



فصلنامه علوم محیطی، دوره بیستم، شماره ۴، زمستان ۱۴۰۱

۱۲۵-۱۴۲

مقاله پژوهشی

## بهینه‌سازی روش SEAR در پاک‌سازی آبخوان‌های آلوده به DNAPL با مدل‌های جایگزین

حسین اژدری بجزستانی و سعید علیمحمدی\*

گروه مهندسی منابع آب، دانشکده مهندسی عمران، آب و محیط زیست، دانشگاه شهید بهشتی، تهران، ایران

تاریخ دریافت: ۱۴۰۰/۱۲/۱ تاریخ پذیرش: ۱۴۰۱/۲/۱۲

اژدری بجزستانی، ح. و س. علیمحمدی. ۱۴۰۱. بهینه‌سازی روش SEAR در پاک‌سازی آبخوان‌های آلوده به DNAPL با مدل‌های جایگزین. فصلنامه علوم محیطی. ۲۰(۴): ۱۲۵-۱۴۲.

**سابقه و هدف:** مایعات با فاز غیرآبی چگال (DNAPL)<sup>۱</sup> در بین رایج‌ترین گونه‌های آلودگی در آب زیرزمینی هستند. روش پمپاژ و تصفیه ارتقاء یافته (SEAR)<sup>۲</sup> یکی از روش‌های مرسوم پاک‌سازی آبخوان‌های آلوده به DNAPL است. باتوجه به هزینه بالای مواد شیمیایی مورد استفاده در این روش (سورفکتانت‌ها و کوسالونت‌ها)، انتخاب الگوی مناسب جانمایی چاه‌ها و نرخ‌های پمپاژ بهینه ضروری است. نرم‌افزار شبیه‌سازی UTCHEM توانایی مدل‌سازی نحوه انتقال و انتشار DNAPL‌ها و قابلیت اجرای روش SEAR را دارد. مشکل اصلی استفاده از این نرم‌افزار، مدت‌زمان زیاد مورد نیاز برای اجرای متعدد مدل در استفاده از آن برای بهینه‌سازی سیستم است. هدف از این پژوهش استفاده از دو روش یادگیری ماشین<sup>۳</sup> (شبکه عصبی مصنوعی و K همسایه نزدیک‌تر) به‌عنوان مدل‌های شبیه‌سازی جایگزین و وارد نمودن بهترین مدل در نرم‌افزار LINGO برای بهینه‌سازی روش SEAR است.

**مواد و روش‌ها:** در اجرای روش SEAR، داده‌های کمی و کیفی آبخوان برای مدل‌سازی نحوه انتشار، انتقال و حذف DNAPLs در نرم‌افزار UTCHEM مورد نیاز است. برای این منظور از اطلاعات سایت Camp Lejeune در کارولینای شمالی، ایالات متحده آمریکا استفاده شد. در این پژوهش با بررسی انواع مدل‌های جایگزین بر مبنای روش‌های یادگیری ماشین و اجرای ۲۵۰ سناریو مختلف در نرم‌افزار UTCHEM، از دو مدل روش شبکه عصبی مصنوعی (ANN)، و نزدیک‌ترین همسایگی‌های K (KNN) به‌منظور شبیه‌سازی روش SEAR و توسعه مدل جایگزین استفاده گردید. به‌منظور صحت‌سنجی دو مدل جایگزین، ۵۰ سناریو جدید در نرم‌افزار UTCHEM اجرا شد و درصد پاک‌سازی آن‌ها به دست آمد. همچنین با استفاده از دو مدل جایگزین نیز درصد‌های پاک‌سازی ۵۰ سناریو مشخص گردید. برای ارزیابی عملکرد مدل‌های جایگزین، از آماره جذر میانگین مربعات خطا (RMSE)<sup>۴</sup> استفاده شد و با نتایج سایر پژوهش‌ها مقایسه گردید. در انتها از مدل جایگزینی که دقت بیش‌تری داشت در نرم‌افزار LINGO به‌منظور بهینه‌سازی روش پمپاژ و تصفیه ارتقاء یافته (SEAR) استفاده شد.

**نتایج و بحث:** مقادیر RMSE در نتایج به‌دست‌آمده از مدل‌های جایگزین ANN و KNN در مرحله صحت‌سنجی به ترتیب برابر ۰/۶۷ و ۶۶/۱ بوده که این بیانگر دقت بالای هر دو مدل جایگزین، به‌خصوص ANN است. میانگین مدت‌زمان هر اجرای نرم‌افزار UTCHEM در این

\* Corresponding Author: Email Address. s\_alimohammadi@sbu.ac.ir

<http://dx.doi.org/10.52547/envs.2022.1164>

<http://dorl.net/dor/20.1001.1.17351324.1401.20.4.4.9>

پژوهش ۴۵ دقیقه بوده است در حالی که در مدل جایگزین به چند ثانیه کاهش یافت؛ همچنین نرم‌افزار LINGO برای مشخص کردن سناریو بهینه حدود ۲۱۵۰۰ سناریو مختلف را در مدت زمان ۳۰ دقیقه بررسی نمود در حالی که در صورت عدم استفاده از مدل جایگزین مدت زمان مورد نیاز برای این کار بیش از ۱۶۰۰۰ ساعت است. بر اساس موقعیت و دبی چاه‌های فعال در سناریو بهینه، مشخص شد که اولاً چاه‌های موجود در بالادست و پایین دست آلودگی بیشترین تأثیر را در پاک‌سازی دارند و ثانیاً عامل زمان بیشتر از دبی پمپاژ چاه‌ها در پاک‌سازی تأثیر دارد. سناریو بهینه به دست آمده در این پژوهش با هزینه کم‌تر نسبت به هزینه‌های گزارش شده در پروژه Camp Lejeune و در مدت زمان ۳۰ روز، منطقه آلوده به DNAPL را تا ۹۵٪ پاک‌سازی می‌کند.

**نتیجه‌گیری:** بر اساس نتایج به دست آمده در این پژوهش مشخص گردید که استفاده از الگوریتم‌های یادگیری ماشین نظیر ANN و KNN، به همراه نرم‌افزار بهینه‌سازی LINGO که از قوی‌ترین نرم‌افزارهای حل مسائل بهینه‌سازی خطی و غیرخطی است، باعث می‌شود علاوه بر داشتن دقت مناسب، مدت زمان مورد نیاز برای یافتن سناریو بهینه تا حد چشم‌گیری کاهش یابد.

**واژه‌های کلیدی:** پمپاژ و تصفیه ارتقاء یافته، DNAPL، شبکه عصبی مصنوعی MLP، K همسایگی نزدیک‌تر، نرم‌افزار LINGO.

## مقدمه

زیرسطحی قرار می‌گیرند، معمولاً یا ناشی از محصولات پالایش نفت و یا ناشی از احتراق سوخت‌های فسیلی هستند. همچنین در مورد حلال‌ها و پاک‌کننده‌های روغن کلردار یا بدون کلر و ترکیبات آلی که به عنوان ماده خام در انواع فرایندهای تولید استفاده می‌شوند، نگرانی‌های زیست‌محیطی زیادی وجود دارد (Shobeyri, 2019). در این میان سیالات ارگانیک با قابلیت انحلال کم در آب که تحت عنوان مایعات فاز غیرآبی (NAPL) شناخته می‌شوند، باتوجه به استفاده روزافزون در بین رایج‌ترین گونه‌های آلودگی در آب زیرزمینی هستند (Qin 2009).

(*et al.*,

NAPL خلاصه شده عبارت Non Aqueous Phase Liquid به معنی مایعات فاز غیرآبی است و به دو دسته DNAPL و LNAPL<sup>۵</sup> تقسیم می‌شوند. NAPL‌های سنگین که چگالی آن‌ها بیشتر از چگالی آب است، مانند حلال‌های کلردار، تار زغال‌سنگ، کرئوزوت، و نفت خام سنگین به عنوان DNAPL شناخته می‌شوند و NAPL‌های سبک که چگالی آن‌ها کمتر از چگالی آب است، مثل گازوئیل، بنزن، تولوئن، زایلن و سایر هیدروکربن‌ها، به عنوان LNAPL شناخته می‌شوند (Mayer and Hassanzade, 2005).

