

## بررسی آلودگی آب زیرزمینی دشت ملکان به آرسنیک

حسین نوروزی<sup>۱</sup>، عطاالله ندیری<sup>۲\*</sup>، اصغر اصغری مقدم<sup>۳</sup>

۱. کارشناس ارشد هیدروژئولوژی، دانشکده علوم طبیعی، گروه علوم زمین، دانشگاه تبریز

۲. استادیار هیدروژئولوژی، دانشکده علوم طبیعی، گروه علوم زمین، دانشگاه تبریز

۳. استاد هیدروژئولوژی، دانشکده علوم طبیعی، گروه علوم زمین، دانشگاه تبریز

(تاریخ دریافت ۱۳۹۵/۰۳/۰۲؛ تاریخ تصویب ۱۳۹۵/۰۶/۱۹)

### چکیده

از آنجا که در زمینه وضعیت پراکنش عناصر سنگین در آب‌های زیرزمینی دشت ملکان اطلاعاتی وجود ندارد، این مطالعه با هدف بررسی فلزات سنگین، به‌خصوص آرسنیک در آب زیرزمینی و تعیین عوامل مهم بر آنومالی آرسنیک این دشت انجام پذیرفت. به همین منظور، نمونه‌برداری از منابع آب زیرزمینی انجام شد و آنالیزهای هیدروشیمیایی در آزمایشگاه آب‌شناسی دانشگاه تبریز انجام شد. همچنین برخی عناصر سنگین از قبیل آهن، آلومینیوم، منگنز، آرسنیک و کروم با روش جذب اتمی-کوره گرافیتی در آزمایشگاه کنترل کیفی آب-فاضلاب تبریز اندازه‌گیری شدند. در مطالعه حاضر روش Random Forest (RF) که یک روش یادگیری مبتنی بر دسته‌ای از درخت‌های تصمیم است، برای ارزیابی احتمال آلودگی آرسنیک پیشنهاد شده که تا کنون در این زمینه استفاده نشده است. روش RF نسبت به روش‌های دیگر مزایایی مانند دقت پیش‌بینی زیاد، توانایی در یادگیری روابط غیرخطی، توانایی زیاد در تعیین متغیرهای مهم در پیش‌بینی و ماهیت غیرپارامتری دارد. براساس پیش‌بینی مدل RF، مقدار قابلیت انتقال، نیترات، هدایت هیدرولیکی و وجود شهرها به‌عنوان تأثیرگذارترین پارامترها در وجود آنومالی آرسنیک شناخته شدند. وجود همبستگی زیاد بین مقادیر نیترات و آرسنیک، به منشأ انسان‌زاد آنومالی آرسنیک، به‌خصوص به دلیل نبود شبکه فاضلاب ارتباط پیدا می‌کند. همچنین آلودگی قسمت جنوب شرقی آبخوان را می‌توان به منشأ زمین‌زاد ناشی از سولفیدهای آرسنیک درون شکستگی‌های موجود در مارن‌های میوسن ربط داد. بر اساس نقشه به‌دست‌آمده از مدل، ۱۳ درصد از وسعت منطقه مطالعه‌شده در محدوده آسیب‌پذیری یا احتمال آلودگی بسیار کم، ۵۳ درصد در محدوده آسیب‌پذیری کم، ۲۱ درصد در محدوده آسیب‌پذیری متوسط، ۱۱/۵ درصد در محدوده آسیب‌پذیری زیاد و ۱/۵ درصد در محدوده آسیب‌پذیری بسیار زیاد نسبت آرسنیک قرار می‌گیرد.

**کلیدواژگان:** آب زیرزمینی، آرسنیک، دشت ملکان، قابلیت انتقال، نیترات، Random Forest.

## مقدمه

در میان آلاینده‌های محیطی، فلزات سنگین به دلیل تأثیرات فیزیولوژیکی بر موجودات زنده اهمیت خاصی دارند. انباشت این عناصر در خاک و آب، سبب ورود آنها به چرخه غذایی و تهدید سلامت انسان و سایر موجودات می‌شود. بنابراین، بررسی توزیع غلظت عناصر کمیاب برای پایش آلودگی آب و حفظ کیفیت محیط زیست ضروری است. عناصر سنگین به دلیل پایداری محیطی، بازیافت بیوژئوشیمیایی، پدیده جذب-واجذبی، پتانسیل اکسایش-کاهش، ته‌نشست، انحلال پذیری و خطرات بوم‌شناختی مورد توجه بیشتری قرار گرفته‌اند [۱۱ و ۱۷]. منابع طبیعی و انسانی دو منبع مهم برای آلودگی آب با فلزات سنگین هستند. منابع طبیعی شامل ورود فلزات سنگین از طریق فرسایش مواد مادری خاک و سنگ می‌شود و بنابراین به زمین‌شناسی منطقه مربوط است. انحلال پذیری و تحرک فلزات سنگین در آب‌های زیرزمینی به فرایندهای طبیعی و تغییرات فیزیوشیمیایی محیط آبخوان بستگی دارد. به علاوه، آب زیرزمینی می‌تواند از شیرابه زباله‌ها، فاضلاب، شست‌وشوی پسماندهای معدنی، نشت از لاگون‌های باطله صنعتی یا دریاچه‌های صنعتی، با فلزات سنگین آلوده شود [۱۶]. یکی از عناصر مهمی که سبب خطرات جدی بر سلامت محیط زیست می‌شود، شبه‌فلز آرسنیک است. آرسنیک به‌عنوان جزء اصلی در بیش از ۲۰۰ کانی، به شکل سولفید، اکسید، آرسنات و آرسنیت وجود دارد که منابع طبیعی آرسنیک را تشکیل می‌دهند [۲۸]. آرسنیک در آب زیرزمینی برخی نقاط جهان مثل آمریکا، شیلی، بنگلادش [۲۰، ۲۲ و ۲۸]، و اخیراً ایران [۲۱ و ۲۴] گزارش شده است. مکانیزم اصلی ورود آرسنیک به فاز آبی، هیدروکسیدی آهن و منگنز و کانی‌های سولفیدی آرسنیک‌دار است [۳۲]. براساس استاندارد آژانس حفاظت محیطی آمریکا [۲۰] و سازمان بهداشت جهانی [۳۰] بیشترین غلظت آرسنیک در منابع آب طبیعی ۱۰ ppb تعیین شده است. با این حال، هنوز استاندارد داخلی ایران ۵۰ ppb است که این مقدار از نظر سازمان بهداشت جهانی نیز زیاد محسوب می‌شود و سلامتی افراد را تهدید می‌کند. اصغری مقدم و برزگر [۱] در بررسی عوامل مؤثر بر غلظت زیاد آرسنیک در آب زیرزمینی آبخوان دشت تبریز، به این نتیجه رسیدند که غلظت آرسنیک در منطقه به عمق چاه، نوع آبخوان و خصوصیات هیدروژئوشیمیایی آن، فاصله از محل تغذیه و زمان ماندگاری آب در زیر زمین

بستگی دارد و منشأ آلودگی آرسنیک در آبخوان دشت تبریز را زمین‌زاد تشخیص دادند و معلوم کردند که حضور آرسنیک در چاه‌های عمیق منطقه نشانگر ارتباط آن با عمق نیست بلکه ارتباط کنترلی زمین‌شناسی و ژئوشیمیایی را نشان می‌دهد. نتایج ندییری و همکارانش [۲] در بررسی آنومالی آرسنیک موجود در منابع آب سد سهند، نشان داد تغلیظ آبی ترمالی آرسنیک در سازندهای میوسن می‌تواند منشأ اصلی آنومالی آرسنیک در منابع آبی منطقه باشد.

امروزه از روش‌های مختلفی برای بررسی غلظت فلزات سنگین استفاده می‌شود. گام نخست در ارزیابی گسترده‌گی و شدت آلودگی عناصر سنگین در مناطق مشکوک به آلودگی، نمونه‌برداری و تعیین غلظت فلزات است و استفاده از یک مدل آماری در کنار این اندازه‌گیری‌ها می‌تواند در مدیریت مناطق یادشده تأثیر بسزایی داشته باشد. تا کنون روش‌های آماری‌ای که برای ارزیابی آلودگی آب زیرزمینی ابداع شده‌اند، به‌طور نسبی از داده‌های کمتری استفاده می‌کنند [۲۷]. در سالیان اخیر، ابزارهای یادگیری<sup>۱</sup> و روش‌های جدیدی برای حل برخی از مشکلات یادشده ارائه شده‌اند و به‌طور گسترده استفاده می‌شوند که این روش‌های یادگیری جدید، با بهره‌گیری از رگرسیون‌های گروهی در حال پیدایش‌اند. یکی از انواع روش‌های یادگیری که از الگوریتم‌های پایه برای پیش‌بینی چندگانه تکراری<sup>۲</sup> استفاده می‌کند، Random Forest نامیده می‌شود [۷ و ۱۸]. این روش در طبقه‌بندی پوشش زمین که از داده‌های سنجش از دور به‌دست می‌آید و نیز در مسائل الکترونیک و پزشکی کاربرد فراوان دارد [۲۳ و ۲۵]. همچنین روش RF پتانسیل لازم به‌عنوان یک ابزار مدل مکانی برای ارزیابی آلودگی در مباحث زیست‌محیطی و منابع آب را داراست [۵]. RF روشی گروهی است که چند الگوریتم درختی را برای تولید یک پیش‌بینی مکرر از هر پدیده ترکیب می‌کند. RF می‌تواند الگوهای پیچیده را یاد بگیرد و ارتباط غیرخطی بین متغیرهای توضیحی و متغیرهای وابسته را در نظر بگیرد. همچنین می‌تواند انواع مختلف داده‌ها را در تجزیه و تحلیل بگنجاند و ترکیب کند که این هم به‌علت نبود توزیع پیش‌فرض‌ها (توزیع نرمال) درباره داده‌های استفاده‌شده است. RF هزاران متغیر ورودی را بدون حذف یکی از آنها می‌پذیرد و اجرا می‌کند، همچنین می‌تواند برآوردی از اینکه کدام

