانسیل آسیب‌پذیری متوسط، زیاد و خیلی زیاد بودند. بر اساس مدل ای وی آی، بیش از ۹۵٪ آبخوان در طبقه با حساسیت بالا قرار داشت. بر اساس مدل گادرز، ۴۵/۳٪ از منطقه در کلاس آسیب‌پذیری کم و مابقی در کلاس متوسط قرار داشت (۱۴).

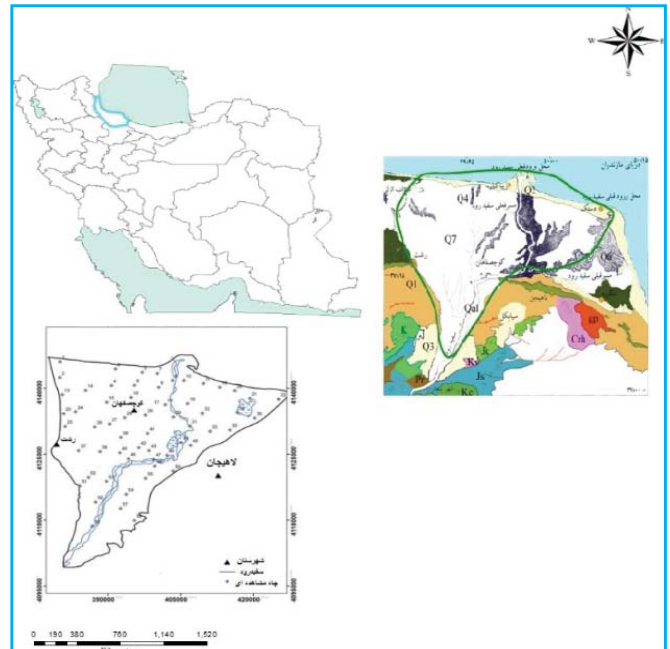
هدف از مطالعه حاضر، افزون بر تهیه نقشه و تعیین مناطقی با پتانسیل بالای آسیب‌پذیری، تطبیق آن با داده‌های آلودگی

آب<sup>۱</sup>، تغذیه خالص<sup>۲</sup>، محیط آبخوان<sup>۳</sup>، بافت خاک<sup>۴</sup>، توپوگرافی<sup>۵</sup>، تأثیر منطقه غیراشباع<sup>۶</sup> و هدایت هیدرولیکی آبخوان<sup>۷</sup> است (۱۶). برای تهیه اطلاعات آماری عمق آب زیرزمینی (D) با استفاده از اطلاعات پیژومترهای موجود در آبخوان، ابتدا نقشه هم‌ارز عمق آب زیرزمینی تهیه و سپس با کمک رتبه‌بندی عمق نقشه، رتبه‌بندی عمق آب زیرزمینی انجام شد. برای تهیه نقشه تغذیه خالص دشت، ابتدا میانگین بارندگی سالیانه<sup>۸</sup>، مقدار نفوذپذیری خاک<sup>۹</sup> و شیب توپوگرافی<sup>۱۱</sup> مناطق متفاوت دشت به دست آمد. این لایه‌ها در GIS با استفاده از دستور UNION همپوشانی شد و سپس نقشه تغذیه خالص به صورت رستری<sup>۱۲</sup> به دست آمد و پس از آن، با توجه به رتبه‌هایی که در محاسبه شاخص آسیب‌پذیری در نظر گرفته شد، برای پارامتر تغذیه و طبق رابطه (۱) نقشه طبقه‌بندی تغذیه به دست آمد.

$$R = \text{Slope}(\%) + R_a(\text{mm}) + I_n(\text{mm}) \quad (1)$$

لایه محیط آبخوان (A) با استفاده از لوگ‌های حفاری چاه‌های محدوده مطالعاتی، نقشه نوع و جنس محیط آبخوان تهیه شد و با توجه به رتبه آن محیط، نقشه رستری رتبه‌بندی آن در محیط GIS به دست آمد. لایه محیط خاک (S) با استفاده از لوگ حفاری چاه‌های پیژومتری، نقشه نوع و جنس خاک تا عمق ۲ متری تهیه و سپس با توجه به رتبه‌بندی این پارامتر، نقشه رستری محیط خاک در محیط GIS حاصل شد. به منظور استخراج لایه شیب، ابتدا نقشه رستری توپوگرافی و سپس به کمک آن، نقشه رستری درصد شیب تهیه شد و در نهایت با استفاده از رتبه‌بندی این نقشه، رتبه‌بندی شیب حاصل گردید. در این دشت، اطلاعات مربوط به محیط غیراشباع از

2. Depth of water (D)
3. Pure Recharge (R)
4. Aquifer media (A)
5. Soil texture (S)
6. Topography (T)
7. Impact of vadose zone (I)
8. Hydraulic Conductivity (C)
9. Rainfall annual (Ra)
10. Soil Infiltration (In)
11. Topography Slope (slope)
12. Raster



شکل ۱. موقعیت منطقه مورد مطالعه

### روش دراستیک استاندارد

مدل دراستیک، یک مدل تجربی و استاندارد جهت ارزیابی آسیب‌پذیری آب زیرزمینی با استفاده از پارامترهای هیدرولوژیکی می‌باشد. این مدل مخفف پارامترهایی است که در یک سیستم هیدرولوژیکی کنترل کننده آلودگی آب زیرزمینی است (۱۶). مدل دراستیک را سازمان حفاظت محیط زیست ایالات متحده آمریکا، با هدف ارزیابی پتانسیل آلودگی آب‌های زیرزمینی، برای سراسر ایالات متحده طراحی کرد (۱۷). این شاخص بدون بُعد و غیرقابل اندازه‌گیری بوده و به خصوصیات آبخوان بستگی دارد. در این روش‌ها برای تعیین اهمیت نسبی، هر کدام از پارامترها نسبت به سایرین ارزیابی می‌شوند. اگرچه این روش‌ها، فرآیندهای مؤثر در آلوده شدن منابع آب زیرزمینی را کاملاً شرح نمی‌دهند، اما به دلیل ساده بودن روش و در دسترس بودن داده‌های مورد نیاز، از اهمیت ویژه‌ای برخوردارند (۳). آلر و همکاران روش دراستیک را ارائه کردند که پس از آن به عنوان روش توصیه شده سازمان محیط زیست آمریکا در سال ۱۹۹۳ ارائه شد. این روش دارای ۷ پارامتر عمق

1. United States Environmental Protection Agency

در نظر گرفته می‌شود. وزن‌های مربوط از ۱ تا ۵ است. به مهم‌ترین آن‌ها وزن ۵ و کم‌اهمیت‌ترین آن‌ها وزن ۱ داده می‌شود (۱۸). بنابراین، شاخص آسیب‌پذیری از مجموع حاصل ضرب وزن و رتبه ۷ پارامتر یاد شده، مطابق رابطه ۲ به دست می‌آید.

(۲)

$$DRASTIC_{Index} = D_r D_w + R_r R_w + A_r A_w + S_r S_w + T_r T_w + I_r I_w + C_r C_w$$

که در آن  $I$  رتبه و  $W$  وزن هر پارامتر می‌باشد. هرچه مقدار شاخص در یک مکان بالاتر باشد، پتانسیل آلودگی در آن محل بیشتر و آسیب‌پذیری آن بالاتر خواهد بود. جدول ۱ به ترتیب رتبه‌بندی و وزن‌دهی عمق سطح آب زیرزمینی، درصد شیب، محیط اشباع آبخوان، محیط غیراشباع آبخوان، محیط خاک، هدایت هیدرولیکی و تغذیه خالص را نشان می‌دهد.