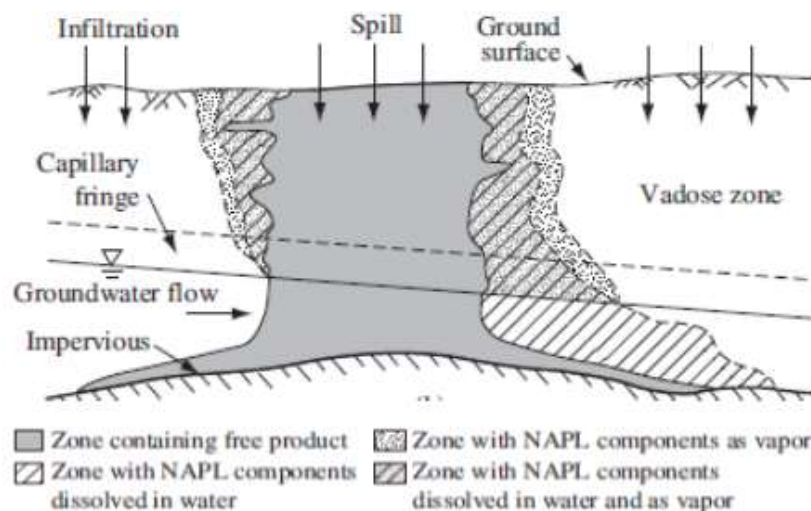
یکی از مهم‌ترین ویژگی‌های آلاینده‌های نامحلول چگال

امروزه آلودگی آبخوان‌ها تبدیل به یک نگرانی بزرگ در بسیاری از کشورها، به خصوص کشورهایی که برای تأمین آب آشامیدنی به منابع آب‌های زیرزمینی متکی هستند شده است. در چند دهه اخیر با پیشرفت صنایع، حجم پساب زیادی از طریق دفع نامناسب و نشت از تانکرهای زیرزمینی، خطوط لوله و زمین‌های با نفوذپذیری بالا وارد منابع آب‌های زیرزمینی شده و غلظت آلودگی را در برخی آبخوان‌ها افزایش داده است؛ در حالی که به طور هم‌زمان تقاضا و معیارهای زندگی در محیط زیست طبیعی در حال افزایش است. این موضوع، اهمیت حفاظت آب‌های زیرزمینی از آلودگی و کنترل مناطق آبخوان آلوده را نشان می‌دهد (Shaafi, 2013; Kueper and McWhorter, 1991).

تکنولوژی‌های بسیاری برای پاک‌سازی آب‌های زیرزمینی آلوده وجود دارد، اما انتخاب تکنولوژی مناسب به ویژگی‌های سایت، نوع آلودگی، الزامات قانونی، هزینه‌ها و محدودیت‌های زمانی بستگی دارد؛ همچنین باید مسائل زیست‌محیطی و اثرات ناشی از این روش‌ها در نظر گرفته شود و باتوجه به ارزیابی نهایی، تکنولوژی مناسب انتخاب گردد (Mayer and Endres, 2007).

بیش از ۱۶۰۰ ترکیب آلی در محیط‌های طبیعی و آلوده شناخته شده است؛ آن‌هایی که در زمره آلودگی‌های

سطح زمين، اين آلاينده فرايند انتقالش را به صورت افقي و عمودي به سمت اعماق زمين شروع نموده (شکل ۱) و در نهايت به شکل لکه‌هاي گسسته رسوب مي‌کند (Mercer and Cohen, 1990).



شکل ۱- رفتار DNAPL هنگامی که در حجم زياد وارد آبخوان می‌شود (Bear and Cheng, 2010)  
 Fig. 1- DNAPL behavior when entering to the aquifer in large volumes (Bear and Cheng, 2010)

آلاينده توسط پمپاژ آب زيرزميني می‌شود (Kueper *et al.*, 2014).

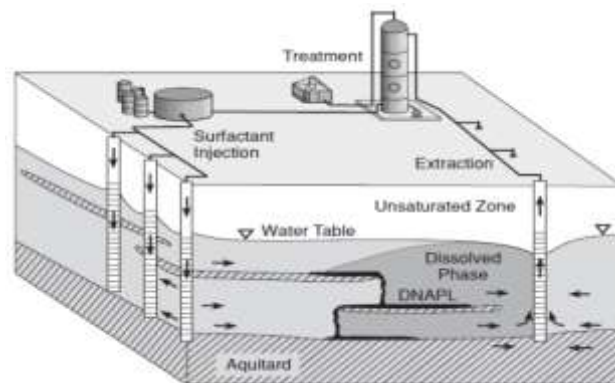
سورفکتانت‌ها عوامل فعال سطحی هستند که متشکل از دو بخش کاملاً متفاوت شیمیایی می‌باشند، این دو بخش عبارت‌اند از یک سر هیدروفیلیک<sup>۶</sup> و یک دم هیدروفوبیک<sup>۷</sup>. هر دوی این بخش‌ها قابلیت انحلال را هم در آب و هم در روغن دارند. این ویژگی منحصر به فرد است که امکان افزایش انحلال‌پذیری مواد غیر قابل انحلال در آب را می‌دهد و همچنین با کاهش شدید نیروی سطحی بین آب و NAPL‌ها، قابلیت جابه‌جایی را ایجاد می‌کند (Battelle *et al.*, 2002).

شکل ۲ یک شمای مفهومی از روش SEAR را نشان می‌دهد. مبنای تکنولوژی SEAR استفاده از سورفکتانت‌ها (مانند SEAR، در روش Alfoterra 145-4 PO Sulfate) است. سورفکتانت در سرتاسر ناحیه آلوده از طریق چاه‌های تزریق وارد می‌شود و از طریق چاه‌های برداشت خارج می‌گردد. برای اینکه تمام محیط آلوده پوشش داده شود باید از تعدادی چاه تزریق و چاه برداشت استفاده شود (Huling and Weaver, 1991).

در آب یا همان DNAPL‌ها، انحلال‌پذیری کم آنها است. این مواد به دلیل سنگینی، به اعماق خاک فرو رفته و در نهايت به آب‌های زيرزميني راه‌یافته و منجر به آلودگی آنها می‌شوند. در واقع به محض انتشار DNAPL بر روی

آلودگی آبخوان‌ها به NAPL به‌خصوص DNAPL یکی از مشکلات جدی در زمینه کنترل کیفیت آب‌های زيرزميني است. ثابت شده که بسیاری از تکنیک‌های رایج پاک‌سازی خاک و آب زيرزميني که به این مواد آلوده هستند، مانند پمپاژ و تصفیه و استخراج بخار ناموفق یا به شکل محدود موفقیت‌آمیز بوده‌اند که علت آن حلالیت کم، کشش سطحی زیاد و تمایل به فرورفتن به زیر تراز آب زيرزميني است (Suthersan *et al.*, 2016; Delshad *et al.*, 1996). علاوه بر قابلیت اندک انحلال در آب، بسیاری از NAPL‌ها از نظر بیولوژیکی سرسخت هستند و به خاطر همین اعمال روش پاک‌سازی بیولوژیکی به دلیل زمان‌بر بودن، سیاست موفق‌تری نمی‌باشد (Huling and Weaver, 1991).

روش پاک‌سازی بهبودیافته آبخوان با سورفکتانت (SEAR) به‌عنوان یک تکنولوژی نویدبخش که بر بسیاری از محدودیت‌های پاک‌سازی آبخوان از DNAPL‌ها غلبه کرده است، در سالیان اخیر تحت بررسی و توسعه قرار گرفته است. SEAR یک روش توسعه‌یافته جهت افزایش قابلیت انحلال و انتقال NAPL‌ها است که منجر به افزایش نرخ حذف



شکل ۲- یک طرح مفهومی برای نمایش روش SEAR (Battelle et al., 2002)  
Fig. 2- A conceptual design for a SEAR method (Battelle et al., 2002)

تحت آلاینده‌های نفتی است و همچنین توانایی پاک‌سازی آبخوان‌ها را از طریق توسعه و استفاده از شبیه‌سازی ترکیبی دارد (Delshad et al., 1996). UTCHEM یک مدل سه‌بعدی چند فاز و چند جزئی است که در آن ویژگی‌های جذب شیمیایی، فیزیکی و زیستی، واکنش‌های ژئوشیمیایی و همچنین حذف و کاهش آلودگی‌های نفتی آبخوان از طریق روش SEAR بررسی شده است (Qin et al., 2007; Ouyang et al., 2017b)

مشکل اصلی استفاده از این نرم‌افزار و سایر مدل‌های شبیه‌سازی کیفی و توزیعی آبخوان‌ها، مدت‌زمان موردنیاز برای اجرای متعدد مدل در شرایط استفاده از آن برای بهینه‌سازی سیستم است و این موضوع کار را مشکل و در بعضی از مواقع ناممکن می‌سازد (Jiang et al., 2015). برای حل این مشکل می‌توان از مدل‌های جایگزین بر مبنای روش‌های یادگیری ماشین استفاده کرد. در صورت استفاده از روش‌های یادگیری ماشین می‌توان در کنار داشتن دقت مناسب، مدت‌زمان موردنیاز را تا حد زیادی کاهش داد. روش‌های یادگیری ماشین که بر مبنای هوش محاسباتی و با الهام گرفتن از طبیعت ابداع شده‌اند، امروزه در دامنه وسیعی از مهندسی و علوم کاربرد دارند. سرعت و دقت این روش‌ها در مقایسه با روش‌های شبیه‌سازی فرایند مینا، علت اصلی توجه به این روش‌ها است (Luo et al., 2013; Ouyang et al., 2017b). از جمله تفاوت‌های اصلی پژوهش حاضر با سایر پژوهش‌های انجام شده