1. Machine learning  
2. Repeated Multiple Prediction

متغیر در پیش‌بینی مدل، مهم است را تشخیص دهد [۲۷]. RF نسبت به شبکه‌های عصبی مصنوعی در برابر گیرافتادن در مینیمم محلی و داده‌های پرت حساسیت کمتری دارد و می‌تواند تخمین بهتری از پارامترها داشته باشد. RF اهمیت نسبی متغیرها را ارزیابی می‌کند و قادر است متغیرهای مهم را نیز انتخاب کند و در عین حال پارامترسازی آن نسبت به روش‌های دیگر مانند شبکه‌های عصبی، محاسبات ساده‌تری دارد [۲۷]. برای جلوگیری از تطابق  $RT_s$  های مختلف، RF تنوع درختان را از طریق درست کردن زیرمجموعه‌های مختلف از داده‌های آموزشی کم می‌کند که اصطلاحاً کیسه‌بندی<sup>۱</sup> نامیده می‌شود. اصطلاح Bagging از مخفف Bootstrap Aggregating به دست آمده است [۹]. Bagging تکنیکی است که از طریق نمونه‌برداری دوباره تصادفی از مجموعه داده‌های اصلی و همراه با جایگزینی، برای ایجاد داده‌های آموزشی استفاده می‌شود و در این مرحله هیچ‌یک از داده‌های انتخاب‌شده از نمونه‌های ورودی را برای تولید زیرمجموعه بعدی، حذف نمی‌کند و بدین ترتیب واریانس نیز کاهش می‌یابد. از این رو، برخی داده‌ها ممکن است بیش از یک بار در شاخه‌های آموزشی استفاده شوند درحالی که برخی داده‌های دیگر که در مدل‌سازی مؤثر نیستند، هرگز استفاده نشوند. بنابراین، ثبات بیشتری برای مدل به دست می‌آید و مدل را در برابر تغییرات جزئی در داده‌های ورودی قابل اعتمادتر می‌سازد و دقت پیش‌بینی آن را افزایش می‌دهد [۸]. از سوی دیگر، هنگامی که RF، یک درخت رشد به وجود می‌آورد از بهترین متغیرها یا نقاط تقسیم داخل زیرمجموعه‌های متغیرها بهره می‌برد که به صورت تصادفی از مجموعه‌های کلی متغیرهای ورودی انتخاب می‌کند. بنابراین قدرت هر RT منفرد را کاهش می‌دهد و میزان تطابق را پایین می‌آورد و بدین صورت خطای کلی مدل را کاهش می‌دهد [۸]. این روش یک متالگوریتم است که برای بهبود یادگیری ماشین رده‌بندی و مدل‌های پس‌رقتی برحسب پایداری و دقت رده‌بندی به کار می‌رود. این روش همچنین واریانس را کاهش می‌دهد و به دوری از اورفیتینگ کمک می‌کند. اگرچه این روش در درخت تصمیم به کار می‌رود، می‌تواند در هر نوع مدل استفاده شود. Bagging حالت مخصوصی از روند مدل میانگین است. یکی دیگر از

ویژگی‌های خوب RF این است که درختان RF بدون پرونینگ یا هرس کردن رشد می‌کنند و در این روش آموزش بیش از اندازه بر دقت مدل تأثیر نمی‌گذارد بلکه آن را از دیدگاه محاسباتی سبک‌تر می‌کند. علاوه بر این، نمونه‌هایی که در آموزش درختان در فرایند Bagging انتخاب نمی‌شوند بخشی از زیرمجموعه‌هایی هستند که الگوهای خارج از کیسه<sup>۲</sup> نامیده می‌شوند و این قسمت در روش RF می‌تواند برای ارزیابی عملکرد مدل استفاده شود. ممکن است اطلاعات زیاد برای مدل‌سازی مفید باشند، اما افزایش تعداد پارامترهای ورودی پیچیدگی‌های اضافی و افزایش زمان محاسبات و مشکلات ابعادی را به سیستم تحمیل می‌کند [۴]. در تلاش برای رسیدن به یک تقسیم بهینه، تصمیم‌ساز<sup>۳</sup> می‌تواند به وسیله مقدار بسیار عظیمی از داده‌های مختلف، اطلاعات شرایطی که در منطقه موجود هستند را به سرعت تحت‌الشعاع قرار دهد. تعداد زیادی از متغیرهای وابسته به خواص و رفتار سیستم آب زیرزمینی و نیروهای راندگی، می‌توانند توانایی مدل را افزایش دهند. ابعاد یادشده در مجموعه داده‌ها می‌توانند سبب کاهش دقت مدل شوند. برای جلوگیری از این خطاها و ابعاد زیاد داده‌ها، انتخاب ویژگی<sup>۴</sup> به کار برده می‌شود. FS یک روش برای انتخاب زیرمجموعه‌های پارامترهای مختص به آن برای آموزش بهتر مدل است [۱۹]. در مطالعات آب زیرزمینی، از تعداد زیادی متغیرهای وابسته به خصوصیات فیزیکی و شیمیایی آبخوان استفاده می‌شود که برخی از آنها می‌توانند مرتبط یا غیرمرتبط باشند. بنابراین، در ارزیابی یک آبخوان نسبت به آلودگی، انتخاب متغیرهای توضیحی که ارتباط بیشتری با آلاینده دارند، خیلی مهم است. FS با بالابردن سرعت فرایند آموزش، افزایش قابلیت تعمیم، کاهش اثر از بین رفتن ابعاد و افزایش قابلیت تفسیر، دقت مدل‌های پیش‌بینی را افزایش می‌دهد. روش‌های زیادی برای FS بیان شده است. در مطالعه حاضر، مدل RF برای آبخوان دشت ملکان با استفاده از یک پایگاه داده‌های سیستم اطلاعات جغرافیایی شامل داده‌های خصوصیات هیدروژئولوژیکی آبخوان، فاصله از فعالیت‌های انسانی و منابع بالقوه تغذیه، متغیرهای سنجش از دور و خصوصیات فیزیکی و شیمیایی اندازه‌گیری شده در صحرا، توسعه داده شده است و

2. Out Of Bag  
3. Decision-Maker  
4. Feature Selection

1. Bagging

است که بیشتر از پادگانه‌های آبرفتی قدیمی، پادگانه‌های آبرفتی جدید، مخروط‌افکنه‌ها و رسوبات رودخانه‌ای تشکیل شده است و مواد اصلی تشکیل‌دهنده آبخوان، رسوبات ماسه، سیلت و رس است. سطح تراز و عمق آب زیرزمینی از رأس مخروط‌افکنه‌ها به سمت پای دشت کاهش می‌یابد. براساس نقشه‌های هم‌ضخامت رسوبات آبرفتی، لاگ‌های حفاری و داده‌های ژئوفیزیکی، در قسمت‌های بالایی دشت، ذرات تشکیل‌دهنده سفره دانه‌درشت است و هر چه به سمت مرکز دشت و نواحی خروجی و به سمت دریاچه ارومیه نزدیک می‌شویم، رسوبات دانه‌ریزترند.

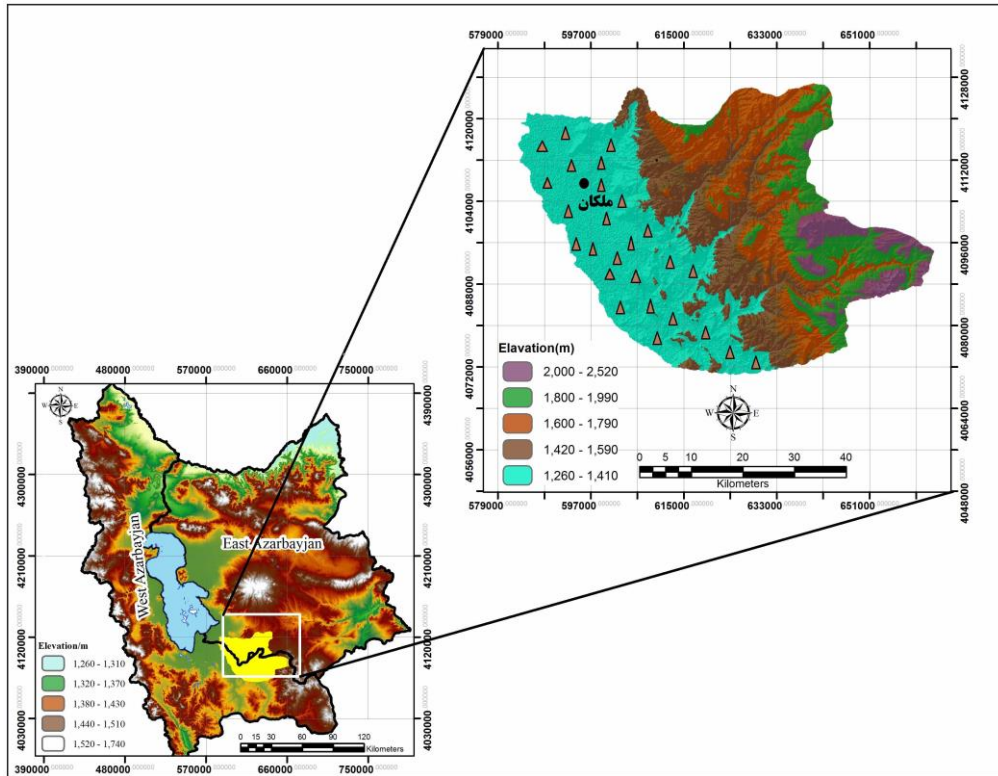
براساس بررسی‌های هیدروژئولوژیکی، رسوبات کواترنری بیشتر سازندهای آب‌دار منطقه را دربرگرفته‌اند و آبدهی خوبی دارند. این نهشته‌ها شامل رسوبات رودخانه‌ای (Qal) که در طرفین رودخانه‌های دائمی و فصلی به‌صورت کم شیب گسترش دارند، رسوبات واریزه‌ای Qt2 و Q1 که بیشتر در دامنه‌ها دیده می‌شوند و رسوبات Qt1 که تراس‌های مرتفع را تشکیل می‌دهند، هستند. سنگ‌های با توان آبدهی متوسط که بیشتر در سازندهای شمشک، لالون و واحدهای ماسه سنگی - شیلی کرتاسه یافت می‌شوند، قسمت اعظم منابع آب منطقه ملک‌ان را تشکیل می‌دهند. سازندهای با آبدهی متوسط به بالا نیز در واحدهای آهکی اربیتولین‌دار، تشکیلات تفکیک‌ناپذیر ژوراسیک و کرتاسه، واحدهای آهکی لار و دلچای یافت می‌شوند. همچنین گروه سنگ‌های تخریبی - آذرآواری با توان آبدهی ضعیف نیز که شامل واحد خاکستر آتشفشانی همراه با سنگ‌های آذرآواری، کنگلومرا و ماسه‌سنگ میوسن، واحد مارن‌های خاکستری با کمی سنگ آهک متعلق به کرتاسه و شیل و آهک دلومیتی متعلق به پرکامبرین هستند، حدود ۳۴ کیلومترمربع از منطقه را تشکیل می‌دهند. سایر سازندها وسعت زیادی نیز ندارد، آبدهی ضعیفی دارند. این سازندها که در قسمت‌های غربی به‌طور گسترده در حاشیه دریاچه ارومیه قرار گرفته‌اند، سن کواترنری (Qs) و شامل رسوبات رسی هستند که علاوه بر اینکه نفوذپذیری بسیار کمی دارند، به‌علت داشتن لایه‌های گچ و نمک روی کیفیت آب‌ها اثر نامطلوب می‌گذارند.

