

تخمین ارتفاع سطح LNAPL در آبخوان‌های آلوده به نفت با استفاده از برنامه‌نویسی بیان ژن (GEP)، سیستم استنتاج فازی (ANFIS) و روش رگرسیون چند متغیره (MLR)

فاطمه ابراهیمی^(۱)، محمد نخعی^(۲)، حمیدرضا ناصری^(۳) و کمال خدایی^(۴)

۱. دانشجوی دکتری هیدروژئولوژی، دانشکده علوم زمین، دانشگاه خوارزمی، تهران
۲. استاد گروه زمین‌شناسی کاربردی، دانشکده علوم زمین، دانشگاه خوارزمی، تهران
۳. استاد گروه زمین‌شناسی معدنی و آب، دانشکده علوم زمین، دانشگاه شهید بهشتی، تهران
۴. استادیار، گروه زمین‌شناسی محیطی، پژوهشکده علوم پایه کاربردی جهاد دانشگاهی

تاریخ دریافت: ۹۸/۰۷/۰۷

تاریخ پذیرش: ۹۸/۰۹/۱۱

چکیده

یکی از مهم‌ترین نگرانی‌ها در آبخوان‌های مجاور به تاسیسات نفتی، نشت LNAPL‌ها می‌باشد. بازیافت LNAPL‌ها همواره مشکل و پرهزینه است. نخستین مرحله در برنامه‌ریزی چنین سیستم‌هایی، تعیین اهداف طراحی می‌باشد، اغلب بیشینه‌سازی برداشت آلاینده، و کمینه‌سازی هزینه به‌عنوان مهم‌ترین اهداف طراحی در نظر گرفته می‌شوند. شناسایی ضخامت LNAPL و نوسانات آن می‌تواند تعیین‌کننده روش بازیافت، بیشینه‌سازی برداشت و کاهش هزینه آن شود. در این مطالعه از سه روش برنامه‌نویسی بیان ژن^(۱)، سیستم استنتاج تطبیقی فازی^(۲)، و رگرسیون چند متغیره^(۳)، برای تخمین و پیش‌بینی ارتفاع سطح LNAPL استفاده شده است. متغیرهای ورودی شامل ارتفاع سطح آب زیرزمینی و نرخ تخلیه LNAPL و متغیر خروجی ارتفاع سطح LNAPL می‌باشد. نتایج اجرای سه مدل توسط پارامترهای آماری مورد تحلیل و بررسی قرار گرفت و مشخص شد که برنامه‌نویسی بیان ژن دارای نتایج بهتری می‌باشد و می‌تواند به‌طور موفقیت‌آمیزی در پیش‌بینی نوسانات سطح LNAPL در فرایندهای بازیافت مورد استفاده قرار گیرد. همچنین توسط مدل GEP یک معادله برای پیش‌بینی سطح LNAPL ارائه شد که می‌توان آن را در سرچاه برای پیش‌بینی ارتفاع سطح LNAPL استفاده کرد.

واژه‌های کلیدی: نوسانات LNAPL، برنامه‌نویسی بیان ژن، سیستم استنتاج تطبیقی فازی، رگرسیون چند متغیره.

1. Light Non-Aqueous Phase Liquid (LNAPL)
2. Gene Expression Programming (GEP)
3. Adaptive Neuro-Fuzzy Inference System (ANFIS)
4. Multivariate Linear Regression (MLR)

* نویسنده مرتبط: Ebrahimii.edu@gmail.com

مقدمه

این مدل ها شامل هوش مصنوعی^۱، هوش محاسباتی^۲، محاسبات نرم^۴، یادگیری ماشین^۵ و داده کاوی^۶ می باشند (Solomatine et al., 2009).