روش SEAR یک روش بسیار کارا و مؤثر در زمینه حذف DNAPL می‌باشد ولی با بررسی پروژه‌های انجام شده با این روش، مشخص شده که روشی گران‌قیمت است (Luo et al., 2013). در این روش علاوه بر هزینه‌های موردنیاز برای کارهای تحقیقاتی، نمونه‌برداری و آزمایشگاه، تجهیزات، تأسیسات و همچنین حفر چاه‌ها، بخش عمده‌ای از هزینه‌ها مربوط به تهیه ماده روان‌ساز یا سورفکتانت است (Mayer and Endres, 2007). با توجه به هزینه بالای اجرای این روش، فعالیت در زمینه بهینه‌سازی و بهبود عملکرد آن یک موضوع بسیار ضروری می‌باشد (Jiang et al., 2015). آلاینده‌ها آب به دو گروه محلول و نامحلول تقسیم‌بندی می‌شوند (Lerner et al., 2003). در زمینه آلاینده‌های محلول در آب، به دلیل آنکه پیش‌بینی جریان و رفتار آنها به‌مراتب شناخته‌شده‌تر از آلاینده‌های نامحلول است، تحقیقات بیشتری انجام گرفته و مدل‌های مناسبی جهت پیش‌بینی رفتار آنها در آب‌های زیرزمینی ارائه شده است. اما در مورد آلاینده‌های نامحلول به علت چند فاز بودن محیط اختلاط، مدل‌های توسعه‌یافته کمتری موجود است و اغلب نیز محدودیت‌هایی در مدل‌سازی دارند (Shobeyri, 2019).

از بین نرم‌افزارهای موجود، نرم‌افزار UTCHEM که حاصل یک تحقیق پیش‌گامانه در دانشگاه تگزاس است، توانایی مدل‌سازی نحوه انتشار و انتقال DNAPL‌ها و قابلیت اجرای روش SEAR را دارد. این نرم‌افزار یک مبنای علمی و مهندسی برای مدل‌سازی و بازسازی نواحی بهبودیافته

مدل شبیه‌سازی استفاده شده است. در پژوهشی دیگر (Luo *et al.* (2014) از سه مدل جایگزین بر مبنای روش‌های یادگیری ماشین که عبارت‌اند از رگرسیون چندمتغیره، کریجینگ و شبکه عصبی مصنوعی شعاع مینا جهت بهینه‌سازی پاک‌سازی آبخوان آلوده به DNAPL استفاده کردند. نتایج این پژوهش نشان داد که دو روش کریجینگ و شبکه عصبی مصنوعی شعاع مینا از دقت تقریب مناسب برخوردار هستند و عملکرد مدل کریجینگ به میزان کمی از شبکه عصبی مصنوعی شعاع مینا بهتر است. اما دقت مدل رگرسیون چندجمله‌ای مناسب گزارش نشده است.

مدل‌های دیگری مانند رگرسیون بردار پشتیبان (SVR) نیز به‌عنوان مدل جایگزین توسط (Hou *et al.* (2017) مورد استفاده قرار گرفته است. آن‌ها علاوه بر مدل SVR از مدل کریجینگ و مدل شبکه عصبی شعاع مینا برای مقایسه توانایی مدل‌های جایگزین مختلف استفاده کردند و افزایش توانایی مدل‌ها به‌واسطه بهینه‌سازی پارامترهای آن‌ها مورد ارزیابی قرار گرفت. خطای نسبی نرخ حذف آلاینده بین مدل جایگزین بهینه شده و مدل شبیه‌سازی برای ۱۰۰ داده اعتبارسنجی کمتر از ۳ درصد گزارش شده است که نشانگر عملکرد مناسب مدل جایگزین می‌باشد. (Ouyang *et al.* (2017a) از سه مدل جایگزین بر مبنای روش‌های یادگیری ماشین که عبارت‌اند از برنامه‌ریزی ژنتیک چندژنی، کریجینگ و رگرسیون بردار پشتیبان جهت بهینه‌سازی پاک‌سازی آبخوان آلوده به DNAPL استفاده کردند. نتایج این پژوهش حاکی از برتری برنامه‌ریزی ژنتیک چند ژنی نسبت به کریجینگ و رگرسیون بردار پشتیبان می‌باشد.

در سال‌های اخیر در ایران نیز پژوهشگران به بررسی و تحقیق درباره موضوعات مربوط به شبیه‌سازی و بهینه‌سازی پاک‌سازی آلاینده‌های DNAPL پرداخته‌اند. به‌عنوان مثال (Shams *et al.* (2021) با استفاده از مدل‌های جایگزین از قبیل شبکه عصبی، SVR، فرایند گاوسی و

در این است که در انجام آن از دو روش شبکه عصبی مصنوعی (ANN) و K همسایگی نزدیک‌تر (KNN) به‌عنوان مدل جایگزین به‌منظور شبیه‌سازی روش SEAR استفاده شده است و سپس از مدل جایگزینی که عملکرد آن بهتر بوده به همراه نرم‌افزار LINGO برای بهینه‌سازی روش SEAR استفاده گردیده.

### پیشینه مطالعاتی استفاده از مدل‌های جایگزین در

#### بهینه‌سازی پاک‌سازی آلاینده‌های DNAPL

در دو دهه اخیر توجه ویژه‌ای به موضوع استفاده از مدل‌های جایگزین در پاک‌سازی آلاینده‌های DNAPL شده است. (Qin *et al.* (2009) از اولین افرادی بودند که از مدل‌های جایگزین در پاک‌سازی آبخوان‌های آلوده به آلاینده‌های DNAPL استفاده کردند. آنها سپس اقدام به بهینه‌سازی سیستم پاک‌سازی آبخوان به روش SEAR با استفاده از مدل‌های جایگزین بر مبنای روش‌های یادگیری ماشین نمودند. مدل جایگزین مورد استفاده در این پژوهش رگرسیون چندمتغیره بوده است.

(Luo and Lu (2013) از شبکه عصبی مصنوعی شعاع مینا برای ساخت مدل جایگزین به‌منظور تقریب مدل شبیه‌سازی جریان چند فازی آبخوان آلوده به DNAPL استفاده نمودند؛ مدل توسعه داده شده در مسئله بهینه‌سازی پاک‌سازی آبخوان آلوده به مواد نفتی مورد استفاده قرار گرفت. در این پژوهش خطای نسبی متوسط نرخ حذف مواد نفتی، بین مدل جایگزین بر مبنای روش یادگیری ماشین (شبکه عصبی مصنوعی شعاع مینا) و مدل شبیه‌سازی برای ۱۰ داده اعتبارسنجی کمتر از ۵ درصد گزارش شده که نشان‌دهنده دقت بالای مدل جایگزین است؛ همچنین بار محاسباتی فرایند شبیه‌سازی و بهینه‌سازی با استفاده از روش یادگیری ماشین به‌صورت قابل‌ملاحظه‌ای کاهش یافت. آن‌ها همچنین با استفاده از مدل جایگزین کریجینگ اقدام به بهینه‌سازی پاک‌سازی آبخوان آلوده به DNAPL نمودند. در این پژوهش از روش SEAR و مدل UTCHEM به‌عنوان روش پاک‌سازی و

انتقال و حذف DNAPLها در نرم‌افزار UTCHEM، داده‌های کمی و کیفی آبخوان (خصوصیات فیزیکی - هیدروژئولوژیکی و شیمیایی) موردنیاز است. به همین خاطر در این پژوهش از گزارش‌ها و اطلاعات مربوط به پروژه Camp Lejeune که در یک آبخوان سطحی کم‌عمق در محل استقرار پایگاه نیروی دریایی آمریکا در سایت شماره ۸۸ واقع در کارولینای شمالی، استفاده شده است (Camp 1999 Lejeuneet). دلیل استفاده از گزارش‌ها و اطلاعات مربوط به پروژه Camp Lejeune در این پژوهش این بوده است که اولاً داده‌های مورد استفاده برای یک منطقه واقعی بوده باشند، نه یک منطقه فرضی و ثانیاً بتوان نتایج به‌دست‌آمده را با سایر پژوهش‌های مشابه مقایسه نمود. در پروژه Camp Lejeune مجموعه‌ای از فعالیت‌ها جهت حذف DNAPL به روش SEAR انجام گردیده است. موقعیت آبخوان مورد مطالعه و محل آلوده، مقطع زمین‌شناسی و سطح پتانسیومتری آبخوان در شکل ۳ مشخص شده است. هیچ منبع آب سطحی در نزدیکی این سایت وجود ندارد.