لوگ حفاری چاه‌های پیرومتری به دست آمد. بقیه مراحل همانند اثر بافت خاک تهیه شد. نقشه GIS قابلیت انتقال و ضخامت آبخوان از آب منطقه‌ای استان دریافت شد. به‌طور کلی، با توجه به اینکه در آزمایشات پمپاژ، مقدار پارامتر ضریب قابلیت انتقال آب اندازه‌گیری می‌شود، لذا با استفاده از ضخامت اشباع آبخوان، مقدار هدایت هیدرولیکی از تقسیم ضریب قابلیت انتقال بر ضخامت اشباع به دست آمد. نقشه هدایت هیدرولیکی به صورت رستری به دست آمد و پس از آن، با توجه به رتبه‌هایی که در محاسبه شاخص آسیب‌پذیری در نظر گرفته شد، نقشه رتبه‌بندی حاصل گردید.

در برآورد آسیب‌پذیری آبخوان با استفاده از روش دراستیک، ارقام بین ۱ تا ۱۰ به هر یک از پارامترهای فوق تعلق می‌گیرد، ضمن اینکه بنا بر اهمیت متفاوت پارامترها، برای هر یک وزنی نیز

جدول ۱. رتبه‌بندی و وزن‌دهی پارامترهای هفت‌گانه دراستیک (۱۸)

تغذیه (میلی‌متر بر سال)		هدایت هیدرولیکی (متر بر روز)		محیط خاک		محیط غیراشباع آبخوان		محیط اشباع آبخوان		درصد شیب توپوگرافی		عمق آب زیرزمینی (متر)	
دامنه	رتبه	دامنه	رتبه	دامنه	رتبه	دامنه	رتبه	دامنه	رتبه	دامنه	رتبه	دامنه	رتبه
۰-۵۰/۸	۱	۰/۰۴-۴/۱	۱	نازک/ نبود لایه خاک	۱۰	لایه محبوس	۱	شیل توده‌ای	۲	۰-۲	۱۰	۰-۱/۵	۱۰
۵۰/۸-۱۰۱/۶	۳	۴/۱-۱۲/۳	۲	شن	۱۰	سیلت- رس	۳	آذرین- دگرگونی	۳	۲-۶	۹	۱/۵-۴/۶	۹
۱۰۱/۶-۱۷۷/۸	۶	۱۲/۳-۲۸/۷	۴	ماسه	۹	شیل	۳	آذرین- دگرگونی هوازده	۴	۶-۱۲	۵	۴/۶-۹/۱	۷
۱۷۷/۸-۲۵۴	۸	۲۸/۷-۴۱	۶	کود گیاهی (peat)	۸	سنگ آهک	۳	یخرفت‌ها (رسوبات یخچالی)	۵	۱۲-۱۸	۳	۹/۱-۱۵/۲	۵
> ۲۵۴	۹	۴۱-۸۲	۸	رس فشرده/ متراکم	۷	ماسه‌سنگ	۶	ماسه‌سنگ، سنگ آهک لایه- لایه	۶	> ۱۸	۱	۱۵/۲-۲۲/۸	۳
-	-	> ۸۲	۱۰	لوم ماسه‌ای	۶	سنگ آهک لایه‌لایه، ماسه‌سنگ	۶	ماسه‌سنگ توده‌ای	۶	-	-	۲۲/۸-۳۰/۴	۲
-	-	-	-	لوم	۵	شن و ماسه با مقدار زیادی سیلت و رس	۶	سنگ آهک توده‌ای	۸	-	-	> ۳۰/۴	۱
-	-	-	-	لوم سیلتی	۴	شن و ماسه	۸	شن و ماسه	۸	-	-	-	-
-	-	-	-	لوم رس‌دار	۳	بازالت	۹	بازالت	۹	-	-	-	-
-	-	-	-	کود (muck)	۲	سنگ آهک کارستی	۱۰	سنگ آهک کارستی	۱۰	-	-	-	-
-	-	-	-	رس غیرمتراکم	۱	-	-	-	-	-	-	-	-
وزن: ۴		وزن: ۳		وزن: ۲		وزن: ۵		وزن: ۳		وزن: ۱		وزن: ۵	

با تهیه نقشه‌های رستری پارامترهای دراستیک در محیط GIS و به‌کارگیری قابلیت Calculator Raster نرم‌افزار GIS، می‌توان لایه‌های متفاوت را تلفیق و نقشه پهنه‌بندی آسیب‌پذیری را تهیه کرد. در نهایت، با تلفیق پارامترها طبق رابطه ۲، شاخص

این درخت متظاهر می‌شوند (۲۰). روش درخت تصمیم، یک روش چند مرحله‌ای می‌باشد که در آن به صورت بازگشتی مجموعه داده‌ها به تقسیمات فرعی و کوچک‌تر تقسیم می‌شوند تا زمانی که تقسیمات فرعی نهایی نتوانند بیش‌تر از آن تجزیه شوند (۲۱). درختان تصمیم، مجموعه‌ای از داده‌های معلوم را می‌گیرند و یک درخت تصمیم را از آن استنتاج می‌کنند. مجموعه داده‌های اولیه که درخت به وسیله آن‌ها ایجاد می‌شود، به عنوان داده‌های آموزشی شناخته می‌شوند. درخت تصمیم از بالا به پایین ترسیم می‌شود. در بالا ویژگی اول و مقدارش قرار می‌گیرد و از آن به بعد شاخه یا منجر به یک ویژگی و یا منجر به یک تصمیم می‌شود (۲۲).

اساس مدل‌های درختی مبتنی بر روش تصمیم و غلبه کردن هستند. جایگزینی معادله رگرسیون خط به جای برچسب در گره‌ها، شیوه‌ای است که در مدل M5 اجرا می‌شود که می‌تواند متغیرهای عددی پیوسته را پیش‌بینی و یا برآورد کند. ساختار درخت تصمیم‌گیری شبیه یک درخت بوده که از ۴ بخش ریشه، شاخه، گره‌ها و برگ‌ها تشکیل شده است. گره اول در درخت تصمیم به عنوان ریشه درخت در نظر گرفته می‌شود. هر گره مربوط به یک خصوصیت معین می‌باشد و شاخه‌ها به معنای بازه‌ای از مقادیر هستند. این بازه‌های مقادیر باید بخش‌های مختلف مجموعه مقادیر معلوم را برای خصوصیت‌ها به دست دهند. ساخت مدل درخت در دو مرحله انجام می‌گیرد. در مرحله اول، درخت تصمیم با انشعاب‌سازی داده‌ها تشکیل می‌شود. معیار انشعاب در مدل M5 بیشینه‌سازی کاهش انحراف معیار داده‌ها در گره فرزند است. زمانی که امکان کاهش انحراف معیار داده‌های گره فرزند میسر نبود، گره والد آن منشعب نشده و به گره پایانی و یا برگ رسیده است. فرمول مورد استفاده برای محاسبه کاهش انحراف معیار استاندارد به صورت زیر است (۲۳).

$$SDR = sd(T) - \sum_{i=1}^N \frac{|T_i|}{|T|} sd(T_i) \quad (3)$$

که در آن، T نشان‌دهنده مجموعه‌ای از نمونه‌هاست که بر هر گره وارد شده‌اند، T<sub>i</sub> نشان‌دهنده زیرمجموعه‌ای از نمونه‌هاست که

نهایی از حاصل ضرب ارزش عددی رتبه‌بندی شده هر پارامتر در وزن آن پارامتر به دست می‌آید. حاصل جمع تمامی پارامترها برای هر قسمت بیانگر درجات آسیب‌پذیری خواهد بود. بنابراین می‌توان شاخص آسیب‌پذیری را به ۸ گروه تقسیم‌بندی نمود: آسیب‌پذیری قابل چشم‌پوشی از ۲۳-۵۳؛ آسیب‌پذیری بسیار کم از ۵۴-۷۴؛ آسیب‌پذیری اندک از ۷۵-۱۰۵؛ آسیب‌پذیری اندک تا متوسط از ۱۰۶-۱۳۶؛ آسیب‌پذیری متوسط تا زیاد از ۱۳۷-۱۶۷؛ آسیب‌پذیری زیاد از ۱۶۸-۱۹۸؛ آسیب‌پذیری خیلی زیاد از ۱۹۹-۲۳۰ و کاملاً مستعد آلودگی با مقدار آسیب‌پذیری بیشتر از ۲۳۰.