در دهه های اخیر، تکنیک های جدید مدل های داده مبنا همچون هوش مصنوعی به عنوان تکنیک جایگزین در مدل سازی سیستم های هیدروژئولوژی استفاده شده است (Emamgholizadeh et al., 2014). چهار مورد از مهم ترین دسته بندی های هوش مصنوعی که در شبیه سازی های هیدروژئولوژیک استفاده شده اند شامل: روش های یادگیری ماشین، مجموعه های فازی، محاسبات تکاملی و موجک می باشند (Yaseen et al., 2015). همکاران (۱۳۹۸). هوش مصنوعی در مقایسه با سایر مدل ها، به دلیل اجرای راحت و سریع و عدم نیاز به خصوصیات سیستم هیدروژئولوژیکی، توانسته توجه هیدروژئولوژیست ها را به خود جلب کند (Coppola et al., 2005). تکنیک های مختلفی از مدل های شبکه عصبی و هوش مصنوعی برای اهداف مختلفی همچون: شبیه سازی و پیش بینی سطح آب در سیستم های هیدرولیکی (Coppola et al., 2005; Coppola et al., 2007; Nayak et al., 2006; Cimen and Kisi, 2009; Noori et al., 2010; Guldal and Tongal, 2010; Adamowski and Chan, 2011; Mpallas et al., 2011; Yoon et al., 2011; Karimi et al., 2012; Kisi et al., 2012; Danandeh Mehr et al., 2014; Yaseen et al., 2016a, 2016b, 2017; Tiri et al., 2018; Ghorbani et al., 2017)، تعیین پارامترهای هیدرولیکی (Azari and Samani, 2018) و تعیین خصوصیات فیزیکی و شیمیایی (Elzwayie et al., 2016; Sanikhani et al., 2018; Zaqoot et al., 2018) توسط هیدروژئولوژیست ها به کار گرفته شده است.

همچنین در سال های اخیر، استفاده از روش محاسبات تکاملی به دلیل سرعت و دقت بالا در پیش بینی سطوح آب

آب زیرزمینی یکی از منابع مهم نوشیدنی، خانگی و صنعتی در سرتاسر جهان می باشد. امروزه با پیشرفت صنعت، کیفیت آب زیرزمینی در معرض آلودگی قرار گرفته است. یکی از آلوده کننده های کیفی آب زیرزمینی نشت ترکیبات نفتی به آبخوان می باشد که به عنوان مهم ترین نوع آلودگی در آبخوان های مجاور به تاسیسات نفتی در نظر گرفته می شود (Nadim et al., 2000). برخی از هیدروکربن های نفتی که سبک تر از آب هستند از طریق زون غیراشباع نفوذ کرده و خود را به سطح ایستابی رسانده و تشکیل یک لایه معلق بر روی سطح ایستابی می دهند که LNAPL نامیده می شود. LNAPL ها یک فاز جداگانه و غیر همزیست با آب تشکیل می دهند و فضاهای خالی آبخوان را اشغال می کنند. سوخت های گازوئیل، دیزل و موتور و مواد مشابه نمونه هایی از LNAPL ها می باشند. اگر نشت LNAPL متوقف شود، آنگاه نوسانات LNAPL متأثر از نوسانات سطح ایستابی خواهد بود. از آنجایی که هیچ مطالعه ای در زمینه پیش بینی نوسانات LNAPL یافت نشد، بنابراین تمرکز اصلی در این بخش روی مطالعات پیش بینی نوسانات سطح آب می باشد. چندین روش برای شبیه سازی فرایندهای هیدروژئولوژی و پیش بینی سطح آب وجود دارد که شامل مدل های ریاضی، فیزیکی و تجربی می باشد. مدل های ریاضی نیازمند داشتن دانش کامل از خصوصیات زمین شناسی و ژئومورفولوژی آبخوان هستند، همچنین روش های فیزیکی شامل ساخت مدل های آزمایشگاهی کوچک مقیاس و اندازه گیری های مستقیم می باشد (ناصری و همکاران، ۱۳۹۹). اما مدل های تجربی بدون نیاز به این خصوصیات، تنها از طریق آنالیز داده های سری زمانی به دست می آیند. انواع مختلفی از مدل های سری زمانی می توانند در شبیه سازی های هیدروژئولوژی استفاده شوند اما باید توجه کرد که این مدل ها، نمی توانند برای مسائل غیرخطی استفاده شوند (Yaseen et al., 2015). با پیشرفت های اخیر هوش محاسباتی در زمینه یادگیری ماشین، روش های تجربی توسعه یافتند که مدل های داده مبنا امیده می شوند.