درخت تصمیم به شبیه‌سازی مدل‌های پاک‌سازی آبخوان‌های آلوده و بهینه‌سازی فرایند پاک‌سازی پرداختند؛ همچنین به منظور افزایش دقت نتایج، از راهکار تجمیع مدل‌ها استفاده کردند. باتوجه به مقالات بررسی شده می‌توان نتیجه گرفت که امروزه پژوهش در زمینه شبیه‌سازی و بهینه‌سازی سیستم‌های پاک‌سازی آلاینده DNAPL باتوجه به نقش و اهمیت آن‌ها در حفظ و نگهداری منابع آب و محیط‌زیست و همچنین هزینه‌های بالای موردنیاز برای این سیستم‌ها، از اهمیت زیادی برخوردار است. ارائه‌ی روش‌های نوین شبیه‌سازی و بهینه‌سازی پاک‌سازی آلاینده DNAPL جز مباحثی است که از یک سو جزو معضلات بزرگ زیست-محیطی است و از سوی دیگر همچنان از مباحث جذاب مطالعاتی و پژوهشی می‌باشد.

## مواد و روش‌ها

### منطقه مورد مطالعه و اطلاعات مدل UTCHEM

برای اجرای روش SEAR به منظور مدل‌سازی نحوه انتشار و



شکل ۳- موقعیت آبخوان مورد مطالعه و محل آلوده، مقطع زمین‌شناسی و سطوح پتانسیومتری آبخوان (Shams et al., 2021)

Fig. 3- The location of the studied aquifer and contaminated site, geological cross-section, and the potentiometric surface of the aquifer (Shams et al., 2021)

خشک‌شویی سایت Camp Lejeune می‌باشد. پرکلرواتیلن ماده شیمیایی اصلی مورد استفاده در صنعت خشک‌شویی است. شرکت مهندسی و خدمات Duke در سال ۱۹۹۷ منطقه

محل دقیق آلودگی در شکل ۳ مشخص شده است. این آلودگی ناشی از نشت یک بشکه پرکلرواتیلن (PCE) و نفوذ آن به زمین در مدت ۳۰۰ روز می‌باشد. آلودگی ایجاد شده به خاطر نشت این ماده شیمیایی از مخازن موجود در

ادامه جدول ۱- پارامترهای فیزیکی و شیمیایی استفاده شده در مدل شبیه‌سازی SEAR

Table 1. Cont. Physical and chemical parameters used in the SEAR simulation model

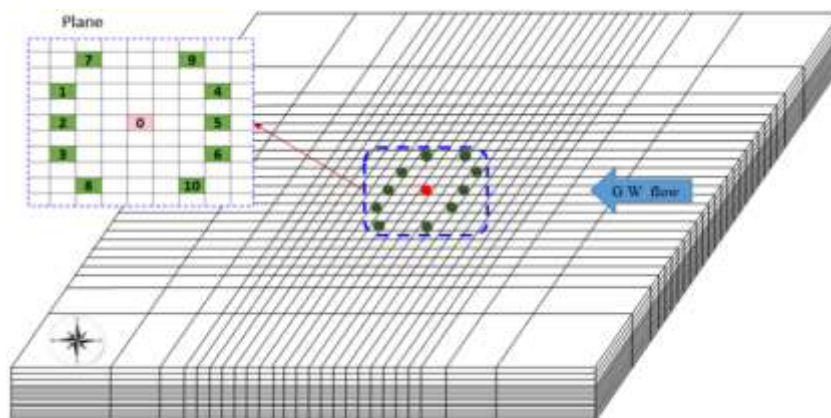
پارامتر Parameter	مقدار Value
چگالی IPA (g/cm <sup>3</sup> )	0.786
IPA density (g/cm <sup>3</sup> )	
حلالیت PCE در آب (g/L)	0.15
PCE solubility in water (g/L)	
کشش سطحی PCE-آب خالص (dynes/cm)	45
Interfacial tension PCE-pure water (dynes/cm)	
تراکم پذیری اجزا (1/psi)	0
Compressibility of components (1/psi)	
پراکندگی طولی (m)	0.1
Longitudinal dispersivity (m)	
پراکندگی عرضی (m)	0.02
Transverse dispersivity (m)	
ضریب انتشار مولکولی (cm <sup>2</sup> /s)	8.2e-6
Molecular diffusion coefficient (cm <sup>2</sup> /s)	
اشباع باقیمانده فاز آبی	0.24
Residual saturation of aqueous phase	
اشباع باقیمانده فاز اولئیک	0.15
Residual saturation of oleic phase	
ثابت قانون هنری	0.765
Henry's Law constant	

مطابق شکل ۴ ساختار مدل در سطح، شبکه‌ای ۲۵ در ۲۵ است که ابعاد طول آن به ترتیب ۷.۳۲، ۳.۶۶، ۱.۸۳، ۰.۹۱، ...، ۰.۹۱، ۱.۸۳، ۳.۶۶، ۷.۳۲ متر و ابعاد عرض آن به ترتیب ۳.۶۶، ۱.۸۳، ۰.۹۱، ۰.۶۱، ...، ۰.۶۱، ۱.۸۳، ۳.۶۶ متر به صورت متقارن است. طول کل محدوده مورد بررسی ۴۳ متر و عرض آن ۲۴.۴ متر می‌باشد. عمق این محدوده نیز ۴ متر و متشکل از ۱۶ لایه است؛ سه لایه فوقانی هر کدام به عمق ۰.۶۱، لایه چهارم عمق ۰.۳۰ و ۱۲ لایه تحتانی هر کدام به عمق ۰.۱۵ متر.

آلوده شده در سایت Camp Lejeune را با همکاری Baker Environmental بررسی کرده است (Battelle and Duke, 2002). با استفاده از داده‌های به دست آمده از تحقیقات میدانی در این سایت به عنوان پارامترهای ورودی، مدل ژئوسیستم<sup>۸</sup> این سایت توسط دانشگاه تگزاس در نرم‌افزار UTCHEM ساخته شد. در این مدل از فرایند SEAR به منظور از بین بردن آلودگی DNAPL موجود در این سایت در مدت زمان ۱۴۰ روز استفاده گردید. این مدل تحت عنوان مدل ۵۲ نرم‌افزار UTCHEM نیز شناخته می‌شود. مقادیر پارامترهای فیزیکی و شیمیایی مورد استفاده در این پژوهش در جدول ۱ مشخص شده است.

جدول ۱- پارامترهای فیزیکی و شیمیایی استفاده شده در مدل شبیه‌سازی SEAR  
Table 1. Physical and chemical parameters used in the SEAR simulation model

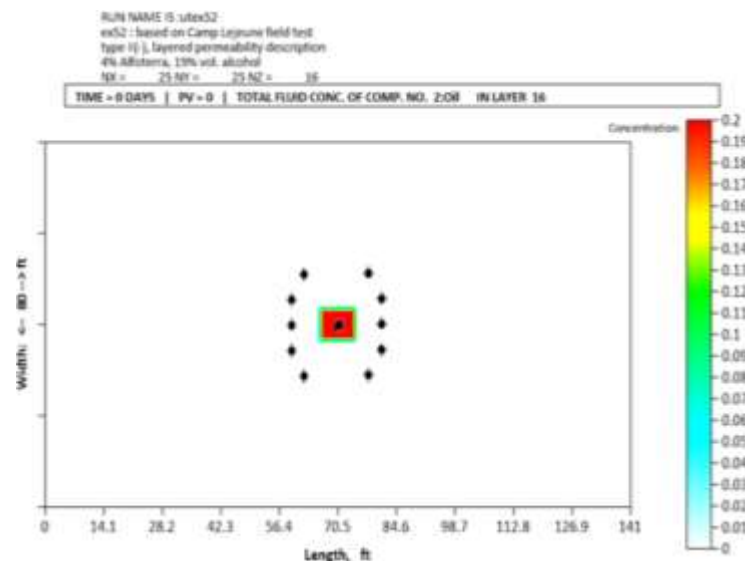
پارامتر Parameter	مقدار Value
ویسکوزیته آب (mPa.s)	1
Water viscosity (mPa.s)	
ویسکوزیته PCE (mPa.s)	0.89
PCE viscosity (mPa.s)	
چگالی آب (g/cm <sup>3</sup> )	0.998
Water density (g/cm <sup>3</sup> )	
چگالی PCE (g/cm <sup>3</sup> )	1.632
PCE density (g/cm <sup>3</sup> )	
چگالی سورفکتانت (g/cm <sup>3</sup> )	1.057
Surfactant density (g/cm <sup>3</sup> )	



شکل ۴- نمای سه بعدی محل قرارگیری چاه‌ها در محدوده انجام عملیات SEAR

Fig. 4- Three dimensional view of the location of wells in the area of SEAR operation

چیدمان و جانمایی چاه‌های تزریق و پمپاژ در این مدل بر اساس نتایج به‌دست‌آمده توسط Shams و همکاران می‌باشد (Shams *et al.*, 2021). به عبارتی در این پژوهش از آرایش k-spot همگرا استفاده گردیده است. در آرایش k-spot همگرا چاه‌های بیرونی چاه‌های تغذیه هستند و به‌صورت دایره‌ای (بر روی یک دایره) قرار دارند و چاه برداشت در مرکز قرار دارد. در شکل ۵ توزیع اولیه آلودگی در لایه ۱۶ (لایه تحتانی) به همراه محل قرارگیری چاه‌ها مشخص شده است.