آسیب‌پذیری، پتانسیل آبخوان را برای آلودگی نشان می‌دهد و نباید با خطر آلودگی اشتباه گرفته شود؛ بدین معنا که ممکن است در منطقه‌ای آسیب‌پذیری اندک و متوسط باشد، ولی به دلیل وجود گسترده منابع آلاینده، از نظر آلودگی، منطقه پرخطری محسوب شود. برعکس، ممکن است در منطقه‌ای آسیب‌پذیری بالا باشد، ولی به علت نبود منابع آلاینده، هیچ‌گونه خطر آلودگی آب زیرزمینی را تهدید نکند.

شاخص آسیب‌پذیری آبخوان آب زیرزمینی با استفاده از نقشه‌های رتبه‌بندی پارامترها و وزن هر یک از پارامترهای هفت‌گانه دراستیک نقشه پهنه‌بندی آسیب‌پذیری محدوده مورد مطالعه به دست آمد. برای نمونه‌گیری، در مجموع ۲۰ نمونه از ۲۰ حلقه چاه در منطقه کوچصفهان، به همراه مختصات جغرافیایی آن‌ها دریافت شد. برای اندازه‌گیری غلظت نیترات نمونه‌ها، روش استاندارد و از دستگاه اسپکتروفتومتر (Spectrophotometer U.V/Visible Jesco 7800) به کار گرفته شد. مقادیر نیترات در خرداد ماه سال ۱۳۹۳ اندازه‌گیری شد که در بازه ۰/۳۹-۲۸/۶۴ متغیر بود. حد مجاز نیترات در آب شرب ۴۵ میلی‌گرم بر لیتر می‌باشد (با توجه به مقدار استاندارد، در حال حاضر خطری از لحاظ نیترات متوجه آب زیرزمینی نیست) (۱۹).

### مدل‌های ناپارامتریک

#### ساختار درخت تصمیم M5

مدل درخت تصمیم M5، یک مدل درختی برای پیش‌بینی صفات عددی پیوسته است که در آن توابع رگرسیونی خطی در برگ‌های

$\hat{\mu}$  آیین نتیجه تست پتانسیلی را دارند و  $Sd$  انحراف معیار استاندارد داده‌های ورودی است و به صورت زیر محاسبه می‌شود (۲۳).

$$SDR(T) = \sqrt{\frac{1}{N} \left( \sum_{i=1}^N y_i^2 - \frac{1}{N} \left( \sum_{i=1}^N y_i \right)^2 \right)} \quad (۴)$$

که در آن،  $y_i$  مقدار عددی ویژگی هدف نمونه  $i$  و  $N$  تعداد داده‌ها را نشان می‌دهد. از آنجایی که فرآیند انشعاب (دسته‌بندی) در گره‌های فرزند دارای انحراف معیار کمتری نسبت به گره‌های والد است، از این رو دارای نتیجه دقیق‌تری هستند و از همگنی بیشتری برخوردارند. پس از بررسی همه دسته‌بندی‌های ممکن مدل  $M5$  یکی را انتخاب می‌کند که دارای حداقل خطای مورد انتظار باشد. مرحله دوم طراحی مدل درختی، شامل کوچک کردن درخت بیش از حد بزرگ شده از طریق هرس کردن شاخه‌ها و جایگزین شدن با توابع رگرسیون خطی است.

#### ساختار مدل یادگیری بر پایه نمونه IBK

روش‌های یادگیری بر پایه نمونه، روش‌های ناپارامتریکی هستند که از روابط از پیش تعیین شده برای پیش‌بینی مجهول استفاده نمی‌کنند و به جای آن از نمونه‌های مشخصی برای پیش‌بینی متغیر استفاده می‌کنند. برخی الگوریتم‌های این روش شامل  $IB1$ ،  $IB2$  و  $IB3$  هستند (۲۴). این روش‌ها فرض می‌کنند که نمونه‌های مشابه دارای دسته‌بندی مشابه هستند. بنابراین دسته‌بندی نمونه‌های جدید، براساس نمونه‌هایی که بیشترین شباهت را با آن نمونه دارند، انجام می‌شود. همچنین برون داشتن هیچ اطلاعاتی از متغیرها، فرض می‌کنند که متغیرها دارای وزن یکسان در تابع تشابه هستند. روش‌های یادگیری بر پایه نمونه از روش‌های نزدیک‌ترین همسایه توسعه یافته‌اند. الگوریتم  $IB1$  ساده‌ترین این الگوریتم‌ها و مشابه الگوریتم‌های نزدیک‌ترین همسایه هستند، با این تفاوت که  $IB1$  دامنه متغیرها را نرمال می‌کند و روشی ساده در برخورد با مقادیر گم‌شده دارد (۲۴). در این روش، با استفاده از تعداد اندکی از نمونه‌های مشابه، فرض می‌شود که تابع مورد نظر به صورت محلی خطی است و با ایجاد توابع تکه‌ای و خطی سعی در پیدا کردن تابع ناشناخته دارند. در این پژوهش، از  $IBK$  که بر

اساس  $IB1$  کار می‌کند، برای پیش‌بینی نمونه هدف استفاده شده است. روش  $IBK$  یک الگوریتم یادگیری بر اساس مشاهدات و نمونه‌ها می‌باشد (۲۵).  $IBK$  یک رده‌بند با  $K$  همسایه نزدیک است که معیار فاصله ذکر شده را استفاده می‌کند. تعداد نزدیک‌ترین فاصله‌ها (پیش فرض  $K=1$ )، می‌تواند به طور صریح در ویرایشگر شیء تعریف شود. پیش‌بینی‌های متعلق به پیش از یک همسایه می‌تواند بر اساس فاصله آنها تا نمونه‌های آزمایشی وزن دار گردد. دو فرمول متفاوت (روابط ۵ و ۶) برای تبدیل فاصله به وزن، پیاده‌سازی شده‌اند. تعداد نمونه‌های آموزشی که به وسیله رده‌بندی نگهداری می‌شود، می‌تواند با تنظیم گزینه اندازه پنجره محدود گردد. زمانی که نمونه‌های جدید اضافه می‌شوند، نمونه‌های قدیمی حذف شده تا تعداد کل نمونه‌های آموزشی در اندازه تعیین شده باقی بماند (۲۶).

$$w_i = \frac{1}{d_i} \quad (۵)$$

$$w_i = 1 - d_i \quad (۶)$$

که در آن‌ها،  $w_i$  وزن نمونه و  $d_i$  فاصله نمونه تا هدف می‌باشد. بدین صورت که همسایه‌های نزدیک‌تر، وزن بیشتری به خود اختصاص می‌دهند و تأثیر بیشتری در پیش‌بینی متغیر هدف خواهند داشت.