1. Data-Driven models (DDM)
2. Artificial Intelligence (AI)
3. Computational Intelligence (CI)
4. Soft Computing (SC),
5. Machine Learning (ML)
6. Data Mining (DM)

روش مطالعه

آنالیز داده‌ها

برداشت آلودگی LNAPL از آبخوان اغلب مشکل و پرهزینه است. یک از راه‌ها برای این منظور برداشت آلودگی و انتقال آن به سطح زمین می‌باشد. نخستین مرحله در برنامه‌ریزی چنین روش‌هایی، تعیین اهداف طراحی می‌باشد. اغلب کمینه‌سازی هزینه و بیشینه‌سازی برداشت آلاینده به‌عنوان مهم‌ترین اهداف طراحی در نظر گرفته می‌شود. منطقه مورد مطالعه در این تحقیق روستای اسماعیل‌آباد است. این روستا در جنوب تهران واقع شده است و در بخش‌های شمالی در مجاورت تاسیسات نفتی می‌باشد شکل ۱ موقعیت جغرافیایی محدوده مورد مطالعه را نشان می‌دهد. نشت LNAPL سبب آلوده شدن خاک و آب زیرزمینی این روستا شده است. بدین منظور چندین چاه برای بازیافت LNAPL حفر شده است. یکی از اهداف این مطالعه، تخمین و پیش‌بینی نوسانات سطح LNAPL درون چاه می‌باشد آن‌چنان‌که بتوان با کمک نتایج به‌دست‌آمده هزینه پمپاژ را کمینه کرد.