پتانسیل RF برای ایجاد یک نقشه آسیب‌پذیری نسبت به آرسنیک با توجه به معیارهای مختلف بررسی می‌شود که وابسته به تغییرات پارامترهای الگوریتم، صحت و دقت نقشه‌هاست. هدف کلی از این تحقیق، توسعه یک مدل دقیق و صحیح برای ارزیابی و پیش‌بینی آلودگی آب زیرزمینی نسبت به آرسنیک با استفاده از داده‌های مکانی مختص به خصوصیات ذاتی و خاص آبخوان، نیروهای محرک و خصوصیات فیزیکی - شیمیایی است.

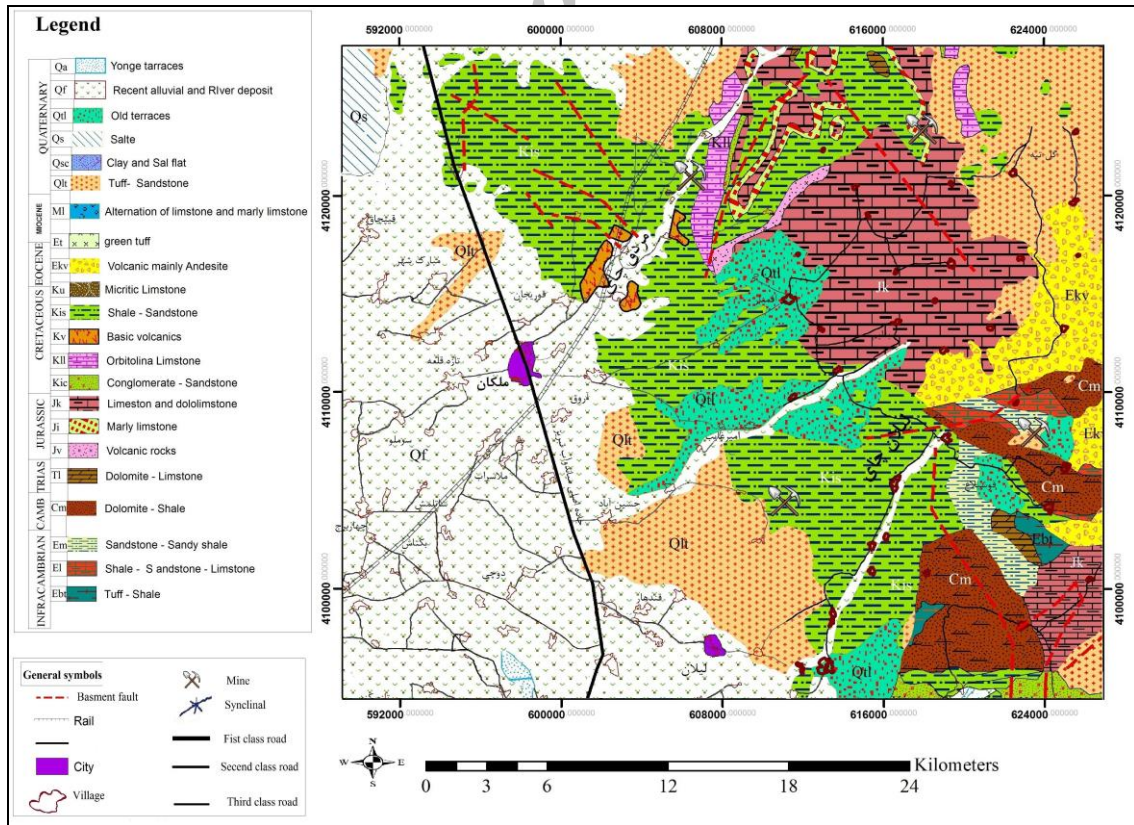
## مواد و روش‌ها

### منطقه مطالعه‌شده و زمین‌شناسی منطقه

دشت ملک‌ان با وسعتی حدود ۴۵۰ کیلومترمربع در جنوب استان آذربایجان شرقی و در جنوب شرق دریاچه ارومیه واقع شده است (شکل ۱) و جزء زون زمین‌ساختاری البرز - آذربایجان محسوب می‌شود. متوسط ارتفاع دشت ملک‌ان ۱۳۲۰ متر و اختلاف ارتفاع آن از رأس تا انتهای دشت ۴۵ متر است. دشت ملک‌ان از نظر تقسیمات طبیعی در حوضه آبریز دریاچه ارومیه قرار دارد و این حوضه براساس روش تجربی آمبرژه [۱۵] و با استفاده از آمار ایستگاه ملک‌ان اقلیم سرد و نیمه‌خشک دارد. منطقه ملک‌ان سازندهای زمین‌شناسی مختلفی دارد. سازند لالون که در قسمت شرق و سازند روتنه در دوره پرمین در جنوب منطقه برونزد دارد، سازند شمشک قسمت شرق و شمال شرقی منطقه را پوشانده است و سازند لار مختص به دوره تریاس و ژوراسیک در قسمت شمال و شمال شرق منطقه دیده می‌شود. همان‌طور که در شکل ۲ نشان داده شده است، بخش بزرگی از مساحت منطقه در بخش غربی مختص به رسوبات آبرفتی دوره کواترنر است و بخش کمی از آنها در بخش شمال غرب پهنه‌های رسی - نمکی دارند. در قسمت جنوب و جنوب شرقی دشت مجموعه رسوبات آهک و مارن - ژئوپس دیده می‌شود که مختص به دوره کرتاسه و مجموعه سنگ‌های پیروکلاستیک و رس‌سنگ‌ها مختص به دوره پلیوسن هستند. در قسمت شمال و شمال غربی دشت ملک‌ان، سنگ‌های آهکی تا توده خاکستری روشن مختص به دوره تریاس و ژوراسیک و توده آهک‌های کرتاسه به‌چشم می‌خورد. آبخوان دشت از نوع آزاد



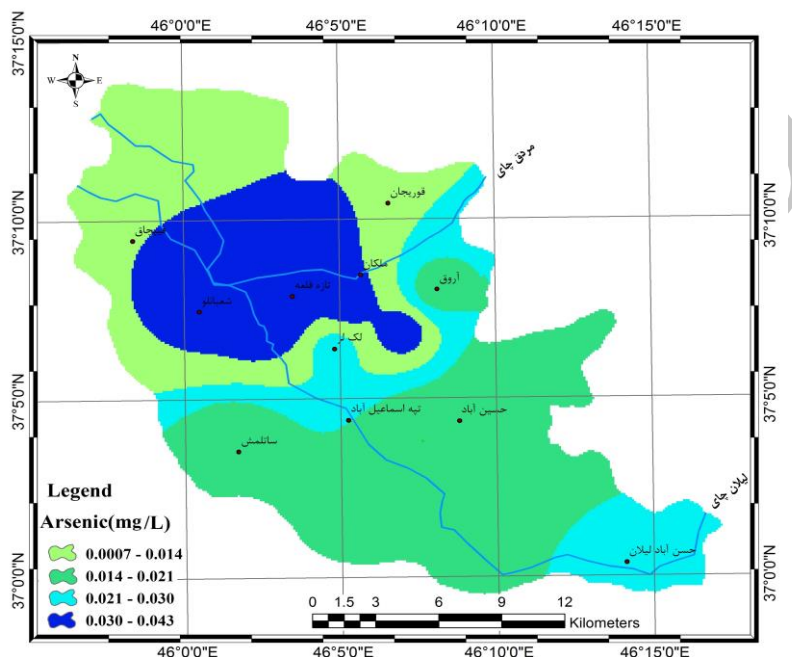
شکل ۱. موقعیت منطقه مطالعه شده و نقاط نمونه برداری



شکل ۲. نقشه زمین شناسی منطقه مطالعه شده

شبکه فاضلاب شهری و وجود چاه‌های جذبی زیاد، میزان آلودگی آب زیرزمینی، به خصوص میزان آرسنیک منطقه بیشتر از حد استاندارد جهانی (WHO) است (شکل ۳). خلاصه نتایج آنالیز برخی عناصر سنگین در جدول ۱ و همچنین خصوصیات آماری داده‌های هیدروشیمیایی در جدول ۲ آورده شده است (واحد غلظت یون‌ها برحسب میلی‌اکی‌والان بر لیتر).

عمق کم آب زیرزمینی در قسمت‌های کم‌ارتفاع دشت سبب تبخیر زیاد و افزایش شوری بیش از اندازه آب‌های زیرزمینی می‌شود، ضمن اینکه ریزدانه‌بودن رسوبات نیز موجب می‌شود که آب بر اثر نیروی کاپیلاری بالا بیاید و در عمق کم واقع شود و به تبع آن میزان تبخیر نیز افزایش یابد. به علت فعالیت شدید کشاورزی به‌ویژه وجود باغ‌های انگور و استفاده زیاد از کودهای شیمیایی و نبود



شکل ۳. توزیع مکانی آرسنیک دشت ملکان

جدول ۱. نتایج آنالیز برخی عناصر سنگین (میلی‌گرم بر لیتر)