#### انتخاب ورودی‌های بهینه با استفاده از آزمون گاما

آزمون گاما مقدار خطای برآورد شده (واریانس خطا) را با توجه به داده‌های مستقیم نشان می‌دهد. وقتی که مقدار گاما صفر باشد، هیچ محدودیتی برای ساخت یک مدل خوب وجود نخواهد داشت، بنابراین از نقاطی که دارای گامای کمتر در مقایسه با دیگر متغیرها باشند، جهت مدل‌سازی استفاده خواهد شد (۲۷). برای استفاده از آزمون گاما ۲۰ سناریو تعریف شد تا بهترین ترکیب ورودی از پارامترها انتخاب گردد. جدول ۴ ترکیب پارامترهای ورودی به آزمون گاما را نشان می‌دهد. در این پژوهش، برای به دست آوردن بهترین ترکیب متغیرهای ورودی، از بسته نرم‌افزاری آزمون گاما موجود در نرم‌افزار WinGamma استفاده شد.

در نهایت برای انتخاب بهترین مدل، از شاخص‌های آماری، جذر



جدول ۴. پارامترهای ورودی در سناریوهای مختلف

تعداد ورودی	پارامترهای ورودی	سناریو	تعداد ورودی	پارامترهای ورودی	سناریو
۳	طول جغرافیایی - عرض جغرافیایی - توپوگرافی	۱۱	۸	محیط آبخوان - تغذیه خالص - عمق سطح ایستابی - هدایت هیدرولیکی	۱
۳	طول جغرافیایی - عرض جغرافیایی - هدایت هیدرولیکی	۱۲	۷	محیط آبخوان - بافت خاک - عمق سطح ایستابی - هدایت هیدرولیکی	۲
۲	طول جغرافیایی - عرض جغرافیایی	۱۳	۶	محیط غیراشباع - بافت خاک - عمق سطح ایستابی - هدایت هیدرولیکی	۳
۱	عمق سطح ایستابی	۱۴	۵	محیط آبخوان - بافت خاک - محیط غیراشباع	۴
۱	تغذیه خالص	۱۵	۴	بافت خاک - محیط غیراشباع	۵
۱	محیط آبخوان	۱۶	۳	بافت خاک - عرض جغرافیایی - طول جغرافیایی	۶
۱	بافت خاک	۱۷	۳	عمق سطح ایستابی - عرض جغرافیایی - طول جغرافیایی	۷
۱	محیط غیراشباع	۱۸	۳	تغذیه خالص - عرض جغرافیایی - طول جغرافیایی	۸
۱	توپوگرافی	۱۹	۳	محیط آبخوان - عرض جغرافیایی - طول جغرافیایی	۹
۱	هدایت هیدرولیکی	۲۰	۳	عرض جغرافیایی - محیط غیراشباع - طول جغرافیایی	۱۰

### تغذیه خالص

نقشه تغذیه خالص دشت بر اساس مدل دراستیک در شکل ۲ (ب) مشاهده می‌شود. این نقشه بر اساس همپوشانی سه مؤلفه مقدار بارش سالانه، شیب زمین و نفوذپذیری خاک ترسیم شد. آبخوان به ۴ منطقه تقسیم‌بندی شد که به ترتیب وسعتشان در دشت شامل منطقه زردرنگ با محدوده تغذیه خالص (۵۰/۸-۱۰۱/۶) که قسمت وسیعی از دشت می‌باشد؛ منطقه آبی‌رنگ که دارای بیشترین میزان تغذیه خالص (>۲۵۴) بود و منطقه سبز رنگ که دارای تغذیه خالص در محدوده (۱۷۷/۸-۱۰۱/۶) بود. تغذیه در این سه منطقه بیشتر تحت تأثیر بارندگی و نفوذپذیری زیاد دشت است. بر اساس مدل دراستیک، به این مناطق رتبه ۳، ۹ و ۶ اختصاص داده شد.

### محیط آبخوان

با استفاده از اطلاعات به‌دست آمده از حفاری‌های اکتشافی، نقشه محیط آبخوان دشت در شکل ۳ (الف) نشان داده شده است. آبخوان دشت از نظر محیط آبخوان به دو منطقه که به رنگ‌های مختلف نشان داده شده، تقسیم شد. منطقه بنفش که بخش وسیعی از آبخوان را در برمی‌گیرد، از دانه‌های رسوبات متوسط‌دانه تا گراول درشت‌دانه

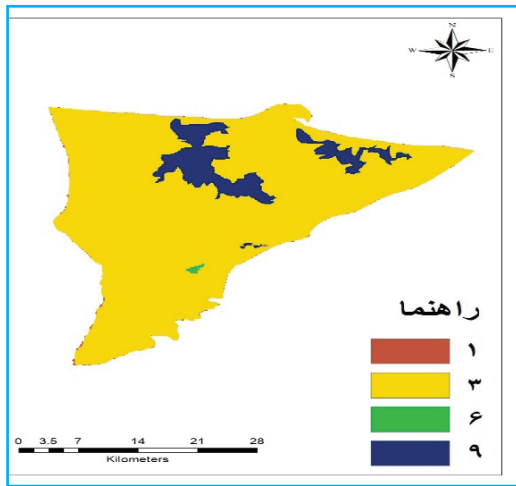
میانگین مربعات خطا<sup>۱</sup>، ضریب تبیین<sup>۲</sup> و متوسط قدرمطلق خطا<sup>۳</sup> استفاده گردید. کم بودن میزان RMSE و MAE و بالا بودن  $R^2$ ، بیانگر دقت قابل قبول مدل و ملاک برتری آن خواهد بود.

### یافته‌ها

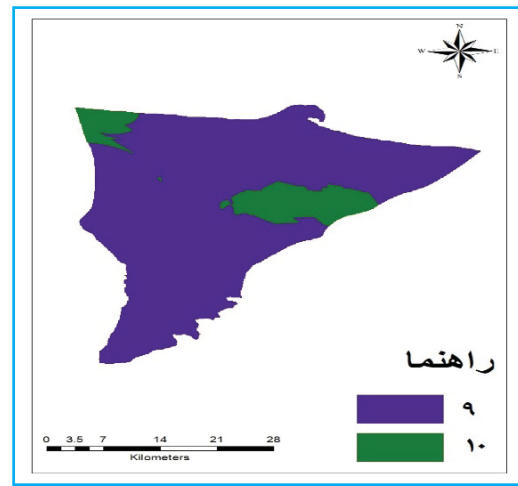
#### عمق سطح ایستابی

نقشه لایه‌بندی عمق تا سطح ایستابی در دشت آستانه-کوچصفهان بر اساس مدل دراستیک در شکل ۲ (الف) نشان داده شده است. آبخوان دشت آستانه-کوچصفهان را می‌توان به دو منطقه تقسیم کرد که با دو رنگ متفاوت در نقشه نشان داده شده است. منطقه رنگ بنفش که بخش وسیعی (مرکز و قسمت‌های شمالی و جنوبی) از دشت را شامل می‌شود و عمق ایستابی در محدوده ۴/۵-۱/۵ متر بوده و در مدل دراستیک دارای رتبه‌بندی ۹ می‌باشد و منطقه سبزرنگ که نواحی کوچکی (شمال غرب و شرق) را شامل می‌شود، دارای سطح ایستابی بین ۱/۵-۰ متر بوده و در مدل دراستیک دارای رتبه‌بندی ۱۰ می‌باشد.