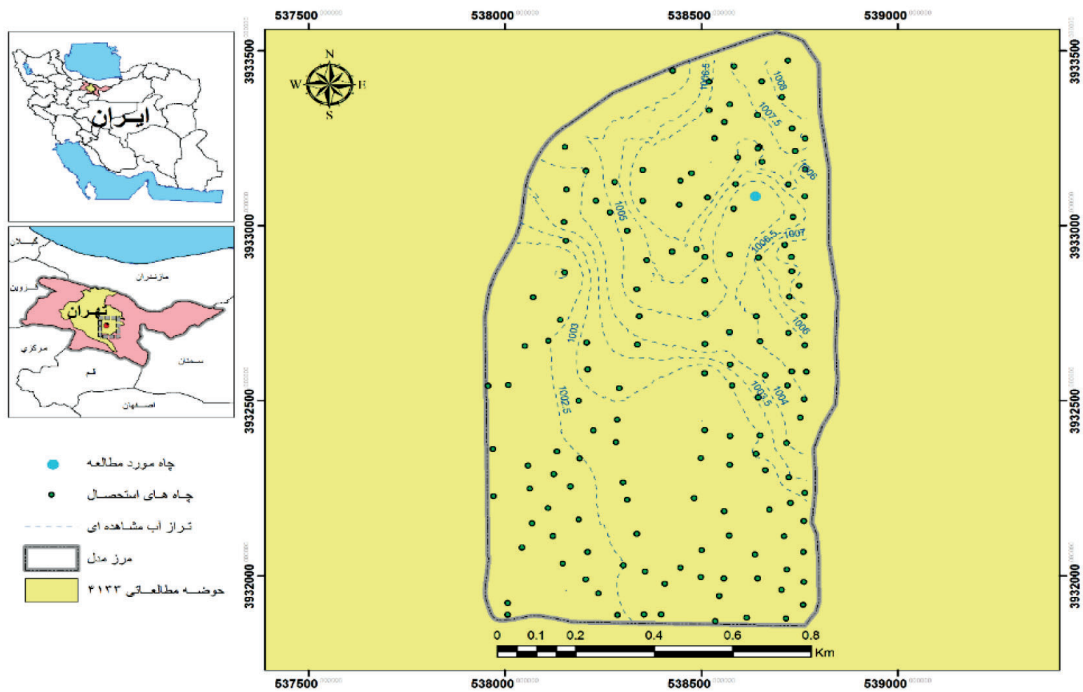
تصحیح سطح آب زیرزمینی

یکی از داده‌های ورودی مورد استفاده در پیش‌بینی نوسانات سطح LNAPL، ارتفاع سطح آب زیرزمینی می‌باشد. در آبخوان‌های آلوده به نفت، ارتفاع سطح آب مشاهده شده کمتر از سطح آب واقعی می‌باشد آن‌چنان‌که این سطح بایستی تصحیح شود. در اینجا از رابطه ۱ برای تصحیح ارتفاع سطح آب زیرزمینی استفاده شده است که ρ_n چگالی نفت، ρ_w چگالی آب، z_{aw} ارتفاع فصل مشترک سطح آب و هوا، z_{an} ارتفاع فصل مشترک سطح نفت و هوا، z_{nw} ارتفاع فصل مشترک سطح نفت و آب و h_w ارتفاع واقعی سطح آب زیرزمینی می‌باشد.

$$h_w = z_{aw} = \frac{\rho_n}{\rho_w} z_{an} + \left(1 - \frac{\rho_n}{\rho_w}\right) z_{nw} \quad (1)$$

زیرزمینی افزایش یافته است. این روش‌ها شامل برنامه‌نویسی ژنتیک^۱، برنامه‌نویسی بیان ژن، الگوریتم ژنتیک^۲ و الگوریتم‌های هوش ازدحامی می‌باشند. نتایج اجرای این مدل‌ها در پیش‌بینی سطوح آب نشان‌دهنده دقت بالاتر آنها در مقایسه با مدل‌های دیگر است (Yaseen et al., 2017; Ghorbani et al., 2017; Aytek et al., 2008; Wang et al., 2009, Karimi et al., 2012., Kisi et al., 2012, Danandeh Mehr et al., 2014). این مطالعه با هدف پیش‌بینی ارتفاع سطح LNAPL با استفاده از روش‌های هوش مصنوعی انجام شده است. ارتفاع سطح LNAPL به‌طور مستقیم متأثر از نوسانات سطح ایستابی است. هنگامی که سطح آب بالا می‌آید، لایه LNAPL سیال روی سطح ایستابی نیز به سمت بالا مهاجرت خواهد کرد و قطرات کوچک و جداگانه‌ای به نام LNAPL پسماند را درون منافذ خاک تشکیل خواهد داد که به راحتی قابل بازیافت نمی‌باشد. همچنین، هنگامی که سطح ایستابی به پایین می‌رود، لایه LNAPL نیز از درون منافذ زهکش شده و تا رسیدن به سطح ایستابی پایین خواهد رفت. به‌طور کلی نوسانات سطح آب زیرزمینی و LNAPL می‌تواند متغیر به تشکیل زون آلوده^۳ شده و موجبات گسترش آلودگی به نواحی غیر آلوده آبخوان و خاک را فراهم آورد. یک مولفه اساسی در اجرای سیستم‌های بازیافت در نواحی آلوده شامل تعیین نوسانات سطح آب و LNAPL است، آن‌چنان‌که آگاهی از ارتفاع دقیق این سطوح می‌تواند مدیریت صحیح عملیات بازیافت را به دنبال داشته باشد. همان‌گونه که در بالا دیده شد، تاکنون مطالعه‌ای برای آنالیز و پیش‌بینی نوسانات سطح LNAPL گزارش نشده است. این مقاله با هدف پیش‌بینی نوسانات سطح LNAPL صورت گرفته است که برای این منظور از سه روش ANFIS، GEP و MLR استفاده شده است و نتایج آنها با هم مورد مقایسه قرار گرفته است. در این مطالعه داده‌های سطح آب زیرزمینی و نرخ تخلیه به‌عنوان پارامترهای ورودی و سطح LNAPL به‌عنوان پارامتر خروجی در نظر گرفته شده است.