X	Y	محل نمونه‌برداری	Mn	Fe	Cr	Al	As
۵۹۰۴۷۸	۴۱۱۸۷۴۱	چپقلو	۸/۹۶۵	۶/۳۲۱	۰/۱۰۲	۲/۳۳	۰/۲۸۵
۵۹۱۴۳۸	۴۱۰۱۹۲۲	دوچی	۴/۲۳۵	۲/۹۸۳	۰/۰۹	۲/۶۹	۰/۰۴۲
۵۹۹۱۹۲	۴۱۰۸۱۲۱	قره‌چال یوزباشکندی	۰/۴۲۳	۰/۵۸۵	۰/۰۳۲	۰/۵۳	۰/۳۷۱
۵۸۹۵۵۱	۴۱۰۹۱۲۹	شعبانلو	۰/۰۲۹	۰/۴۰۹	۰/۰۵۴	۰/۸۹	۰/۳۴۲
۶۰۹۹۱۰	۴۰۹۶۲۷۲	حسن‌آباد لیلان	۰/۰۴۷	۰/۷۴	۰/۰۴۹	۰/۲۳	۰/۲۲۵
۵۹۶۶۰۲	۴۱۰۹۵۸۴	ملکان	۰/۰۵۸	۰/۴۲	۰/۰۵۹	۱/۰۴	۰/۳۶۲
۶۰۰۸۶۷	۴۱۱۰۳۳۳	آروق	۰/۰۲۴	۰/۵۳۶	۰/۰۷۱	۰/۵۹	۰/۱۳۵
۵۹۴۱۵۸	۴۱۱۲۰۶۸	بایقوت	۰/۸۸۳	۱/۷۰۱	۰/۰۵۴	۱/۷۷	۰/۴۰۲
۵۹۲۹۸۰	۴۱۰۸۸۶۵	سرملو	۱/۸۸۷	۰/۴۵۵	۰/۰۴۵	۰/۹۵	۰/۴۲۵
۵۹۵۹۹۴	۴۱۰۷۲۳۲	لک‌ر	۷/۸۵۶	۱/۵۲۸	۰/۰۹۱	۱/۲۱	۰/۲۲۱
۶۰۰۱۷۸	۴۱۰۲۴۶۱	مهماندار	۰/۵۳۴	۰/۲۱۲	۰/۰۵۶	-۰/۲۹	۰/۰۰۷
۶۰۴۳۴۶	۴۱۰۴۸۷۸	دمیرچی	۰/۰۱۴	۰/۰۳۲	۰/۰۳۹	۰/۶۸	۰/۰۲۹

جدول ۲. خصوصیات آماری داده‌های هیدروشیمیایی

Parameter	Min	Max	Mean	Std. Deviation	Skewness	Variance	Kurtosis
Mn	۰/۰۰۱	۰/۸۹۷	۰/۲۰۸	۰/۳۲۰	۱/۳۷۵	۰/۱۰۳	۰/۲۷۹
Fe	۰/۰۰۳	۰/۶۳۲	۰/۱۳۳	۰/۱۷۸	۲/۰۷۰	۰/۰۳۲	۳/۳۳۵
Cr	۰/۰۰۳	۰/۰۱۰	۰/۰۰۶	۰/۰۰۲	۰/۵۸۴	۰/۰۰۰	-۰/۸۵۶
Al	-۰/۰۲۹	۰/۲۶۹	۰/۱۰۵	۰/۰۸۵	۱/۵۵۱	۰/۰۰۷	-۰/۳۷۶
As	۰/۰۰۱	۰/۰۴۳	۰/۰۲۴	۰/۰۱۵	-۰/۳۵۱	۰/۰۰۰	-۱/۳۸۴
Ca	۲/۸۸۳	۲۴/۵۷۰	۱۱/۶۶۲	۷/۴۸۲	۰/۷۶۹	۵۵/۹۷۶	-۰/۸۲۸
Mg	۰/۶۴۰	۶/۶۴۰	۳/۷۱۷	۱/۵۶۰	-۰/۲۳۳	۲/۴۳۳	۰/۱۶۴
K	۰/۱۲۷	۰/۶۳۵	۰/۲۵۵	۰/۱۳۶	۱/۹۹۳	۰/۰۱۹	۳/۱۸۸
Na	۱/۳۸۳	۷۲/۳۱۵	۱۸/۶۴۷	۲۰/۹۸۰	۱/۵۸۰	۴۴۰/۱۷۰	۱/۵۱۷
Hco <sub>3</sub>	۳/۱۹۰	۲۸/۶۵۰	۷/۵۷۰	۶/۷۷۳	۲/۸۰۶	۴۵/۸۶۸	۶/۳۳۹
So <sub>4</sub>	۲/۵۳۸	۱۱/۶۴۰	۵/۵۶۹	۲/۶۶۶	۱/۰۷۵	۷/۱۰۹	۰/۳۷۳
No <sub>3</sub>	۰/۰۲۰	۲/۳۴۵	۰/۶۹۲	۰/۷۳۳	۰/۹۲۵	۰/۵۳۷	-۰/۰۱۹
Cl	۳/۵۳۱	۶۵/۷۶۵	۲۱/۲۳۲	۲۰/۸۴۲	۰/۹۷۶	۴۳۴/۳۷۰	-۰/۳۶۳
Ec	۹۵۰/۰	۱۰۲۹۵	۳۲۹۱/۳	۳۰۴۴/۰	۱/۲۵۶	۹۲۶۵۸۰۰/۰	۰/۲۵۹
pH	۶/۲۸۰	۷/۶۷۰	۷/۰۹۱	۰/۴۴۵	-۰/۵۵۴	۰/۱۹۸	-۰/۶۴۴
TH	۵۰/۱/۹	۵۶۰/۸/۴	۱۷۸۱/۳	۱۶۶۳/۴	۱/۲۵۶	۲۷۶۶۸۰۰/۰	۰/۲۵۹
Sio <sub>2</sub>	۰/۱۹۰	۲/۱۲۳	۱/۱۴۳	۰/۶۰۰	-۰/۰۶۴	۰/۳۶۰	-۰/۹۶۶
Temp	۱۱/۲	۱۹	۱۵/۴۶	۲/۰۱	-۰/۳۹	۴/۰۶	۰/۰۹۷
TDS	۷۱۲/۵۰	۷۷۲۱/۲	۲۴۶۸/۴	۲۲۸۳/۰	۱/۲۵۶	۰۵۲۱۲۰۰۰/۰	۰/۲۵۹

سبب تولید خطر می‌شوند)، محیط خاک، متغیرهای سنجش از دور و پوشش زمین، شاخص تفاضلی نرمال شده پوشش گیاهی<sup>۱</sup> و خصوصیات ذاتی آبخوان به دست آورده شدند. جدول ۳ متغیرهای توضیحی و متغیر پاسخ را که همان غلظت آرسنیک است، بیان می‌کند که منابع داده‌ها و روش‌های برآورد نیز آورده شده است. به منظور به دست آوردن متغیرهای پیوسته و استاندارد شده، برای تمام منطقه مطالعه شده، همه داده‌ها به فرمت رستری تبدیل شدند که در این مرحله از سه روش استفاده شد: الف) تکنیک‌های زمین‌آماری برای به دست آوردن نقشه‌های هدایت الکتریکی، گرادیان هیدرولیکی، بافت خاک و...، که انتخاب روش زمین‌آماری و میان‌یابی براساس نتایج نیم‌تغییرنماها انجام گرفته است. بعد از رسم نیم‌تغییرنما برای ویژگی‌های مطالعه شده، روش میان‌یابی انتخاب شد. به دلیل اینکه در این تحقیق، مدل‌های نیم‌تغییرنمای

#### طراحی اطلاعات و فراخوانی مدل آلاینده

در مطالعه حاضر از ۲۷ چاه مد نظر با پراکندگی مناسب نمونه‌برداری شد. همچنین، تعداد ۱۲ نمونه برای بررسی وضعیت پراکنش یون‌های فرعی برداشته شد. نمونه‌های برداشتی به منظور تجزیه عناصر اصلی و فلزات سنگین در ظروف پلی اتیلنی جداگانه صورت گرفت که به نمونه‌های برداشتی برای تجزیه فلزات سنگین، ۲ میلی لیتر اسید نیتریک با غلظت ۶۵ درصد (غلظتی که pH را به کمتر از ۲ برساند) به عنوان نگه‌دارنده اضافه شد. همچنین، بعضی پارامترهای فیزیکوشیمیایی (pH، EC، T) در محل نمونه‌برداری اندازه‌گیری شد. آنالیزهای هیدروشیمیایی ۲۷ نمونه در آزمایشگاه آب‌شناسی دانشگاه تبریز انجام شد و عناصر سنگین (آهن، آلومینیوم، منگنز، آرسنیک، کروم)، با روش جذب اتمی- کوره گرافیتی در آزمایشگاه کنترل کیفی آب استان آذربایجان شرقی آنالیز و اندازه‌گیری شدند. همچنین، متغیرهای دیگری مانند نیروهای راندگی (بعضی از فعالیت‌هایی که به صورت نقطه‌ای یا غیرنقطه‌ای

1. Normalized Difference Vegetation Index

زیرزمینی بالا سبب تسریع انتقال آلاینده به داخل آبخوان می‌شود). ۳. بافت خاک به دلیل وجود رس یا شن و ماسه که انتقال آلاینده را کنترل می‌کند. ۴. میزان قابلیت انتقال آبخوان و هدایت هیدرولیکی آن که در سرعت آب و انتقال آلاینده نقش دارند. ۵. گرادیان هیدرولیکی که در جهت انتقال آلاینده و میزان مهاجرت آلاینده مؤثر است. ۶. متغیرهای سنجش از دور نیز می‌توانند برای ارزیابی پتانسیل رواناب (نوع و میزان پوشش گیاهی موجود در منطقه بر میزان رواناب تأثیر دارد)، منابع پراکنده و آلاینده‌های کشاورزی کمک کنند. در این تحقیق برای غلبه بر مشکلات یادگیرنده‌های پایه، الگوریتم Random Forest که یک روش یادگیری مبتنی بر دسته‌ای از درخت‌های تصمیم است، پیشنهاد می‌شود.