شکل ۵- توزیع اولیه آلودگی در لایه ۱۶ و محل قرارگیری چاه‌ها

Fig. 5- The initial distribution of contamination in the 16th layer and location of wells

SEAR استفاده کمتری شده است.

### شبکه عصبی مصنوعی پرسپترون چندلایه

شبکه عصبی مصنوعی پرسپترون چندلایه (MLP-ANN) پرکاربردترین نوع شبکه عصبی است که در آن ارتباط نورون‌های هر لایه فقط به‌صورت جلو رونده با نرون‌های لایه بعدی برقرار می‌باشد. برای آموزش شبکه عصبی مصنوعی MLP از الگوریتم پس انتشار خطا که یک روش یادگیری با نظارت است استفاده می‌گردد. منظور از آموزش یا کالیبراسیون یک شبکه عصبی، تنظیم پارامترهای شبکه (وزن‌ها و بایاس‌ها) با ارائه الگوهای آموزشی است به‌طوری که با ارائه مجدد همان الگوها، خطای حاصل بین پاسخ مطلوب و شبکه حداقل گردد (Harun *et al.*, 2002). به‌طور کلی در شبکه پرسپترون

آلودگی در مرکز محدوده قرار دارد؛ به همین علت برای افزایش دقت و بررسی بهتر تغییرات آلودگی در طی فرایند پاک‌سازی، ابعاد شبکه در مرکز کوچک‌تر فرض شده است. در شکل ۴ سلول‌های سبزرنگ بیانگر محل قرارگیری چاه‌های پمپاژ هستند که از ۱ تا ۱۰ نام‌گذاری شده‌اند. همچنین سلول قرمز رنگ نشان‌دهنده محل چاه تخلیه و یا برداشت است. جهت حرکت آب زیرزمینی نیز مطابق شکل از شرق به غرب منطقه می‌باشد.

### روش تحقیق

در این پژوهش پس از مطالعه کتابخانه‌ای و بررسی دقیق مدل‌های جایگزین بر مبنای روش‌های یادگیری ماشین، از دو مدل جایگزین بر مبنای شبکه عصبی مصنوعی MLP-ANN Multi Layer Perceptron و K همسایگی نزدیک‌تر (KNN) به‌منظور شبیه‌سازی روش SEAR و نرم‌افزار LINGO برای بهینه‌سازی آن استفاده شده است. علت استفاده از دو مدل جایگزین ANN و KNN این بوده است که در پژوهش‌های مشابه بررسی شده مشخص گردید که در اغلب آنها مدل جایگزین بر مبنای شبکه عصبی مصنوعی (ANN) دقت بهتری نسبت به سایر مدل جایگزین دارد و همچنین تاکنون از مدل جایگزین KNN که یک روش ناپارامتری است، به‌منظور شبیه‌سازی روش



بهترین تخمین برای متغیر وابسته جدید استفاده می‌شود؛ بنابراین یک فرم یا معادله از پیش تعیین شده برای برآورد وجود ندارد بلکه بر اساس اطلاعات به دست آمده از داده‌ها، تخمین انجام می‌شود (Shmueli *et al.*, 2011). بر اساس آمارهای ارائه شده، الگوریتم K همسایه نزدیک‌تر یکی از ده الگوریتمی است که بیشترین استفاده را در پروژه‌های گوناگون یادگیری ماشین، هم در صنعت و هم در پژوهش داشته است (Swikatek *et al.*, 2019). در روش KNN با استفاده از یک معیار فاصله تعداد K همسایه نزدیک‌تر از داده‌ها که کم‌ترین فاصله را با داده‌های جدید دارند انتخاب شده و با میانگین‌گیری وزنی مقدار خروجی آن‌ها، مقدار خروجی داده جدید برآورد می‌شود (Araghinejad *et al.*, 2014). به منظور یافتن بهترین مدل KNN، ضمن بررسی تعداد مختلف همسایگان (مقادیر K متفاوت) و استفاده از فاصله اقلیدسی به عنوان معیار فاصله، می‌توان برای هر یک از متغیرهای ورودی با اجرای عملیات تحلیل حساسیت<sup>۱۱</sup> یک وزن خاص در نظر گرفت. این نوع محاسبه فاصله در واقع فقط به کمیت مقدار متغیرهای ورودی نمی‌پردازد بلکه به کیفیت آن‌ها نیز اهمیت می‌دهد و باعث می‌شود که دقت محاسبات بیشتر گردد. واضح است هر چه این وزن‌ها دقیق‌تر باشند، دقت محاسبات بیشتر می‌شود ولی اگر وزن‌ها به درستی انتخاب نشوند، حتی دقت محاسبات نسبت به قبل کاهش نیز می‌یابد (Moradian and Sepehrifar, 2009). در صورت استفاده از وزن‌های مختلف برای متغیرهای ورودی، رابطه محاسبه فاصله به صورت زیر خواهد بود (Huda *et al.*, 2014).

$$D_i = \frac{1}{\sqrt{w_1(x_{new1} - x_{i1})^2 + w_2(x_{new2} - x_{i2})^2 + \dots + w_7(x_{new7} - x_{i7})^2}} \quad (3)$$

### نرم‌افزار LINGO

نرم‌افزار LINGO یکی از قوی‌ترین نرم‌افزارهای حل مسائل بهینه‌سازی خطی و غیرخطی است و توسط شرکت LINDO تهیه و ارائه می‌گردد (Lindo, 2022). LINGO دارای مزایای زیادی است که آن را از سایر

چندلایه عموماً دو نوع سیگنال، یعنی سیگنال‌هایی که در مسیر رفت (از لایه ورودی به طرف لایه خروجی) و دیگری سیگنال‌های برگشت (از لایه خروجی به طرف لایه ورودی) حرکت می‌کنند، استفاده می‌شود که به ترتیب به نام سیگنال تابعی و خطا معروف هستند. تنظیم پارامترها در شبکه پرسپترون چندلایه توسط سیگنال خطا و سیگنال ورودی صورت می‌گیرد. تعیین تعداد لایه‌ها، نرون‌ها و توابع تحریک موجود در آن‌ها از مهم‌ترین مسائل در مدل‌سازی با شبکه عصبی مصنوعی است (Menhaj *et al.*, 2007).

نحوه آموزش در شبکه‌ی عصبی پرسپترون به صورت آموزش نظارت شده است که در آن ضمن مقایسه خروجی‌های واقعی و خروجی‌های به دست آمده از شبکه عصبی در تکرارهای مختلف، سعی می‌شود میزان اختلاف موجود بین آن‌ها (میزان خطا) به کمترین میزان ممکن کاهش پیدا کند. E تابع خطا است که در رابطه ۱ مشخص گردیده. (Araghinejad, 2013)

$$E = \frac{1}{n} \sum_{p=1}^n (y_p - \hat{y}_p)^2 \quad (1)$$

در این رابطه n بیانگر تعداد داده‌های ورودی (یا خروجی) و  $y_p$  خروجی‌های واقعی و  $\hat{y}_p$  خروجی‌های بدست آمده از شبکه عصبی است. اثبات شده است که یک شبکه پرسپترون سه‌لایه با تابع تحریک سیگموئید در لایه میانی و تابع تحریک خطی در لایه خروجی، قادر است هر نوع رابطه غیرخطی بین ورودی و خروجی را آموزش دیده و تقریب بزند؛ و به همین خاطر به آن تقریب زننده جهانی<sup>۱۰</sup> گفته می‌شود (Araghinejad, 2013).

### مدل برازش K همسایه نزدیک‌تر (K Nearest Neighbour)

K همسایه نزدیک‌تر یا به اختصار KNN یک روش ناپارامتری است که در کلاس‌بندی داده‌ها و تشخیص الگو مورداستفاده قرار می‌گیرد (Brownlee, 2016). در روش برازش K همسایه نزدیک‌تر، شباهت بین داده‌های جدید پیش‌بینی و داده‌های مشاهداتی به منظور به دست آوردن

برای ساخت این شبکه داده‌ها به دو دسته تقسیم شدند؛ ۸۰ درصد داده‌ها برای فرایند آموزش (واسنجی) و از ۲۰ درصد باقی داده‌ها برای آزمون (صحت‌سنجی) استفاده شد.

در مدل KNN از فرمول فاصله اقلیدسی وزن‌دار بر اساس عملیات تحلیل حساسیت استفاده شده است. همچنین مشاهده شد که با انتخاب ۱۰ همسایه، بهترین نتیجه (کمترین خطا) حاصل گردید.

### ساختار مدل بهینه‌سازی

قیود در نظر گرفته شده در مدل بهینه‌سازی این پژوهش عبارت‌اند از:

تنها ۶ چاه از ۱۰ چاه پمپاژ مشخص شده در شکل ۴ فعال هستند.

مجموع دبی تمامی چاه‌های تزریق برابر با دبی برداشت چاه تخلیه و برابر با  $14/159 \text{ (m}^3/\text{day)}$  است.