1. Root Mean Square Error (RMSE)
2. Coefficient of Determination ( $R^2$ )
3. Mean Absolute Error (MAE)



(ب)



(الف)

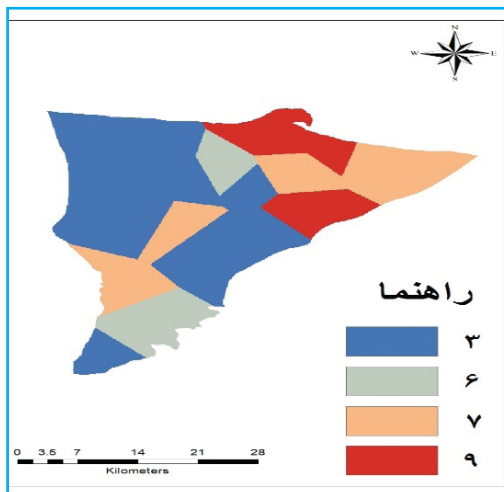
شکل ۲. (الف) ارزش گذاری عمق سطح ایستابی، (ب) ارزش گذاری تغذیه خالص

شامل بخش‌هایی از شمال غرب، مرکز و قسمت‌های جنوبی که دارای نفوذپذیری کمتر بوده و در نتیجه قابلیت انتقال آلاینده‌ها به سیستم آب زیرزمینی نیز کاهش می‌یابد. بخش دوم خاک‌هایی با بافت ماسه با رتبه ۹ که محدوده شمال و بخشی از ناحیه شرقی دشت را شامل گردید. بخش سوم خاک‌هایی با رتبه ۷ و بافت رس متراکم که بخش‌هایی از شمال شرق و جنوب غرب دشت را در برمی‌گیرد. بخش آخر خاک‌هایی با بافت لوم ماسه‌ای و رتبه ۶ که بخش‌هایی از مرکز و جنوب شرقی دشت را شامل می‌شود.

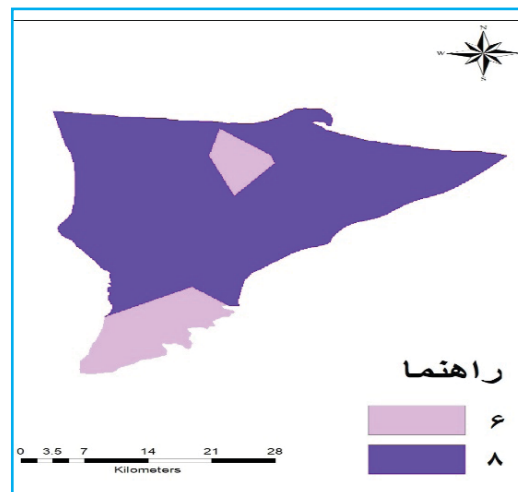
تشکیل شده است. بر اساس مدل دراستیک، رتبه ۸ به این بخش‌ها اختصاص داده شد. منطقه بنفش کم‌رنگ (یاسی) شامل بخش‌هایی از شمال و جنوب آبخوان که حاوی ماسه‌سنگ‌های توده‌ای بوده و در مدل دراستیک رتبه ۶ به این بخش‌ها اختصاص داده شد.

#### محیط خاک

نقشه لایه‌بندی محیط خاک بر اساس مدل دراستیک در شکل ۳ (ب) نشان داده شده است. لایه خاک وجود ۴ نوع طبقه مختلف را نشان می‌دهد. بخش اول خاک با بافت لوم رس دار با رتبه ۳ که



(ب)



(الف)

شکل ۳. (الف) ارزش گذاری محیط آبخوان، (ب) ارزش گذاری محیط خاک

### محیط غیراشباع

با توجه به اطلاعات هیدرولوژی منطقه و بر اساس لوگ‌های چاه‌های موجود در منطقه، رسوبات ناحیه غیراشباع در آبخوان دارای تنوع می‌باشد. نقشه لایه‌بندی بخش غیراشباع مورد استفاده در مدل در شکل ۴ (الف) نشان داده شده است. بر اساس این نقشه، سه منطقه در دشت آبخوان مورد نظر قابل مشاهده بود. منطقه آبی‌رنگ شامل بخش‌های شمال‌شرقی و مرکزی دشت متشکل از سیلت، رس، شیل با بخش‌هایی از سنگ آهک می‌باشد. رتبه اختصاص یافته به این مناطق ۳ بود. منطقه زرد رنگ، بخشی از شمال، جنوب، غرب و شرق دشت بوده و حاوی بافت درشت‌تر ماسه‌سنگ، گراول با بخش مهمی از سیلت و رس بوده و رتبه اختصاص یافته به این مناطق ۶ بود. هم‌چنین منطقه قرمز رنگ که شامل بخش‌هایی از جنوب و جنوب‌غربی دشت آستانه-کوچصفهان بوده و متشکل از شن و ماسه و دارای رتبه‌بندی ۸ در مدل دراستیک می‌باشد.

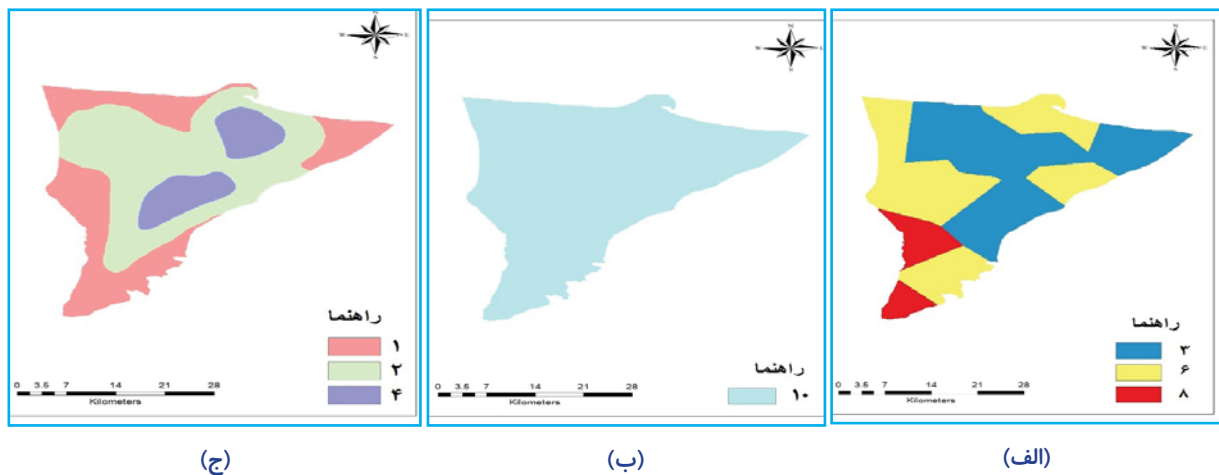
### توپوگرافی

نقشه ارزش‌گذاری توپوگرافی دشت آستانه-کوچصفهان بر اساس مدل دراستیک در شکل ۴ (ب) نشان داده شده است. کل منطقه

مورد مطالعه دارای شیب کمتر از ۲٪ بوده که در مدل دراستیک با رتبه ۱۰ مشخص شد. بنابراین میزان شیب در دشت کم بوده و این عامل پتانسیل آلودگی آب زیرزمینی را افزایش می‌دهد.