بیشتر ویژگی‌های مطالعه‌شده، خطی و با شیب کم به دست آمده بود، از روش IDW با توان ۲ برای میان‌یابی ویژگی‌ها استفاده شد؛ ب) محاسبات فاصله اقلیدسی رستری برای به دست آوردن لایه‌های پتانسیل منابع نقطه‌ای آلودگی؛ ج) طبقه‌بندی پوشش زمین از سنجش از دور و NDVI با تصاویر ماهواره‌ای. برخی ویژگی‌هایی مهم از خصوصیات ذاتی منطقه که بیان‌کننده انتخاب مدل رگرسیونی RF است به صورت زیر توضیح داده شده است: ۱. شیب سطح زمین که نشان می‌دهد آیا رواناب روی سطح باقی خواهد ماند تا به نفوذ آلاینده به داخل زون اشباع کمک کند. ۲. سطح آب زیرزمینی و ضخامت زون وادوز نشان می‌دهند آیا شست‌وشوی عناصر آلاینده به داخل زون وادوز سریع اتفاق می‌افتد یا اینکه دیرتر صورت می‌گیرد (سطح آب

جدول ۳. متغیرهای توضیحی و متغیر پاسخ، همراه با منبع و روش استفاده‌شده در برآورد لایه رستری

نوع متغیر	نام متغیر	داده‌های استفاده‌شده	روش استفاده‌شده
متغیرهای توضیحی خصوصیات سفره	تراز آب زیرزمینی ارتفاع سنگ کف آبخوان ضخامت اشباع قابلیت انتقال آبخوان، هدایت هیدرولیکی گرادیان هیدرولیکی ضخامت زون وادوز محیط خاک، محیط آبخوان لایه تغذیه	اطلاعات ۳۰ پیزومتر موجود در منطقه اطلاعات ژئوفیزیک و گمانه‌های حفاری تراز سطح آب و ارتفاع سنگ کف استفاده از اطلاعات ۲۷ چاه پمپاژشده در منطقه منحنی‌های تراز آب زیرزمینی و فاصله بین منحنی‌ها عمق سطح آب زیرزمینی لاگ‌های حفاری منطقه لایه‌های شیب، بارندگی و نفوذپذیری خاک	روش‌های میان‌یابی معکوس وزنی فاصله و کریجینگ پهنه‌بندی لایه‌ها با روش‌های زمین‌آمار و استفاده از روش Pisco
متغیرهای توضیحی فلزات سنگین	کروم، منگنز، آهن و آلومینیوم	آنالیز ۲۷ نمونه برداشتی از آب زیرزمینی	روش میان‌یابی معکوس وزنی فاصله
متغیرهای توضیحی نیروهای محرک	فاصله از شهرها، رودخانه‌ها، باغ‌های انگور، محل‌های دفن زباله	تصاویر ماهواره‌ای و گوگل ارث، نقشه‌های با مقیاس کم و مختصات برداشتی نقاط انتخابی	اندازه‌گیری فاصله اقلیدسی با استفاده از GIS
متغیرهای توضیحی سنجش از دور	شاخص NDVI، شیب سطح زمین	تصاویر ماهواره‌ای سال ۲۰۱۴ ماه نوامبر Landsat8	طبقه‌بندی پوشش زمین با داده‌های Landsat8
متغیرهای توضیحی فیزیکی-شیمیایی	pH، هدایت الکتریکی، دما	اندازه‌گیری‌های صحرایی	استفاده از روش میان‌یابی معکوس وزنی فاصله
متغیر پاسخ	آرسنیک	آنالیز ۲۷ نمونه برداشتی از آب زیرزمینی	پهنه‌بندی با استفاده از روش میان‌یابی معکوس وزنی فاصله

ساخته شد. به دلیل اینکه با افزایش درخت‌ها میزان کاهش می‌یابد، تعداد ۵۰۰ درخت برای فراخوانی مدل استفاده شد. پارامتر  $m$  نیز به وسیله تغییر تعدادی متغیرهای تقسیم بین یک و ماکزیمم متغیرهای هر زیرمجموعه، بهینه شد. نتایج مدل به وسیله برآورد خطای

### طراحی مدل

نخستین مرحله در ایجاد مدل پیش‌بینی آلودگی به شبه‌فلز آرسنیک، تنظیم پارامترهای مدل است. به منظور تنظیم مقدار  $k$ ، طوری که مقدار خطا همگرا شود و تخمین قابل اعتمادتر باشد، مدل از ۱ تا ۵۰۰ درخت،

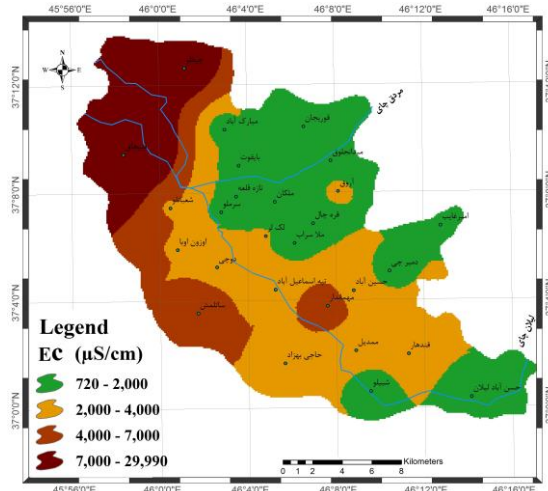
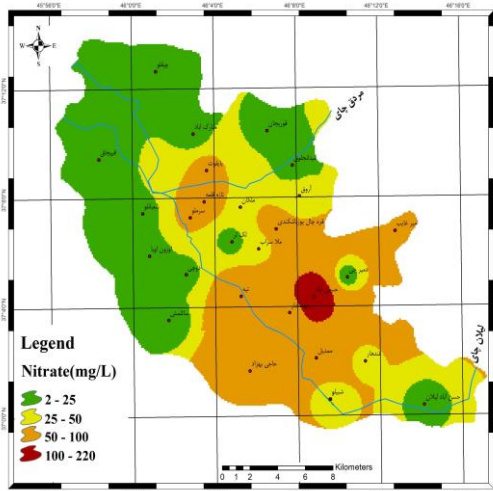


ضریب ذخیره این دشت که با روش بیلان جزء-به-جزء محاسبه شده به طور متوسط حدود ۳/۲ درصد است. مقدار گرادیان هیدرولیکی مطلق که مسئول حرکت آب زیرزمینی و مهاجرت ماده حل شونده در آن است در دشت ملکان در نقاط مختلف متفاوت و به طور کلی در ورودی‌های دشت میانگین ۰/۰۰۶ و در خروجی‌ها نیز حدود ۰/۰۰۱ است. یون‌هایی مانند آهن، منگنز و آلومینیوم نیز در آزادسازی آرسنیک در آب اثرگذارند. جذب آرسنیک (As (V)) روی سطوح کانی‌ها فرایند مهمی است که بر تحرک و تثبیت آرسنیک تأثیر می‌گذارد. گروه‌های هیدروکسی (به‌ویژه آهن و منگنز) روی سطح بیشتر کانی‌ها زیانند و مکان‌های جذبی و واکنشی هستند؛ بنابراین اکسیدها و هیدروکسیدهای آهن ارتباطی قوی با آرسنیک (As (V)) دارند [۲۸]. ولی حضور آهن در غلظت‌های خیلی زیاد، مانع جدایش آرسنیک از آب می‌شود و در نتیجه به‌عنوان جاذب عمل می‌کند. بنابراین، همبستگی منفی بین آهن و منگنز و آرسنیک در نمونه‌های آبی قابل مشاهده است. وزن‌دهی پارامترها براساس تأثیر در احتمال آلودگی آرسنیک با یک کلاس‌بندی مناسب انجام گرفت. به‌طور مثال، کلاسی که در آن فاصله از شهر کمتر باشد، تأثیر بیشتری بر آلودگی دارد، بنابراین هر چه به شهرها نزدیک می‌شویم، وزن بیشتری اختصاص داده شده است و هر چه این فاصله بیشتر شود، وزن کمتری به خود می‌گیرد. یا در بحث خواص هیدرولیکی آبخوان، هر چه مقدار قابلیت انتقال، هدایت هیدرولیکی و گرادیان هیدرولیکی بیشتر باشند، براساس چارچوب هیدروژئولوژیکی وزن بیشتری به خود می‌گیرند و هر چه کمتر باشند، به‌دلیل اینکه در احتمال آلودگی تأثیر کمتری دارند وزن کمتری به آنها نسبت داده می‌شود. در نهایت، براساس امتیاز پیش‌بینی‌شده توسط مدل، همپوشانی تمام لایه‌ها انجام می‌گیرد و مکان‌هایی را که احتمال تجاوز غلظت آرسنیک بیش از حد استاندارد جهانی وجود دارد، شناسایی می‌شوند.

OOB ارزیابی شد. علاوه بر این، برای کاهش ابعاد و افزایش دقت و قابل تفسیر بودن مدل، روش FS به‌کار برده شد و متغیرهای مهم در پیش‌بینی نیز شناسایی شدند. برای فراخوانی مدل آلاینده فلزات سنگین، داده‌های به‌دست‌آمده از لایه‌های رستری ۲۴ متغیر توضیح داده‌شده که برای هر پیکسل از لایه رستری یک مقدار استخراج شده بود با هم و با متغیر هدف استفاده شد. حد مجاز آرسنیک برای تمایز نقاط آلوده و غیرآلوده براساس اعلام سازمان بهداشت جهانی [۳۱]، ۰/۰۱ میلی‌گرم بر لیتر در نظر گرفته شد. غلظت آرسنیک به متغیر پاسخ جدید برای هر نمونه تجربی دوباره مقیاس‌بندی شد. نمونه‌های با غلظت مساوی یا بیشتر از حد آستانه (۰/۱ میلی‌گرم بر لیتر) مساوی یک و نمونه‌های کمتر از حد آستانه مساوی صفر در نظر گرفته شد. متغیرهای توضیحی (پیش‌بینی‌کننده) و متغیر پاسخ با هم در مجموعه‌ای از بردارهای ویژگی ورودی ترکیب شدند. این بردارها ورودی الگوریتم RF را تشکیل دادند و به‌عنوان بردارهای Input-feature شناخته شدند. متغیر پاسخ دوتایی (متغیر توضیحی و متغیر پاسخ) به‌عنوان مقادیر هدف برای آموزش الگوریتم استفاده شدند.