نرخ پمپاژ هر کدام از چاه‌های تزریق بین  $0/566$  تا  $4/106 \text{ (m}^3/\text{day)}$  است.

حداقل درصد پاک‌سازی قابل قبول  $95\%$  است.

باتوجه به اینکه مجموع دبی پمپاژ تمامی چاه‌های تزریق برابر با دبی برداشت چاه تخلیه برابر است و سایر هزینه‌ها (مانند هزینه‌های طراحی و اجرا) باتوجه به شروط بیان شده ثابت هستند، در نتیجه تنها عاملی که مقدار هزینه‌ها را افزایش یا کاهش می‌دهد و به‌وسیله آن می‌توان فرایند SEAR را به لحاظ هزینه بهینه کرد، زمان انجام این فرایند است. افزایش زمان انجام فرایند SEAR باعث افزایش مقدار مواد شیمیایی (سورفکتانت) مصرف شده می‌شود؛ همچنین هزینه‌های جاری (همانند هزینه‌های آب و برق و ...) نیز افزایش پیدا می‌کند. باتوجه به مطالب گفته شده، مدل بهینه‌سازی مربوط به این مسئله به صورت زیر است:

$$\text{Minimiz } Z = T \quad (3)$$

$$ST :$$

$$Q_{\min} = 0.566 \quad (4)$$

نرم‌افزارهای مشابه متمایز می‌سازد؛ از جمله می‌توان به ابزارها و حل‌کننده‌های قدرتمند، مستندات گسترده برای کاربران، زبان برنامه‌نویسی جبری (مدل‌سازی جبری)، ابزارهای مناسب برای کار با داده‌ها، شناسایی نوع مدل (خطی یا غیرخطی) توسط نرم‌افزار، نوشتن برنامه‌های پیچیده و بزرگ تنها با تعداد محدودی دستور ساده، و امکان ارتباط و تعامل با بانک‌های اطلاعاتی و نرم‌افزارهای صفحه گسترده مانند Excel اشاره نمود.

### فرضیات مدل‌های اجرا شده در نرم‌افزار

#### UTCHEM

در تمامی سناریوهای اجرا شده در نرم‌افزار UTCHEM سه محدودیت همواره رعایت شده است که عبارت‌اند از:

مجموع دبی تمامی چاه‌های تزریق برابر با دبی برداشت چاه تخلیه و برابر با  $14.159 \text{ (m}^3/\text{day)}$  است.

زمان اجرای هر کدام از سناریوها بین ۱۵ تا ۴۰ روز است. نرخ پمپاژ هر کدام از چاه‌های تزریق بین  $0/566$  تا  $4/106 \text{ (m}^3/\text{day)}$  است.

ابتدا باتوجه به محدودیت‌های گفته شده، ۲۵۰ سناریو مختلف ساخته شد؛ سپس هر کدام از آنها در نرم‌افزار UTCHEM اجرا گردیده و درصد پاک‌سازی‌شان به دست آمد. در ادامه به منظور بهره‌بردن از مدل‌های جایگزین، از دو روش MLP-ANN و KNN استفاده گردید.

### ساختار مدل‌های جایگزین

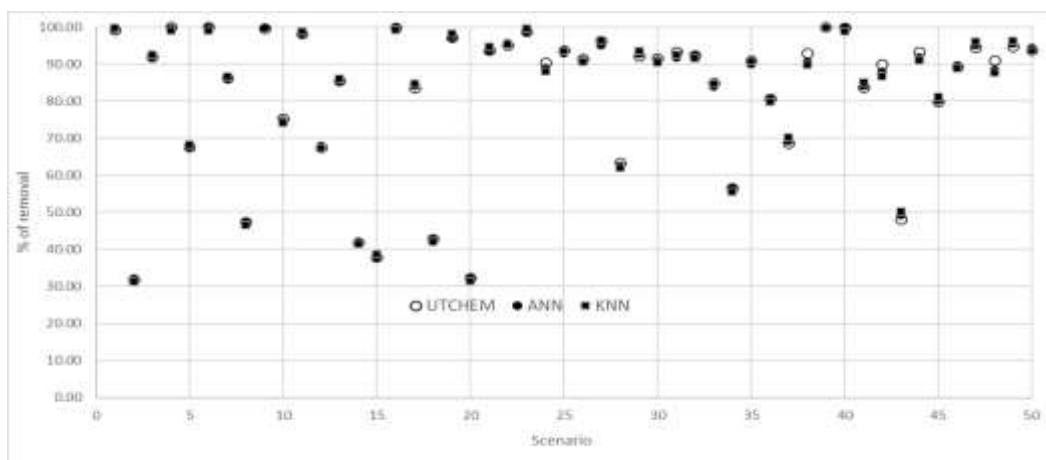
به‌منظور ساخت و انجام محاسبات دو مدل جایگزین MLP-ANN و KNN از نرم‌افزار Matlab 2018b استفاده شده است. شبکه عصبی دارای ۱ لایه پنهان (میانی) با ۷ نرون است. در واقع عملکرد شبکه با تعداد نرون‌های متفاوت در لایه میانی (از ۱ تا ۲۰ نرون) بررسی گردید و مشخص شد که بهترین عملکرد، یعنی کمترین جذر میانگین مربعات خطا، وقتی حاصل شد که لایه میانی دارای ۷ نرون بوده است. همچنین از تابع انتقال سیگموئید در لایه میانی و تابع انتقال خطی در لایه خروجی به‌عنوان تابع انتقال یا تحریک استفاده گردید.

بين ۰/۵۶۶ تا ۴/۱۰۶ (m<sup>3</sup>/day) و دبي مجموع آن‌ها برابر با ۱۴/۱۵۶ (m<sup>3</sup>/day) باشد، و در کوتاه‌ترين زمان به حداقل درصد پاک‌سازي موردنظر، يعني ۹۵ درصد برسد.

### نتايج و بحث

#### مقايسه ميزان دقت دو مدل جايزين

پس از ساخت و آموزش مدل‌هاي جايزين، درصد پاک‌سازي ۵۰ سناريو جديد که در فرايند ساخت و آموزش مدل‌ها نقشي نداشتند توسط اين دو مدل جايزين MLP-ANN و KNN و همچنين نرم‌افزار UTCHEM محاسبه شدند. شکل ۶ مقادير درصد پاک‌سازي حاصل از مدل UTCHEM و دو مدل جايزين MLP-ANN و KNN را نشان مي‌دهد. مقدار خطاي (RMSE) برآورد درصد پاک‌سازي دو مدل جايزين MLP-ANN و KNN نسبت به خروجي UTCHEM به ترتيب برابر ۰/۶۷ و ۱/۶۶ به دست آمد.



شکل ۶- درصد پاک‌سازي به‌دست‌آمده از UTCHEM و دو مدل جايزين MLP-ANN و KNN

Fig. 6- Percentage of removal obtained from UTCHEM and two alternative models MLP-ANN and KNN

بر اساس نتايج به‌دست‌آمده از نرم‌افزار LINGO در سناريو بهينه، چاه‌هاي شماره ۴، ۷، ۸ و ۹ غيرفعال بوده و چهار چاه شماره ۱، ۳، ۵ و ۱۰ به ترتيب بيشتري دبي پمپاژ را دارند. اين سناريو ضمن رعايت تمامي قيود، در کمترین مدت‌زمان يعني ۳۰.۴ روز محدوده مورد بررسي را به ميزان ۹۵٪ پاک‌سازي مي‌کند. در جدول ۲ دبي پمپاژ هرکدام از چاه‌ها در سناريو بهينه مشخص شده است.

$$Q_{max} = 4.106 \quad (6)$$

$$\sum_{i=1}^{10} Z_i = 6; Z_i\text{'s are binary (0,1) variables} \quad (7)$$

$$\sum_{i=1}^{10} Z_i \times Q_i = 14.159 \quad (8)$$

$$(Z_i \times Q_{min}) \leq Q_i \leq (Z_i \times Q_{max}) \quad (9)$$

$$CP = f(T, Q_1, Q_2, Q_3, Q_4, Q_5, Q_6, Q_7, Q_8, Q_9, Q_{10}) \geq 95 \quad (10)$$

در مدل بهينه‌سازي فوق، تابع هدف کمينه کردن زمان است. در اين مدل T نشان‌دهنده کل زمان اجرائي عمليات پاک‌سازي و Q بيانگر دبي چاه‌ها مي‌باشد. متغيرهاي Z متغير باينري هستند و درصورتی که براي یک چاه صفر باشند، يعني آن چاه غيرفعال است و اگر برابر یک باشند، يعني آن چاه فعال است. در اين مدل بهينه‌سازي تابع f بر اساس نرم‌افزار UTCHEM و يا مدل‌هاي جايزين، مقدار درصد پاک‌سازي حاصل شده از اجرائي فرايند SEAR را محاسبه مي‌کند. سناريو بهينه، سناريویی است که در آن ۶ چاه از ۱۰ چاه فعال بوده و نرخ پمپاژ هرکدام