### هدایت هیدرولیکی

شکل ۴ (ج) نقشه نرخ‌گذاری شده هدایت هیدرولیکی را نشان می‌دهد. با توجه به شکل، سه ناحیه مشخص در دشت تقسیم‌بندی گردید. ناحیه اول دارای کمترین میزان هدایت هیدرولیکی (۰/۰۴-۴/۱) با رنگ صورتی شامل بخش‌هایی از شمال‌غرب، غرب، جنوب و شمال‌شرق دشت می‌باشد که بر اساس نرخ‌گذاری مدل دراستیک، کمترین نرخ آسیب‌پذیری یک به این بخش‌ها اختصاص داده شد. ناحیه دوم یا منطقه آبی کم‌رنگ که قسمت عمده دشت و دارای هدایت هیدرولیکی بین ۴/۱-۱۲/۳ متر بر روز بوده، با رتبه ۲ مشخص شد. ناحیه سوم، منطقه آبی‌رنگ که شامل بخش‌های مرکزی دشت بوده و دارای هدایت هیدرولیکی ۱۲/۳-۲۸/۷ متر بر روز بوده و بر اساس نرخ‌گذاری مدل دراستیک، رتبه ۴ به این بخش‌ها اختصاص داده شد.

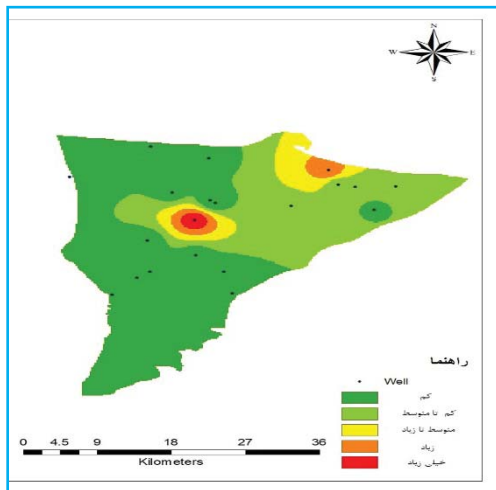


شکل ۴. (الف) ارزش‌گذاری محیط غیراشباع، (ب) ارزش‌گذاری توپوگرافی، (ج) ارزش‌گذاری هدایت هیدرولیکی

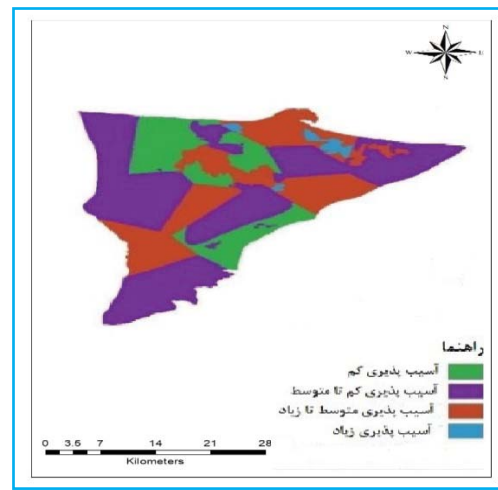
### محاسبه شاخص آسیب‌پذیری شاخص دراستیک

بعد از اینکه پارامترهای دراستیک در منطقه به‌دست آمد، با یکدیگر ترکیب و نقشه آسیب‌پذیری آبخوان به‌دست آمد. شکل ۵ (الف) نقشه پهنه‌بندی آسیب‌پذیری آبخوان را با استفاده از مدل دراستیک نشان می‌دهد. همانطور که در شکل مشاهده می‌شود،

فعالیت‌های صنعتی و کودهای شیمیایی به دشت قابل پیش‌بینی بود. به نظر می‌رسد منشأ اصلی نیترات آب‌های زیرزمینی در دشت مورد مطالعه ناشی از فعالیت‌های کشاورزی باشد. استفاده از کودهای شیمیایی به ویژه اوره در کشت برنج، فضولات حیوانی که برای بهبود کیفیت خاک در شالیزارها و زمین‌های کشاورزی استفاده می‌شوند، شایع‌ترین راه‌های ورود نیترات به آب‌های زیرزمینی می‌باشند. میزان همبستگی بین شاخص دراستیک با غلظت نیترات ۸۰٪ به‌دست آمد.



(ب)



(الف)

شکل ۵. (الف) نقشه پهنه‌بندی آسیب‌پذیری با استفاده از مدل دراستیک استاندارد، (ب) نقشه پهنه‌بندی مقدار نیترات اندازه‌گیری شده در نقاط مختلف حوضه

### نتایج مدل‌های ناپارامتریک IBK و M5

درخت تصمیم M5، پیش‌بینی مطلوب‌تری را حاصل می‌نماید.

### جدول ۵. نتایج به‌دست آمده از مدل‌های ناپارامتریک IBK و M5

آزمون			آموزش			مدل
MAE	RMSE	R <sup>2</sup>	MAE	RMSE	R <sup>2</sup>	
۰/۵۱	۰/۵۴	۰/۹۸	۰/۳۵	۰/۳۴	۰/۹۸	درخت تصمیم M5
۲/۵۹	۶/۳۷	۰/۸۳	۰/۵۱	۰/۵۴	۰/۹۷	IBK

در این پژوهش با در نظر گرفتن پارامترهای ورودی، ۲۰ ترکیب مختلف ایجاد شد که با استفاده از آزمون گاما و بررسی آماره‌های آن، ترکیبی که دارای مقدار آماره گاما، گرادیان، انحراف معیار و

نقاطی با آسیب‌پذیری زیاد با رنگ آبی نمایش داده شده است. آسیب‌پذیری آبخوان این دشت در ۴ محدوده قرار دارد. ۱۸/۵۶٪ دشت دارای آسیب‌پذیری اندک، ۵۱/۲۹٪ دارای آسیب‌پذیری اندک تا متوسط، ۲۸/۴۶٪ دارای آسیب‌پذیری متوسط تا زیاد و ۱/۶۷٪ دارای آسیب‌پذیری زیاد می‌باشد. شکل ۵ (ب) مقدار نیترات اندازه‌گیری شده در نقاط مختلف حوضه را نشان می‌دهد. افزایش میزان نیترات در چاه‌های منطقه مورد مطالعه به دلیل پتانسیل آسیب‌پذیری بالا و متوسط برای بخش گسترده‌ای از منطقه، تداوم و ورود کنترل نشده مواد نیتراتی به ویژه از طریق

برای این منظور، داده‌های ورودی (پارامترهای دراستیک، غلظت نیترات اندازه‌گیری شده، طول جغرافیایی و عرض جغرافیایی) و خروجی (مقدار نیترات اندازه‌گیری شده و پهنه‌بندی شده در ۳۰ حلقه چاه موجود در منطقه) مدل و شبکه به دو دسته آموزش (۸۰٪) و آزمون (۲۰٪) برای کارایی مدل، تقسیم‌بندی شد. با توجه به وجود امکان ساختار مکانی بین داده‌ها، طول و عرض جغرافیایی به این پارامترها افزوده شده است (۲۸). جدول ۵ نتایج حاصل از الگوریتم‌های مختلف را با مقدار نیترات اندازه‌گیری شده در منطقه نشان می‌دهد. با توجه به نتایج ضریب تمییز، RMSE و MAE مدل

درخت تصمیم M5 به دلیل داشتن همبستگی زیاد و خطای کمتر، نتایج مطلوب تری را برآورد می‌کند.