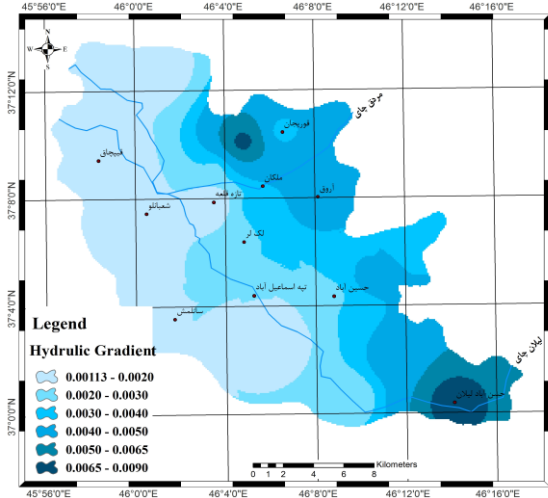
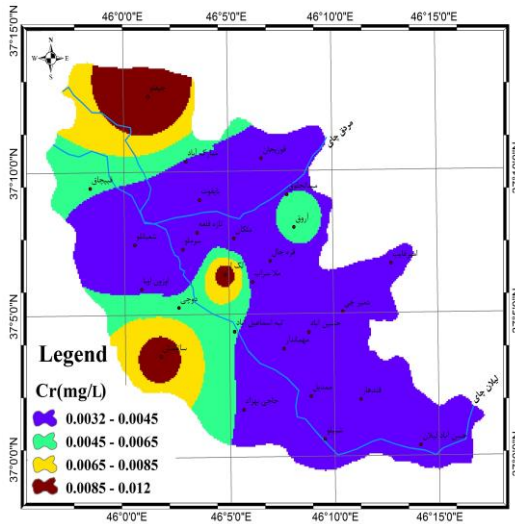
### بحث و نتایج

در منطقه ملکان که یکی از قطب‌های تولید انگور کشور محسوب می‌شود، هر ساله برای افزایش بازدهی محصول باغ‌ها از کودهای شیمیایی و حیوانی با مقدار بیش از اندازه لازم استفاده می‌شود که این کودها مدام آلاینده‌های زیادی به سفره آب زیرزمینی وارده می‌کنند و پتانسیل آلودگی را در منطقه افزایش می‌دهند. همچنین به‌دلیل وجود چاه‌های جذبی و نبود شبکه فاضلاب در شهر ملکان، فاضلاب شهری نیز در گسترش آلودگی تأثیر زیادی دارد. شکل ۴ برخی لایه‌های رستری متغیرهای توضیحی را که توسط روش RF به‌عنوان متغیر مهم شناسایی شدند را نشان می‌دهد که در لایه‌های ایجادشده توزیع مکانی متغیرهای توضیحی مشاهده می‌شود. با توجه به نقشه قابلیت انتقال، بیشترین قابلیت انتقال در مرکز دشت است.



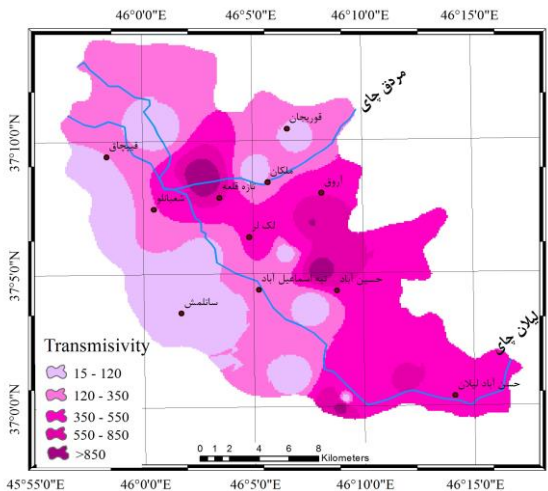
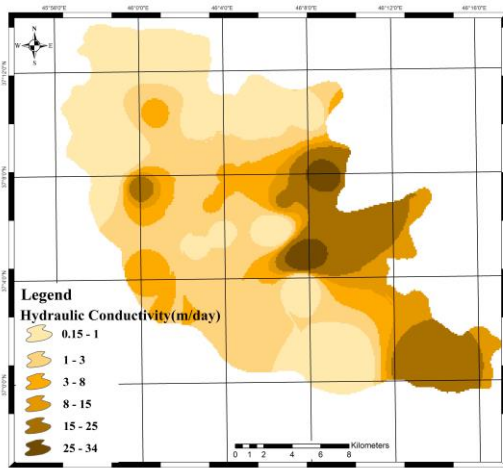
الف

ب



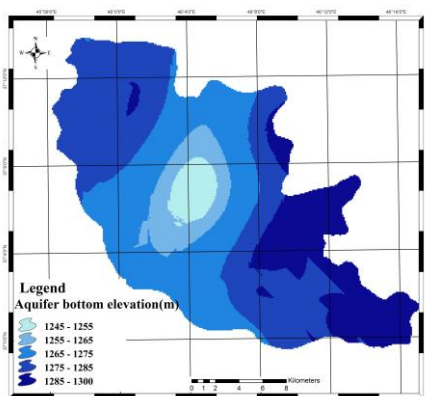
ج

د

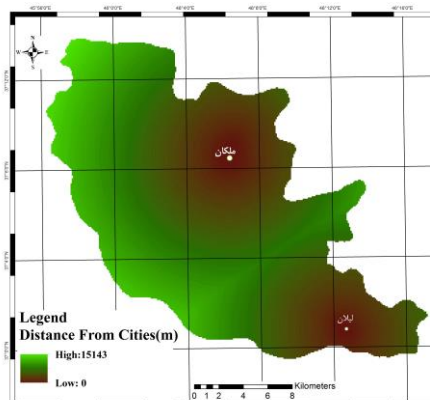


ه

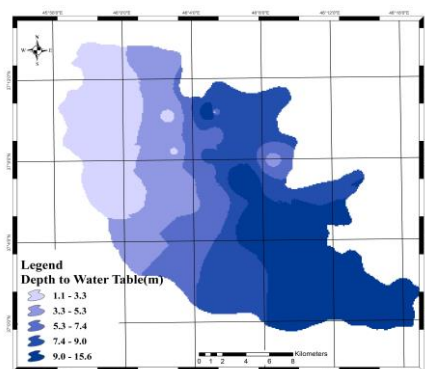
ز



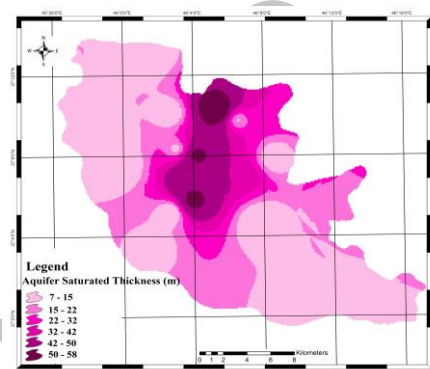
ن



و



ی



م

شکل ۴. برخی لایه‌های رستری مهم در پیش‌بینی مدل: الف) نیترات؛ ب) هدایت الکتریکی؛ ج) کروم؛ د) گرادیان هیدرولیکی؛ ه) هدایت هیدرولیکی؛ ر) قابلیت انتقال؛ و) ارتفاع سنگ کف؛ ن) فاصله از شهرها؛ م) عمق سطح ایستابی؛ ی) ضخامت اشباع آبخوان

دارند که به وجود آرسنیک در قسمت‌های جنوب شرقی منطقه تأثیر می‌گذارند.

در این میان تأثیر هریک از این آنیون‌ها در آزادسازی توانایی تشکیل کمپلکس برای جداسازی آرسنیک جذب‌شده به سطح هیدروکسیدهای آهن، منگنز و آلومینیوم متفاوت است، به طوری که بی‌کربنات سدیم با بیشترین توانایی و فسفات و نیترات به ترتیب در رتبه‌های بعدی قرار دارند [۱۰]. در شرایط کاهش متوسط، قابلیت انحلال آرسنیک می‌تواند به وسیله انحلال این اکسی-هیدروکسیدها کنترل شود. اکسی-هیدروکسیدها به سه طریق بر آزادسازی آرسنیک مؤثر است: الف) انحلال لایه‌های حاوی آرسنیک که سبب آزاد شدن آرسنیک می‌شود؛ ب) جذب آرسنیک روی اکسی-هیدروکسیدها؛ ج) آزاد شدن آرسنیک تحت pH بالا. ولی حضور آهن در غلظت‌های بسیار زیاد، مانع از جدایش آرسنیک به آب می‌شود و در نتیجه به‌عنوان جاذب عمل می‌کند، بنابراین همبستگی منفی بین آهن و منگنز و آرسنیک در

#### منشأ آلودگی آرسنیک

وجود مقادیر زیادی از کانی‌های آبدار مانند آمفیبول و بیوتیت در توده آذرین منطقه، نشانگر توانایی ماگما در تولید سیالات گرمایی است. با وجود این، عموماً در کانسارهای اپی‌ترمال جزء متئوریک مهم‌تر از جزء ماگمایی است، ولی شواهد ماگمایی بیانگر وجود جزء ماگمایی علاوه بر جزء متئوریک است که در زمینه دم‌های سه‌پند است که سازندهای یادشده را تحت تأثیر قرار می‌دهد. همچنین در منطقه انواع دگرسانی‌های سیلیسی شدن، دولومیتی شدن، پیریتی شدن، دکلسی فیکاسیون با شدت و ضعف مختلف دیده می‌شود که تمرکز آرسنیک بیشتر در زمینه زون‌های دگرسانی است. در محدوده مطالعاتی شکستگی‌های کوچک مهم‌ترین کنترل‌کننده‌های ساختاری در امر تثبیت کانی‌های آرسنیک‌دار هستند. وجود این شکستگی‌ها به‌عنوان معابری برای حرکت محلول‌های هیدروترمال تحت فشار بوده و در تمرکز کانی‌های آرسنیک‌دار نقش اساسی

فاصله از کانال‌های آبیاری تأثیری در پیش‌بینی مدل ندارند، اما می‌توانند در برآورد خطای پیش‌بینی کمک کنند. همچنین نبود همبستگی بین آرسنیک با بی‌کربنات نشان می‌دهد بی‌کربنات در آزاد کردن آرسنیک به آب زیرزمینی منطقه دخالتی ندارد و ممکن است هر دو از منشأهای مختلفی باشند [۲۶].