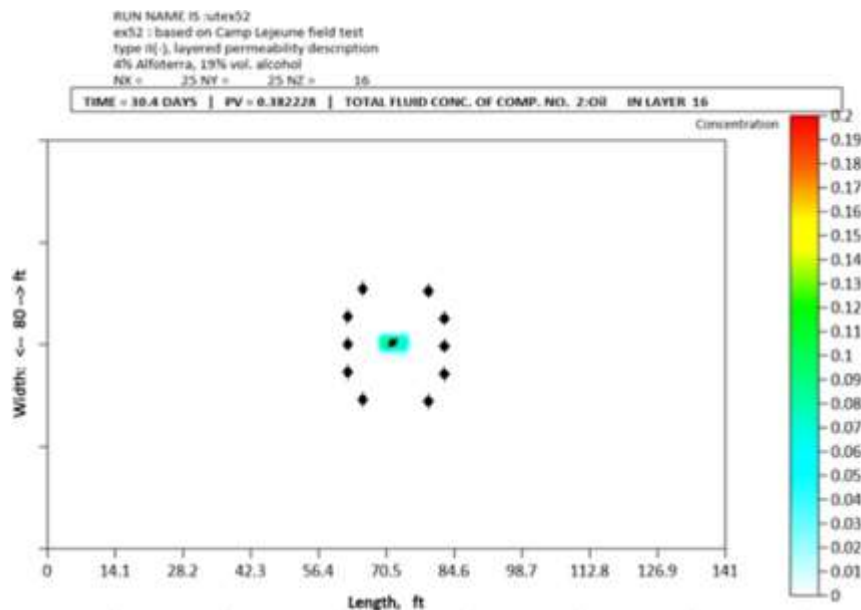
باتوجه به نتايج حاصل شده، خطاي جذر ميانگين مربعات مدل جايزين ساخته شده بر اساس مدل ANN کمتر از مدل جايزين KNN است و اين موضوع بيانگر عملکرد بهتر آن مي‌باشد؛ بر همين اساس در ادامه به‌منظور بهينه‌سازي فرايند SEAR با هدف کاهش هزينه‌ها، از مدل جايزين ANN استفاده شده است.

#### سناريو بهينه فرايند SEAR

جدول ۲- نرخ تزریق چاه‌ها در سناریوی پاک‌سازی بهینه  
Table 3. Injection rate of wells in optimal remediation scenario

شماره چاه Well number	نرخ تزریق (m <sup>3</sup> /day) Injection rate (m <sup>3</sup> /day)
1	4.105
2	1.384
3	3.059
4	0
5	2.582
6	1.216
7	0
8	0
9	0
10	1.809

برای اطمینان از صحت درصد پاک‌سازی به دست آمده از این سناریو، مجدداً از نرم‌افزار UTCHEM استفاده گردید. پس از اجرا مشاهده شد که درصد پاک‌سازی صورت گرفته بر اساس این سناریو برابر با ۹۴.۶۸ درصد می‌باشد. تفاوت اندک موجود نیز مربوط به خطای مدل جایگزین است. در شکل ۷ توزیع آلودگی در لایه ۱۶ (لایه تحتانی) پس از اجرای سناریو بهینه پاک‌سازی به همراه محل قرارگیری چاه‌ها مشخص شده است.



شکل ۷- توزیع آلودگی در لایه ۱۶ بعد از انجام فرایند SEAR و محل قرارگیری چاه‌ها

Fig. 7- The distribution of contamination in the layer 16 after performing the SEAR process and the location of the wells

نرم‌افزار LINGO مشخص شد که چاه‌های موجود در بالادست و پایین دست منطقه آلوده (چاه‌هایی که در راستای جهت جریان آب زیرزمینی هستند) بیشترین تأثیر را در فرایند پاک‌سازی دارند. بر اساس نتایج حاصل، با اجرای سناریوی بهینه، در مدت حدود ۳۰ روز به ۹۵ درصد پاک‌سازی می‌توان رسید.

باید به این نکته نیز اشاره کرد که در انجام این پژوهش فرضیات اولیه‌ای از قبیل ثابت بودن مجموع دبی‌ها، مشخص بودن محل ۱۰ چاه و فعال بودن ۶ چاه از ۱۰ چاه در سناریو بهینه در نظر گرفته شده است؛ علت این موضوع این بوده است که در صورت عدم استفاده از این فرضیات، برای آموزش و ساخت مدل‌های جایگزین نیاز به چندین هزار داده

### نتیجه‌گیری

در این پژوهش از دو روش یادگیری ماشین شبکه عصبی مصنوعی پرسپترون چندلایه (MLP-ANN) و K همسایگی نزدیک تر (KNN) به عنوان مدل‌های جایگزین بر اساس یادگیری ماشین به منظور شبیه‌سازی داده مبنای فرایند SEAR در پاک‌سازی DNAPL استفاده گردید. از نرم‌افزار LINGO به منظور بهینه‌سازی این فرایند استفاده شد تا علاوه بر دقت مناسب، مدت زمان اجرای مدل کاهش یابد. از داده‌های و اطلاعات مربوط به پروژه Camp Lejeune که به عنوان مطالعه موردی استفاده گردید. مشاهده شد که دقت هر دو مدل جایگزین مناسب بوده و دقت مدل ANN بهتر از KNN بود. همچنین بر اساس نتایج به دست آمده از و

برای پژوهش‌های آتی پیشنهاد می‌شود.

### پی‌نوشت‌ها

- <sup>1</sup> Dense Non-Aqueous Phase Liquid
- <sup>2</sup> Surfactant Enhanced Aquifer Remediation
- <sup>3</sup> Machine Learning
- <sup>4</sup> Root Mean Square Error
- <sup>5</sup> Light Non-Aqueous Phase Liquid
- <sup>6</sup> Hydrophilic
- <sup>7</sup> Hydrophobic
- <sup>8</sup> Geosystem
- <sup>9</sup> Biases
- <sup>10</sup> Universal Approximator
- <sup>11</sup> Sensitivity Analysis

Araghinejad, S., 2014. Data-Driven Modeling: Using MATLAB® in Water Resources and Environmental Engineering. Springer Dordrecht Heidelberg New York London.

Battelle and Duke Engineering and Services (Battelle/DE&S), 2002. Surfactant-Enhanced Aquifer Remediation (SEAR) Design Manual, NFESC Technical Report TR-2206-ENV, Report Prepared for the Naval Facilities Engineering Service Center.

Bear, J. and Cheng, A. H. D., 2010. Modeling groundwater flow and contaminant transport. Springer Science & Business Media.

Brownlee, J., 2016. Master Machine Learning Algorithms: Discover how they Work and Implement them from Scratch (Machine Learning Mastery).

Camp Lejeune., 1999. FINAL Cost and Performance Report for Surfactant-Enhanced DNAPL Removal at Site 88, Marine Corps Base Camp Lejeune, North Carolina.

Delshad, M., Pope, G.A. and Sepehrnoori, K., 1996. A compositional simulator for modeling surfactant enhanced aquifer remediation, 1 formulation. J. Contam. Hydrol. 23, 303–327. [https://doi.org/10.1016/0169-7722\(95\)00106-9](https://doi.org/10.1016/0169-7722(95)00106-9).

اولیه خواهد بود که این موضوع باتوجه به مدت زمان موردنیاز برای اجرای هر سناریو در نرم‌افزار UTCHEM دشوار و تا حدی غیرممکن است.

در انتها انجام فعالیت‌هایی از قبیل تحقیق و بررسی بر روی ترکیبات جایگزین سورفکتانت به منظور حداقل‌سازی هزینه‌ی فرایند SEAR، تحقیق و بررسی بر روی تغییرات فیزیکی مثل گرم کردن سورفکتانت قبل از تزریق بر روی عملکرد فرایند SEAR و مدل‌سازی سایر روش‌های پاک‌سازی آب‌های زیرزمینی، نظیر استخراج بخار از خاک

### منابع

Foster, S., Hirata, R., Gomes, D., D'Elia, M. and Paris, M., 2002. Groundwater quality protection: a guide for water utilities, municipal authorities, and environment agencies. World Bank, Washington, DC.

Harun, S. and Ahmat, N.I. And Kassim, A. H. M., 2002. Artificial neural network model for RainfallRunoff Relationship. Journal Technology, 37, 1-12.

Hou, Z., Lu, W., Xue, H. and Lin, J., 2017. A comparative research of different ensemble surrogate models based on set pair analysis for the DNAPL-contaminated aquifer remediation strategy optimization. Journal of Contaminant Hydrology.

Huling, S. G. and Weaver, J. W., 1991. Dense nonaqueous phase liquids. Superfund Technology Support Center for Ground Water, Robert S. Kerr Environmental Research Laboratory.