**جدول ۷. مدل درخت تصمیم M5 و IBK در حالت بهترین و بدترین ترکیب پارامترهای ورودی**

درخت تصمیم M5						سناریو
آزمون			آموزش			
MAE	RMSE	R <sup>2</sup>	MAE	RMSE	R <sup>2</sup>	
۰/۵۱	۰/۶۱	۰/۹۸	۰/۳۵	۰/۴۳	۰/۹۸	۱
۱۲/۰۷	۱۵/۲۵	۰/۰۰	۹/۵۱	۱۱/۴۹	۰/۰۰	۱۵

**ادامه جدول ۷. مدل درخت تصمیم M5 و IBK در حالت بهترین و بدترین ترکیب پارامترهای ورودی**

IBK						سناریو
آزمون			آموزش			
MAE	RMSE	R <sup>2</sup>	MAE	RMSE	R <sup>2</sup>	
۲/۵۹	۴/۷۹	۰/۹۱	۰/۵۳	۰/۶۱	۰/۹۸	۱
۱۱/۸۷	۱۵/۱۲	۰/۰۱	۱۱/۹۱	۱۵/۲۵	۰/۰۰	۱۵

### بحث

آب‌های زیرزمینی، تأمین‌کننده بخش قابل توجهی از نیاز آبی در زمینه‌های شرب، صنعت و کشاورزی می‌باشند. آلودگی این منابع آبی منجر به کاهش کیفیت آب آشامیدنی، مشکلات بهداشتی و زیست‌محیطی و افزایش هزینه‌های پاک‌سازی یا جایگزینی منابع تأمین آب می‌گردد. شناسایی مناطق آسیب‌پذیر آبخوان، یکی از راهکارهای مناسب جهت جلوگیری از آلودگی آب‌های زیرزمینی می‌باشد. بنابراین در این پژوهش، آسیب‌پذیری دشت کوچصفهان با استفاده از روش دراستیک و مدل ناپارامتریک WEKA تخمین زده شد که بر اساس نتایج، این دشت در ۵۱/۲۹٪ دارای آسیب‌پذیری اندک تا متوسط و ۲۸/۴۶٪ دارای آسیب‌پذیری متوسط تا زیاد می‌باشد. همچنین مدل درخت تصمیم M5 به منزله مدل برتر در آسیب‌پذیری دشت مورد مطالعه انتخاب گردید. علت آلودگی برای این دشت، زراعت آبی و به‌خصوص شالیکاری می‌باشد. به دلیل اینکه در زمین‌های شالی به علت مصرف بالای آب و وجود مقدار زیاد رس در خاک، جذب آب زیاد و سطح ایستابی بالاست و از

نسبت V کمتر بود، به‌عنوان بهترین ترکیب انتخاب گردید. نتایج آزمون گاما و آماره‌های آن در جدول ۶ ارائه شده است. در این جدول ترکیب ورودی بر اساس کم‌ترین میزان آماره گاما رتبه‌بندی شده‌اند. نتایج نشان می‌دهد سناریوی اول که ترکیبی دارای پارامترهای ورودی طول جغرافیایی، عرض جغرافیایی، عمق سطح ایستابی، تغذیه خالص، محیط آبخوان، بافت خاک، محیط غیراشباع و هدایت هیدرولیکی باشد، آماره‌های پایین‌تری دارد و بهترین برآورد را در مقدار آسیب‌پذیری دشت حاصل می‌کند و بدترین پیش‌بینی زمانی رخ می‌دهد که فقط پارامتر تغذیه خالص (سناریوی ۱۵) مدنظر قرار گیرد.

**جدول ۶. نتایج آزمون گاما**

رتبه	سناریو	پارامتر		
		گاما	گرادیان	انحراف معیار
۱	۱	۰/۰۰۰۶	۰/۰۷۴۰	۰/۰۰۰۷
۲	۲	۰/۰۰۴۹	۰/۰۹۷۷	۰/۰۰۰۸
۳	۳	۰/۰۱۲۴	۰/۱۲۸۸	۰/۰۰۱۱
۴	۱۱	۰/۰۳۹۳	۲۵/۹۳۳۰	۰/۰۰۳۳
۵	۵-۴	۰/۰۵۱۳	۵/۵۷۷۷	۰/۰۰۳۹
۶	۱۰	۰/۰۵۲۴	۹/۰۹۴۴	۰/۰۰۴۱
۷	۶	۰/۰۵۲۷	۸/۵۰۹۸	۰/۰۰۴۳
۸	۸	۰/۰۵۵۷	۵/۶۵۵۳	۰/۰۰۴۸
۹	۱۳-۹	۰/۰۵۸۰	۱۵/۳۳۳	۰/۰۰۵۶
۱۰	۱۲	۰/۰۶۶۸	۷/۲۱۲	۰/۰۰۹۹
۱۱	۷	۰/۰۶۸۷	۱/۲۸۱۱	۰/۰۱۹۷
۱۲	۱۴	۰/۱۸۹۴	۰/۰۷۳۷	۰/۰۰۴۶
۱۳	۱۹	۰/۲۳۸۷	۱۰/۶۸۵۰	۰/۰۰۴۱
۱۴	۲۰	۰/۲۴۶۰	۰/۰۰۷۴	۰/۰۱۰۳
۱۵	۱۶-۱۷	۰/۲۵۰۰	۰/۰۰۰۰	۰/۰۰۰۰
۱۶	۱۵	۰/۳۷۴۰	۰/۰۰۶۰	۰/۰۰۰۰

مدل درخت تصمیم M5 و IBK در حالت بهترین و بدترین ترکیب پارامترهای ورودی انتخاب شده از آزمون گاما اجرا گردید و نتایج و آماره‌های این دو مدل در جدول ۷ مورد تحلیل قرار گرفت. همانطور که مشاهده می‌شود، در حالت بهترین ترکیب ورودی مدل

هم‌چنین نتایج نشان داد که هر دو مدل ناپارامتریک به کار گرفته شده، تخمین مناسبی از مقدار نیترات می‌دهند، اما با توجه به نتایج ضریب تبیین ( $R^2$ )، RMSE و MAE، مدل درخت تصمیم M5 به منزله مدل برتر در آسیب‌پذیری دشت انتخاب شد. نتایج آزمون گاما نشان داد که ترکیبی که دارای همه پارامترهای ورودی باشد، بهترین برآورد و ترکیبی که فقط پارامتر تغذیه خالص را مدنظر قرار دهد، بدترین پیش‌بینی را ارائه می‌دهد.

#### ملاحظات اخلاقی

نویسندگان تمام نکات اخلاقی شامل عدم سرقت ادبی، انتشار دوگانه، تحریف داده‌ها و داده‌سازی را این مقاله رعایت کرده‌اند. همچنین هرگونه تضاد منافع حقیقی یا مادی که ممکن است بر نتایج یا تفسیر مقاله تأثیر بگذارد را رد می‌کنند.

#### تشکر و قدردانی

بدین‌وسیله از سازمان آب منطقه‌ای استان گیلان، اساتید و دوستان که ما را در نگارش این پژوهش یاری نمودند، تشکر و قدردانی می‌گردد.