انتخاب یک مدل پیش‌بینی براساس سه پارامتر دقت مدل پیش‌بینی، تعداد متغیرهای لازم و در دسترس بودن داده‌ها و دقت این داده‌ها پایدار خواهد بود. برای ارزیابی خطای مدل از منحنی‌های ROC و مقدار خطای MSE استفاده می‌شود. منحنی‌های ROC با تغییر مقدار مجاز آستانه در برابر خروجی پیش‌بینی رسم می‌شوند. به‌طور کلی، در منحنی‌های ROC نتایج  $FPR^3$  در محور X در مقابل TPR در محور Y رسم می‌شود. هر یک از نتایج آستانه در جفت (TPR- FPR) و یک سری از چنین جفت‌هایی برای رسم منحنی‌های ROC استفاده می‌شوند که TPR به‌عنوان حساسیت<sup>۴</sup> و FPR - (به‌عنوان اختصاصیت<sup>۵</sup> شناخته می‌شوند. مقدار آستانه زمانی روی خروجی تنظیم خواهد شد که مشخص شود در اینجا آلودگی وجود دارد یا ندارد. اگر احتمال، بزرگ‌تر از مقدار آستانه باشد، طبقه پیش‌بینی ۱ یا آلوده خواهد بود و اگر کمتر از مقدار آستانه باشد، طبقه پیش‌بینی صفر یا آلوده در نظر گرفته می‌شود. زمانی که حساسیت افزایش می‌یابد، مقدار اختصاصیت کاهش یا مقدار FPR افزایش می‌یابد. بنابراین، آستانه بهینه انتخاب می‌شود و برای طبقه‌بندی نقشه‌ها در تمایز نقاط آلوده و غیرآلوده استفاده می‌شود. همچنین مساحت زیرمنحنی<sup>۶</sup> در منحنی‌های ROC نیز به‌عنوان یکی از روش‌های ارزیابی خطا به کار برده می‌شود که هرچه میزان AUC به یک نزدیک باشد، مدل صحت بیشتری دارد. مدلی که کمترین مقدار خطای MSE را داشته باشد به‌عنوان بهترین و دقیق‌ترین مدل شناخته می‌شود که در مدل مد نظر MSE برابر با ۰/۰۲۰ است و دقت زیادی دارد. شکل ۵ میزان خطاهای به‌دست‌آمده از مدل پیش‌بینی و نتایج آنالیز ROC را برای مدل نشان می‌دهد که با رسم TPR (حساسیت) در مقابل FPR (حساسیت -) به‌دست می‌آید. همان‌طور که مشاهده

نمونه‌های آبی قابل مشاهده است. تحرک آرسنیک را می‌توان به انحلال کاهشی اکسید- هیدروکسید آهن و آزاد شدن هم‌زمان آرسنیک جذب‌شده و هم‌نهبستی، مرتبط دانست. شرایط کاهشی در محیط‌های آبی احتمالاً به‌وسیله اکسیدشدن مواد آلی حاضر در رسوبات به‌وجود می‌آیند [۱]. بنابراین، محیط‌های آبی متأثر از آلودگی مواد آلی، با غلظت زیادی از آرسنیک حل‌شده به‌عنوان نتیجه انحلال کاهشی کانی‌های سولفات آرسنیک‌دار است [۲۸]. با توجه به همبستگی منفی آرسنیک با آهن و همبستگی نسبتاً پایین سولفات تأثیر این فرایند در آزادسازی آرسنیک در منابع آبی منطقه کم‌اهمیت به‌نظر می‌رسد. مهم‌ترین بخش آلودگی آرسنیک در منطقه مختص به دفع سنتی فاضلاب‌های شهری است. به‌دلیل نبود شبکه فاضلاب شهری و وجود چاه‌های جذبی زیاد، میزان آلودگی آب زیرزمینی، به‌خصوص میزان آرسنیک منطقه بیشتر از حد استاندارد جهانی (WHO) است که چون جهت جریان آب زیرزمینی به سمت غرب منطقه است، میزان آلودگی نیز به سمت غرب گسترش می‌یابد.

#### نتایج مدل‌سازی

جدول ۴ سهم هر متغیر توضیحی<sup>۱</sup>، همراه با میزان امتیاز<sup>۲</sup> پیش‌بینی کرده برای هر یک از متغیرها را در مدل نشان می‌دهد که به‌عنوان یکی از خروجی‌های مدل است. در این مدل مقدار قابلیت انتقال آبخوان، میزان نیترات، هدایت هیدرولیکی و فاصله از شهرها، نسبت به سایر پارامترها اهمیت بیشتری در مدل پیش‌بینی دارد و هرچه این مقدار بیشتر باشد، پتانسیل آلودگی در منطقه بیشتر است. در دشت ملکان با توجه به اینکه در قسمت‌های خروجی دشت به‌دلیل ریزدانه‌بودن رسوبات که سبب می‌شود آب بر اثر نیروی کاپیلاری بالا بیاید و در عمق کمتر واقع شود، اما به‌دلیل اینکه این مناطق هدایت هیدرولیکی، قابلیت انتقال و گرادیان هیدرولیکی کمتری دارند و از مناطق شهری، صنعتی و کشاورزی دور هستند، میزان آسیب‌پذیری این مناطق به آلودگی آرسنیک نسبت به سایر مناطق کمتر است. همچنین به تبع نزدیک‌بودن آب زیرزمینی به سطح زمین میزان هدایت الکتریکی بر اثر تبخیر زیاد، بالاست و در این مناطق میزان هدایت الکتریکی رابطه عکس با آلودگی آرسنیک دارد. در این مدل متغیرهای مختص به شیب، کاربری اراضی و

3. False positive rate  
4 Sensitivity  
5. Specificity  
6. Area Under Curve

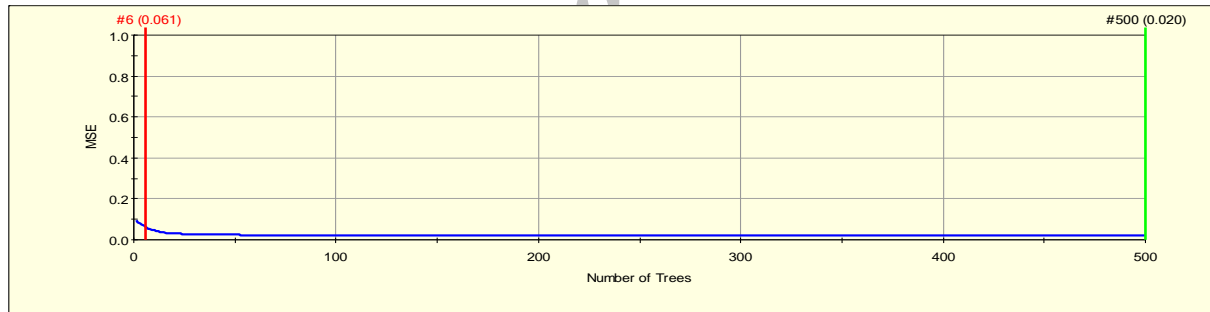
1. Variable importance  
2. Score

مد نظر با داشتن بیشترین سطح زیرمنحنی (AUC) برابر ۰/۹۹۷ و نشان دهنده دقت بیشتر مدل سازی است.

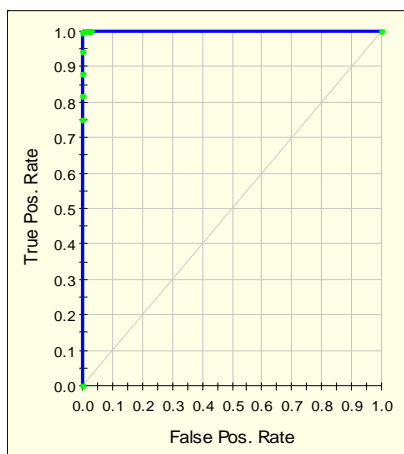
می شود میزان خطای به دست آمده از مدل با افزایش درخت های تصمیم به کمترین مقدار خود می رسد که دقت پیش بینی مدل را زیاد می کند و همچنین منحنی ROC مدل

جدول ۴. سهم هر متغیر توضیحی در مدل پیش بینی

Variable	Score	Variable Importance	Variable	Score	Variable Importance
قابلیت انتقال آبخوان	۱۰۰		هدایت الکتریکی	۱۷/۲	
نیترات	۹۷/۲		ارتفاع سنگ کف آبخوان	۱۶/۸	
هدایت هیدرولیکی	۶۷/۸		محیط آبخوان	۱۳/۱	
فاصله از شهرها	۵۰/۳		دما	۱۲/۴	
عمق سطح ایستابی	۴۴/۲		آلومینیوم	۸/۶	
گرادبان هیدرولیکی	۳۰/۲		فاصله از رودخانه ها	۷/۷	
ارتفاع سطح آب زیرزمینی	۲۹/۵		محیط خاک	۵/۲	
ضخامت اشباع آبخوان	۲۲/۷		شاخص NDVI	۴/۵	
pH	۲۲/۲		تغذیه	۱/۹	
آهن	۲۱/۴		فاصله از باغ های انگور	۱/۷	
منگنز	۱۹/۲		فاصله از محل دفن زباله	۱/۰	
کروم	۱۸/۸		شیب	۰/۹	



(الف)

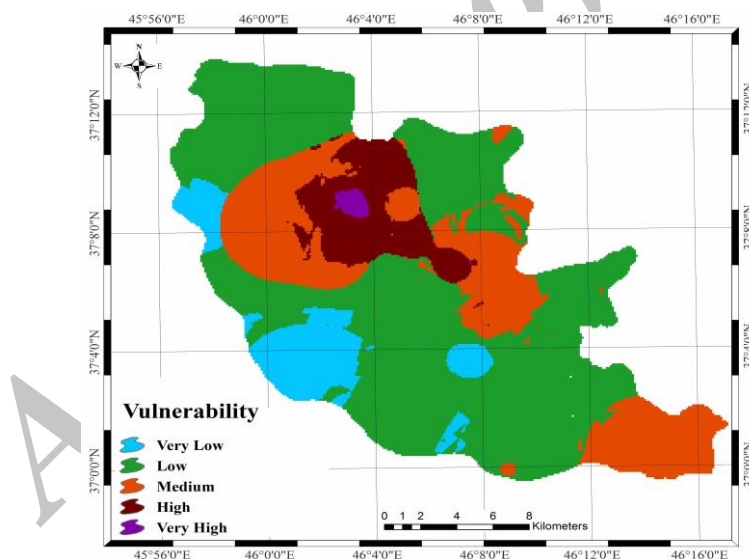


(ب)

شکل ۵. الف) خطای مدل و کاهش آن با افزایش درخت های تصمیم؛ ب) منحنی ROC به دست آمده از مدل پیش بینی

دشت به خصوص در منطقه شهری ملکان و همچنین به دلیل نبود شبکه فاضلاب و وجود چاه‌های جذبی زیاد در این مناطق از دشت، آلودگی آرسنیک و احتمال تجاوز از غلظت مجاز در این مناطق بیشتر است. همچنین آلودگی قسمت جنوب شرقی آبخوان را با اینکه به منطقه شهری نزدیک است، به دلیل اینکه جهت شیب آب زیرزمینی به سمت پایین‌دست است و عملاً در قسمت‌هایی پایینی آلودگی مشاهده نشد، نمی‌توان به وجود منطقه شهری نسبت داد. براساس مطالعات صحرایی که انجام شد، آلودگی این قسمت مختص به منشأ زمین‌زادی آن به صورت سولفیدهای آرسنیک درون شکستگی‌های موجود در مارن‌های میوسن باشد. روش RF با شناسایی مناطق آسیب‌پذیر و همچنین با تشخیص و کنترل عوامل بیرونی تأثیرگذار بر آلودگی آب زیرزمینی، نقش بسزایی در مدیریت کیفی آب زیرزمینی ایفا می‌کند. با استفاده از نتایج این مدل می‌توان حریم کیفی منابع آب زیرزمینی را تعیین و مدیریت مناسبی برای کاربری اراضی مرتبط با سفره آبدار اعمال کرد.