Jiang, X., Lu, W., Hou, Z., Zhao, H. and Na, J., 2015. Ensemble of surrogates-based optimization for identifying an optimal surfactant-enhanced aquifer remediation strategy at heterogeneous DNAPL-contaminated sites. Comput. Geosci. 84, 37–45. <https://doi.org/10.1016/j.cageo.2015.08.003>.

Kueper, B.H. and McWhorter, D.B., 1991. The

- behavior of dense, nonaqueous phase liquids in fractured clay and rock. *Ground Water* 29, 716–728. <https://doi.org/10.1111/j.1745-6584.1991.tb00563.x>.
- Kueper, B. H., Stroo, H. F., Vogel, C. M. and Ward, C. H. (Eds.), 2014. Chlorinated solvent source zone remediation. Springer.
- Lerner, D.N., Kueper, B.H. and Wealthall, G.P., 2003. An illustrated handbook of DNAPL transport and fate in the subsurface. Research Report. Environment Agency, Bristol, UK.
- Lindo Concepts, 2022. LINGO 19.0 - Optimization Modeling Software for Linear, Nonlinear, and Integer Programming. <https://www.lindo.com/index.php/products/lingo-and-optimization-modeling>. Chicago, IL, 60642. USA.
- Luo, J., Lu, W., Xin, X. and Chu, H., 2013. Surrogate model application to the identification of an optimal surfactant-enhanced aquifer remediation strategy for DNAPL contaminated sites. *Journal of Earth Science*. 24, 1023–1032. <https://doi.org/10.1007/s12583-013-0395-1>.
- Luo, J. and Lu, W., 2014. Comparison of surrogate models with different methods in groundwater remediation process. *Journal of earth system science*.
- Mayer, A. and Hassanizadeh, s. m., 2005. SOIL AND GROUNDWATER CONTAMINATION :NONAQUEOUS PHASE LIQUIDS-PRINCIPLES AND OBSERVATIONS. American Geophysical Union. Washington, DC.
- Mayer, A. and Endres, K.L., 2007. Simultaneous optimization of dense non-aqueous phase liquid (DNAPL) source and contaminant plume remediation. *J. Contam. Hydrol.* 91, 288–311. <https://doi.org/10.1016/j.jconhyd.2006.11.009>.
- Menhaj, M. B., 2007. "Fundamentals of Neural Networks", Computational Intelligence, Vol. 1, Amirkabir university of technology, Tehran, Iran.
- Mercer, J. W. and Cohen, R. M., 1990. A review of immiscible fluids in the subsurface: properties, models, characterization and remediation. *Journal of Contaminant Hydrology*.
- Moradian, M. and Sepehrifar, M., 2009. Improving the Accuracy of KNN Algorithm in Data Mining Using Dependency Laws, 15th Annual Computer Conference of Iranian Computer Association.
- Nyer, E.K., 2009. Groundwater Treatment Technology. Published by Van Nostrand Reinhold ISBN 13: 9780442005627
- Ouyang, Q., Lu, W., Lin, J., Deng, W. and Cheng, W., 2017a. Conservative strategy-based ensemble surrogate model for optimal groundwater remediation design at DNAPLs-contaminated sites. *Journal of Contaminant Hydrology*.
- Ouyang, Q., Lu, W., Miao, T., Deng, W., Jiang, C. and Luo, J., 2017b. Application of ensemble surrogates and adaptive sequential sampling to optimal groundwater remediation design at DNAPLs-contaminated sites. *J. Contam. Hydrol.* 207, 31–38. <https://doi.org/10.1016/j.jconhyd.2017.10.007>.
- Qin, X.S., Huang, G.H., Chakma, A., Chen, B. and Zeng, G.M., 2007. Simulation-based process optimization for surfactant-enhanced aquifer remediation at heterogeneous DNAPL-contaminated sites. *Sci. Total Environ.* 381, 17–37. <https://doi.org/10.1016/j.scitotenv.2007.04.011>.
- Qin, X. S., Huang, G. H. and He, L., 2009. Simulation and optimization technologies for

petroleum waste management and remediation process control. *Journal of Environmental Management*.

Shaafi, E., 2013. *Groundwater Rehabilitation System Design Optimization*, Master Thesis in Civil Engineering - Water, Groundwater Orientation, December 2013.

Shams, R., Alimohammadi, S. and Yazdi, J., 2021. Optimizing surfactant-enhanced aquifer remediation based on Gaussian process surrogate model in DNAPL-contaminated sites considering different wells patterns. *Groundwater for Sustainable Development*.

Shmueli, G., Patel, N.R. and Bruce, P.C., 2011. *Data Mining for Business Intelligence: Concepts, Techniques, and Applications in Microsoft Office Excel with XLMiner*. Wiley.

Shobeyri, M., 2019. *Optimum Design of Flood Control Levees, Combining Evolutionary Algorithm and Discrete Differential Dynamic*

*Programming*. Shahid Beheshti University Faculty of Water & Environmental Engineering. february2019.

Suthersan, S.S., Horst, J., Schnobrich, M., Welty, N. and McDonough, J., 2016. *Remediation Engineering, Remediation Engineering: Design Concepts*, second ed. CRC Press. <https://doi.org/10.1201/9781315367088>.

Swikatek, J., Borzowski, L. and Wilimowska, Z., 2019. *Information Systems Architecture and Technology: Proceedings of 40th Anniversary International Conference on Information Systems Architecture and Technology – ISAT 2019: Part II, Advances in Intelligent Systems and Computing*. Springer International Publishing.





Environmental Sciences Vol.20 / No.4 / Winter 2023

125-142

Original Article

## Optimization of SEAR method in DNAPL contaminated aquifers remediation using surrogate models

Hossein Azhdari Bajestani\* and Saeed Alimohammadi

Department of Water Resources Engineering, Faculty of Civil, Water, and Environmental Engineering, Shahid Beheshti University, Tehran, Iran

Received: 2022.02.20 Accepted: 2022.05.02

**Azhdari Bajestani, H. and Alimohammadi, S., 2023.** Optimization of SEAR method in DNAPL contaminated aquifers remediation using surrogate models. *Environmental Sciences*. 20(4): 125-142.

**Introduction:** Dense Nonaqueous Phase Liquids (DNAPL) are the most common types of groundwater pollution. Surfactant Enhanced Aquifer Remediation (SEAR) is one of the most common methods of DNAPL-contaminated aquifer remediation. Due to the high cost of the chemicals used in this method (surfactants or cosolvents), it is necessary to choose the appropriate wells pattern, and the optimal pumping rates. UTCHEM simulation software has the ability to model the fate and transport of DNAPL and the application of the SEAR method. The main problem with this software is the long time required to run multiple scenarios when using optimization algorithms. The purpose of this study is to use two machine learning methods (Artificial Neural Network and K nearest neighbor) as surrogate simulation model and imbedding the best one into the LINGO software to optimize the SEAR method.

**Material and methods:** in the implementation of The SEAR method, the quantitative and qualitative aquifer data are required to model how to diffuse, transmit and delete DNAPLs in UTCHEM software. For this purpose information from the Camp Lejeune site in North Carolina, USA were used. In this study, by examining a variety of alternative models based on machine learning methods and implementing 250 different scenarios in UTCHEM software, two models, Artificial Neural Network (ANN) method, and k-nearest neighbors (KNN) were used to simulate the SEAR method and developing alternative model. In order to validate the two alternative models, 50 new scenarios were implemented in UTCHEM software and their percentage of removal was obtained. Also, using two alternative models, the percentage of removal of 50 scenarios were determined. In order to evaluate the performance of alternative models, the root mean square error (RMSE) was used and was compared with the results of other researchs. Finally, an alternative model with more accuracy was used in LINGO software to optimize the Surfactant Enhanced Aquifer Remediation method (SEAR).

---

\* Corresponding Author: *Email Address.* s\_alimohammadi@sbu.ac.ir



**Results and discussion:** RMSE values in the results obtained from alternative models ANN and KNN in the validation stage were 0.67 and 1.66 respectively, which indicates the high accuracy of both alternative models, especially ANN. The average run time of each UTCHEM software in this study was 45 minutes, while in the alternative model it was reduced to a few seconds; LINGO software also examined about 21,500 different scenarios in 30 minutes to determine the optimal scenario, while the time required for this task is more than 16,000 hours if the alternative model is not used. Based on the position and discharge of active wells in the optimum scenario, it was found that firstly the existing wells upstream and downstream of the Pollution have the most impact on the remediation and secondly, the time factor is more effective than the wells pumping discharge in the remediation. The optimized scenario obtained in this study remediates the DNAPL-contaminated area by up to 95% at a lower cost than the costs reported in the Camp Lejeune project over a period of 30 days.

**Conclusion:** Based on the results obtained in this study, it was found that the use of machine learning algorithms such as ANN and KNN, along with LINGO optimization software, which is one of the most powerful software for solving linear and nonlinear optimization problems, in addition to having the right accuracy, significantly reduces the time required to find the optimal scenario.

**Keywords:** Surfactant enhanced aquifer remediation, DNAPL, Artificial neural network MLP, K nearest neighbours, LINGO software.