## References

1. Vrba J, Zoporozec A. Guidebook on mapping groundwater vulnerability. IAH International Contribution for Hydrogeology 1994; Vol. 16: xxiii, 131 p.
2. Almasri M.N. Assessment of intrinsic vulnerability to contamination for Gaza coastal aquifer, Palestine. Journal of Environmental Management 2008; 88, 577-593.
3. Gogu R.C, Dassargues A. Current trends and future challenge in groundwater vulnerability assessment, using overlay and index methods. Journal of Environmental Geology 2000; 39, 549-558.
4. Ersoy A.F, Gultekin F. DRASTIC- Based Methodology for Assessing Groundwater Vulnerability in the Gumushacikoy and Merzifon Basin (Amasya, Turkey), Earth Sci. Res. S.J.J 2013; 17:1, 33-40.
5. Jafari S.M, Nikoo M.R. Groundwater risk assessment based on optimization framework using DRASTIC method. Arab J Geosci 2016; 9,742
6. Kholghi M, Taki R. Evaluating Groundwater vulnerability in Ghazvin Plain. Journal of Engineering Geology 2004; 1(3), 255-270 (Persian).
7. Amirahmadi, A., Ebrahimi, M., Asadi, M.A. and Akbari, E. (2013). Evaluation of the vulnerability of Neishabour aquifer by DRASTIC method using GIS. Geography and Environmental Hazards, 2(6), 37-56 (In Persian).
8. Sadatipoor E, Noori R, Baghvand A, et al. Application of DRASTIC Model for Groundwater Vulnerability Assessment of the Ghaen Aquifer. Journal Environmental Sciences Studies 2016; 1(2), 63-71 (Persian).
9. Yoosofdoost I, Khashei Siuki A. Determine the Vulnerability of the Aquifer Using the Standard Drastic and Data-Based Methods (Case Study: Kochisfahn). Journal of Remote Sensing & GIS 2017; 9(2), 99-116 (Persian).
10. Khalili Naft Chali A, Shahidi A, Khashei Siuki A. Comparison of Lazy Algorithms and M5 Model to Estimate Groundwater Level (Case Study: Plain Neyshabur). J. Water and Soil Sci (Sci. & Technol. Agric. & Natur. Resour.) 2017; 21(3), 15-26 (Persian).
11. Sadeghzadeh Sadat M, Nazemi A.H, Sadraddini A.A. Impacts of Surface Water Quality on Groundwater Quality (Case study: Tabriz plain). Water and Soil Science 2017; 27(3), 225-237 (Persian).
12. Gesim N.A, Okazaki T. (2018). Assessment of Groundwater

آنجایی که کشاورزان مقدار زیادی کود نیترات استفاده می‌کنند، این کود در آب حل شده و به راحتی در زمین نفوذ می‌کند و به سطح ایستایی می‌رسد و باعث آلودگی زیاد می‌شود. نتایج پژوهش حاضر با نتایج یوسف دوست و خاشعی سیوکی (۹)، آرزومندامیدی لنگرودی و همکاران (۱۵) و مدبری و همکاران (۲۹) همخوانی داشت. همچنین در این خصوص صادق‌زاده سادات و همکاران پژوهشی بر روی رودخانه آچی‌چای واقع در دشت تبریز به منظور شناسایی عوامل مؤثر بر آلودگی رودخانه و آبخوان دشت تبریز با بهره‌گیری از روش‌های شبکه عصبی مصنوعی، درخت تصمیم M5 و زمین آمار در طی دوره ۱۳۸۰-۹۲ انجام دادند که بر اساس نتایج این پژوهش، روش درخت تصمیم M5 به عنوان مدل برتر برای پیش‌بینی معرفی گردید (۱۱).

#### نتیجه‌گیری

بر اساس نتایج این پژوهش، آسیب‌پذیری آبخوان دشت کوچصفهان در ۱۸/۵۶٪ دارای آسیب‌پذیری اندک، ۵۱/۲۹٪ دارای آسیب‌پذیری اندک تا متوسط، ۲۸/۴۶٪ دارای آسیب‌پذیری متوسط تا زیاد و ۱/۶۷٪ دارای آسیب‌پذیری زیاد می‌باشد.

- Vulnerability to Pollution using DRASTIC Model and Fuzzy Logic in Herat City, Afghanistan. *International Journal of Advanced Computer Science and Applications* 2018; 9(10), 181-188.
13. Hosseini M, Saremi A. Assessment and Estimating Groundwater Vulnerability to Pollution Using a Modified DRASTIC and GODS Models (Case Study: Malayer Plain of Iran), *Civil Engineering Journal* 2018; 4(2), 433-442.
  14. Ghanbarian M, Ahmadi Nadoushan M. Determination of Aquifer Vulnerability in Lordegan Aquifer Using DRASTIC, AVI and GODS Models, *Journal of Research in Environmental Health* 2019; 4(4): 257-271.
  15. Arezooman omidi langrudi M, Khashei Siuki A, Javadi S, et al. Groundwater Vulnerability Assessment by the use of Drastic-Nw Modified Model (Case study: Kuchesfehah-Astane Plain). *Iranian Journal of Irrigation and Drainage* 2015, 1(9), 75-82 (Persian).
  16. Babiker I.S, Mohamed M.A.A, Hiyama T, et al. A GIS-based DRASTIC Model for Assessing Aquifer Vulnerability in Kakamigahara Heights, Gifu Prefecture, Central Japan, *Science of the Total Environment* 2005; 345(1-3), 127-140.
  17. Thirumalaivasan D, Karmegam M, Venugopal K. AHP-DRASTIC: software for specific aquifer vulnerability assessment using DRASTIC model and GIS. *Environmental Modeling and Software* 2001; 18, 645-656.
  18. Aller L, Bennet T, Leh R.J.H, et al. DRASTIC: A Standardized System for Evaluating Ground Water Pollution Potential Using Hydrogeologic Settings, EPA/600/2-87/035, U.S. Environmental Protection Agency, Ada, Oklahoma, 1987; PP: 19-25.
  19. Arezooman omidi langrudi M, Khashei Siuki A, Javadi S, et al. Evaluation of vulnerability of aquifers by improved fuzzy drastic method: Case study: Aastane Kochesfahan plain in Iran. *Ain Shams Engineering Journal* 2016; 7, 11-20.
  20. Pal M. M5 model tree for land cover classification. *International Journal of Remote Sensing* 2006; 27(4), 825-831.
  21. Talebi A, Akbari Z. Investigation of Ability of Decision Trees Model to Estimate River Suspended Sediment (Case Study: Ilam Dam Basin). *Journal of Water and Soil Science (Journal of Science and Technology of Agriculture and Natural Resources)* 2013; 17(63), 109-121 (Persian).
  22. Yousefi M, Talebi A.S, Poursabriati R. Application of Artificial Intelligence in Water and Soil Science. Yazd University Press, 2014 (Persian).
  23. Alberg D, Las T.M, Kandel A. Knowledge discovery in data streams with regression tree methods. *WIREs Data Mining Know Discover* 2012; 2: 69-78.
  24. Aha D.W, Kibler D, Albert M.K. Instance-based learning algorithms. *Machine learning* 1991; 6:37-66.
  25. Shirzad A, Soltani F, Zare Abyaneh H. Simulation of scouring in accordance with energy-depleting structures using k-nearest closest neighbor (KNN) algorithm and non-fuzzy adaptive inference system (ANFIS). *First International Water Crisis Summit. University of Zabol, 2008, (Persian).*
  26. Jalali V.R, Homaee M. Introducing a Nonparametric Model Using k-Nearest Neighbor Technique for Predicting Soil Bulk Density. *Journal of Water and Soil Science (Journal of Science and Technology of Agriculture and Natural Resources)* 2011; 15 (56), 181-191(Persian).
  27. Moghaddamnia A., Remesan R., Hassanpour Kashani M., Mohammadi M., Han D and Piri J. Comparison of LLR, MLP, Elman, NNARX and ANFIS Models with a case study in solar radiation estimation. *Journal of Atmospheric and Solar-Terrestrial Physics.* 2009, 71: 975-982.
  28. Khashei Siuki A, Sabazi M. Evaluation of ANFIA, ANN, and geostatistical models to spatial distribution of groundwater quality (case study: Mashhad in Iran). *Arabian Journal of Geosciences* 2015; 8(2), 903-912.
  29. Modaberi H, RahbarHashemi M.M, Ashurnia M. Groundwater Resource Vulnerability Analysis Using Drastic Method and Comparison with Nitrate Parameter in Guilan. *Third National Conference on Environmental Science and Management, 2017, Ardebil (Persian).*