نتایج آسیب‌پذیری با مدل‌های صحیح از روش RF می‌تواند در مدیریت کیفی سفره‌های آب زیرزمینی بسیار کاربرد داشته باشد و با شناسایی مناطق حساس‌تر به آلودگی، نقش بسزایی در این مسئله ایفا می‌کند. شکل ۶ نقشه آسیب‌پذیری به دست آمده از مدل را که در حقیقت بیان‌کننده مناطقی با احتمال تجاوز غلظت آرسنیک از حد مجاز است را نشان می‌دهد. براساس نقشه به دست آمده از مدل، ۱۳ درصد از وسعت منطقه مطالعه شده در محدوده آسیب‌پذیری بسیار کم، ۵۳ درصد در محدوده آسیب‌پذیری کم، ۲۱ درصد در محدوده آسیب‌پذیری متوسط، ۱۱/۵ درصد از منطقه در محدوده آسیب‌پذیری زیاد و ۱/۵ درصد در محدوده آسیب‌پذیری بسیار زیاد نسبت به آرسنیک قرار می‌گیرد. براساس نقشه آسیب‌پذیری، بیشترین آلودگی دشت مختص به نواحی مرکزی منطقه است. با توجه به پیش‌بینی مدل در اهمیت هر یک از متغیرها، که قابلیت انتقال و نیترات به‌عنوان مهم‌ترین عامل در بحث آلودگی آرسنیک شناخته شد، زیادبودن مقدار قابلیت انتقال و نیترات در مناطق مرکزی



شکل ۶. نقشه آسیب‌پذیری آبخوان نسبت به آرسنیک

مدیریت و نظارت صحیح آن استفاده شود. با توجه به پیش‌بینی مدل در اهمیت هر یک از متغیرها، قابلیت انتقال به‌عنوان مهم‌ترین عامل در بحث آلودگی آرسنیک شناخته شد. زیادبودن مقدار قابلیت انتقال، نبود شبکه فاضلاب، وجود چاه‌های جذبی زیاد و فعالیت شدید کشاورزی در مناطق مرکزی دشت، به خصوص در منطقه شهری ملکان، سبب افزایش غلظت آرسنیک به بیش از غلظت حد مجاز می‌شود.

### نتیجه‌گیری

روش RF با داشتن مزایایی مثل یادگیری روابط غیرخطی، توانایی در مقابل داده‌های پرت، برآورد خطای غیرمرتبط داخلی و اجرای هزاران داده ورودی بدون حذف یکی از آنها، به‌عنوان روشی دقیق در مدل‌سازی پیش‌بینی و آسیب‌پذیری شناخته شد. این مدل با شناسایی متغیرهای مهم در آلودگی آب‌های زیرزمینی دشت ملکان به آرسنیک می‌تواند برای

- [9]. Breiman, L; Friedman, J. H; Olshen, R.A, Stone C.J, 1984, Classification and regression trees, Chapman & Hall/CRC, New York.
- [10]. Cai, X; Yu, Y; Huang, Y; Zhang, L; Jia, P.M, Zhao, Q, 2003, Arsenic trioxide-induced mitotic arrest and apoptosis in acute promyelocytic leukemia cells *Leukemia*, 17: pp.1333-1337.
- [11]. Das, P; Samantaray, S; Rout, G.R, 1997, Studies on cadmium toxicity in plants, a review. *Environment Pollution*, 98: pp.29-36.
- [12]. Das, A.K; 1990, Metal ion induced toxicity and detoxification by chelation therapy, A text book on medical aspects of bio-inorganic chemistry, Isted, CBS, Delhi, 21: pp. 17-58.
- [13]. Duda, R.O; Hart, P.E; Stork, D.G, 2011, Pattern classification and introduction to the bootstrap. Vol. 57, pp. CRC press.
- [14]. Duker, A.A; Carranza E.J; Hale, M, 2005, Arsenic geochemistry and health, *Environmental Pollution*, 31 (5): pp. 631-64.
- [15]. Emberger, L; 1952, Sur le quotient pluviothermique, *C.R. Sciences*, 234: pp. 2508-2511.
- [16]. Evanko, F.R; Dzombak, D.A, 1997, Remediation of metals-contaminated soils and groundwater. Technology Evaluation Report prepared for Ground Water Remediation Technologies Analysis Center, *Journal of Hydrology*, 213: pp. 71-87.
- [17]. Facchinelli, A; Sacchi, E; Mallen, L, 2001, Multivariate statistical and GIS-based approach to identify heavy metal sources in soils, *Environment Pollution* 114: pp. 313-324.
- [18]. Friedl, M. A; Brodley, C. E; Strahler, A. H, 1999, Maximizing land cover classification accuracies produced by decision trees at continental to global scales, *IEEE Transvers Geoscience Remote Sensing* 37(2); pp. 969-77.
- [19]. Guyon, I; Elisseeff, A, 2003, an introduction to variable and feature selection, *Journal of machine Learning*, 3: pp. 1157-82.
- [20]. Jack C.N; Wang J; Shraim, A.A, 2003, Global health problem caused by arsenic from natural sources, *Chemosphere*, 52 (9): pp. 1353-1359.
- [21]. Ghassemzadeh, F; Arbab-Zavar, M.H; McLennon, G, 2006, Arsenic and antimony in drinking water in Khohsorkh area, northeast Iran, possible risks for the public health, *Journal of Applied Sciences*, 6 (13): pp. 2705-2714.

آلودگی قسمت جنوب شرقی آبخوان را می‌توان به منشأ زمین‌زادی آن به‌صورت سولفیدهای آرسنیک درون شکستگی‌های موجود در مارن‌های میوسن باشد. براساس نقشه به‌دست‌آمده از مدل، ۱۳ درصد از وسعت منطقه مطالعه‌شده در محدوده آسیب‌پذیری بسیار کم، ۵۳ درصد در محدوده آسیب‌پذیری کم، ۲۱ درصد در محدوده آسیب‌پذیری متوسط، ۱۱/۵ درصد از منطقه در محدوده آسیب‌پذیری زیاد و ۱/۵ درصد در محدوده آسیب‌پذیری بسیار زیاد قرار می‌گیرد. با توجه به نقشه آسیب‌پذیری، بیشترین آلودگی دشت مختص به نواحی مرکزی منطقه است.

#### منابع

- [1]. Asghari Moghaddam, A; Barzegar, R, 2013, considering Factors affecting high concentrations of arsenic in ground water Resources Tabriz plain aquifers, *Journal of Earth Sciences*, 92: pp. 154-147, [Persian].
- [2]. Nadiri, A; Asghari Moghaddam, A; Sadeghi, F; Agaie, H, 2011, investigation of the arsenic anomaly in water resources of Sahand Dam, *Journal of Environmental Studies*, 3: PP. 61-74, [Persian].
- [3]. Bellman, R, 2003, Dynamic programming. Mineola, NY: Dover Publications 366 pp.
- [4]. Ahmed, K.M; Bhattacharya, P; Hasan, M.A., Akhter, S.H; Alam, S.M.M; Bhuyian, M.A.H; Imam, M.B; Khan, A.A; Sracek, O, 2004, Arsenic enrichment in groundwater of the alluvial aquifers in Bangladesh: An overview, *Applied Geochemistry*, 19: pp.181-200.
- [5]. Booker, D.J; Snelder, T. H, 2012, comparing methods for estimating flow duration curves at ungauged sites. *Journal of Hydrology*, 434: pp. 78-94.
- [6]. Boisson, J; Ruttens, A; Mench, M, Vangronsveld, J, 1999, Evaluation of hydroxyapatite as a metal immobilizing Soil additive for the remediation of polluted soils, Part 1. Influence of hydroxyapatite on metal exchange ability in soil, plant growth and plant metal accumulation. *Environmental Pollution* 104: 225-233.
- [7]. Breiman, L; 1996, bagging predictors, *Mach Learn*, 24(2): pp. 123-40.
- [8]. Breiman, L; 2001, Random Forests, *Mach Learn*, 45(1): pp. 5-32.

- [22]. Jain, C.K; Ali, I, 2000, Arsenic Occurrence toxicity and speciation techniques, *Water Resource*, 34(17): pp. 4304-4312.
- [23]. Ko, B; Gim, J; Nam, J, 2011, Cell image classification based on ensemble features and random forest, *Electronics Letters*, 47: pp. 638-648.
- [24]. Mosafieri, M; Yunesian, M; Mesdaghinia, A.R; Nadim, A; Nasseri; and Mahvi, A.H, 2006, Arsenic occurrence in drinking water of L.R. of Iran- the case of Kurdistan province, In: *Fate of Arsenic in the environment, Proceedings of the BUET-UNU international symposium*, 5-6 February, Dhaka, Bangladesh, *Sciences*, 6 (13): pp. 2705-2714.
- [25]. Pal, M; 2005, Random Forest classifier for remote sensing classification, *International Journal of Remote Sensing* 26(1): pp. 217-22.
- [26]. Ramanathan, A.L; Balakrishna, P.M; Chidambaram, S, 2007, Groundwater Arsenic Contamination and its health effect-case studies from India and South East Asia, *Indian Journal Geochemistry*, 22: pp. 371-384.
- [27]. Rodriguez, V. F; Ghimire, B; Rogan, J; Chica-Olmo, M; Rigol-Sánchez, J. P, 2012d, An assessment of the effectiveness of a Random Forest classifier for land-cover classification, *ISPRS Journal of Photogram Remote Sensing*, 67: pp. 94-104.
- [28]. Smedley, P.L; Kinniburgh, D.G, 2002, A review of the source, behavior and distribution of arsenic in natural waters, *Applied Geochemistry*, 17(12): pp. 517- 568.
- [29]. USEPA, 2001, Risk Assessment Guidance for Superfund, Human Health Evaluation. Manual Part A, Interim Final, vol. 1, EPA/540/1-89/002.
- [30]. WHO (World Health Organization), 2009, *Guideline for Drinking Water Quality*.
- [31]. WHO, 2004, *Guidelines for drinking water quality, Third edition*, World Health Organization, Geneva.bnb.
- [32]. Zheng, Y; Stute, M; Gavrieli, A.I; Dhar, R; Simpson, H.J; Schlosser, P; Ahmed, K.M, 2004, Redox control of arsenic mobilization in Bangladesh groundwater, *Applied Geochemistry* 19: pp. 201-214.

Archive of SID