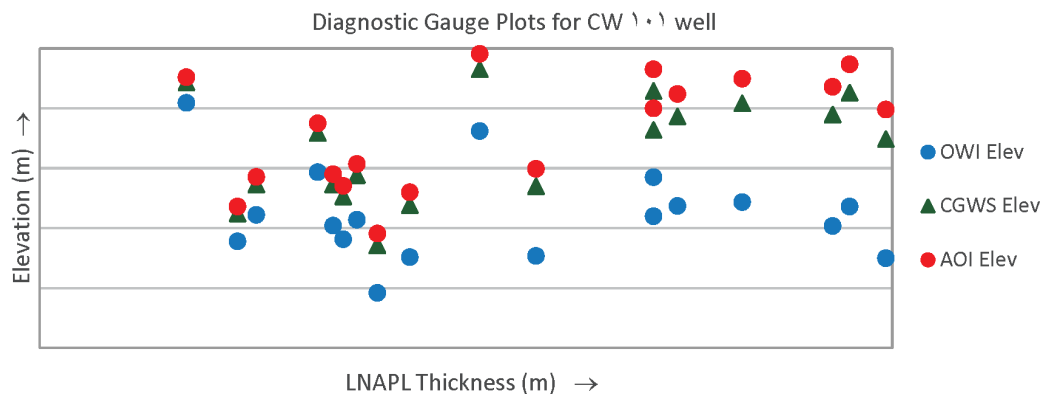
1. Genetic-Programming (GP)
2. Genetic-Algorithms (GA)
3. Smear zone



شکل ۱. نقشه موقعیت جغرافیایی منطقه مورد مطالعه (روستای اسماعیل آباد-ایران)

درون سازند استفاده می شود (Hawthorne, 2011). بر طبق این نمودار، ضخامت و ارتفاع فصل مشترک سطح AOI و CGWS در حال کاهش است، اما ارتفاع سطح فصل مشترک OWI ثابت است بنابراین وضعیت LNAPL در محدوده، محبوس و ارتفاع سطح آب زیرزمینی در حال افت می باشد.

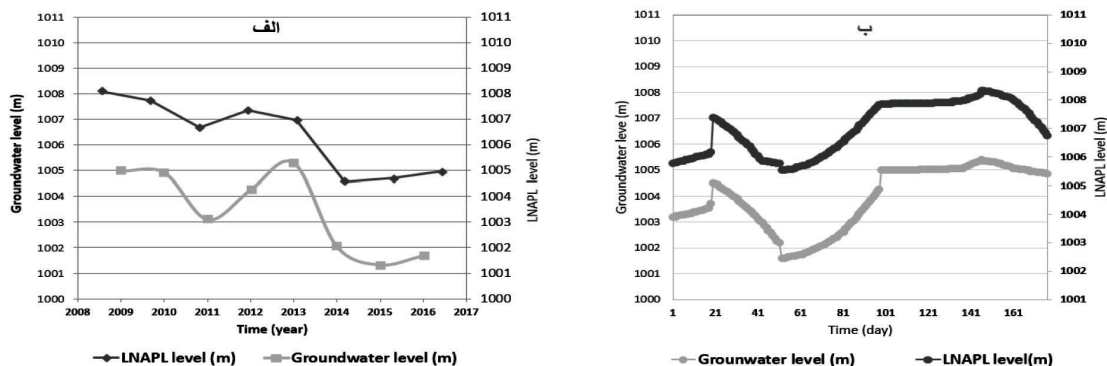
شکل ۲ نمودار وضعیت LNAPL را برای چاه مورد مطالعه نشان می دهد. این نمودار برخاسته از اطلاعات آماري ارتفاع سطوح هوا-نفت (AOI)، آب-نفت (OWI) و سطح آب تصحيح شده (CGWS) برای چاه ۱۰۱ در طول دوره زمانی هشت سال می باشد. از این نمودار به عنوان یک ابزار مهم در تعیین وضعیت LNAPL و ضخامت آن



شکل ۲. نمودار وضعیت LNAPL مربوط به چاه مورد مطالعه

و متغیر خروجی ارتفاع سطح LNAPL (LL) می باشد. در این مطالعه از داده های روزانه شش ماه استفاده شده است. از داده های موجود، داده های ۱۵۰ روز برای آموزش و داده های ۳۰ روز برای تست مدل استفاده شد. جدول ۱ پارامترهای آماری داده های مورد استفاده را نشان می دهد که S_x ، X_{min} ، X_{max} ، X_{mean} و CV به ترتیب دلالت بر میانگین، ماکزیمم، مینیمم، انحراف معیار و ضریب تغییرات داده های مورد استفاده دارند.

همچنین شکل ۳ هیدروگراف های سطح آب و LNAPL را در طول آماربرداری ۱۰ ساله و دوره مورد مطالعه نشان می دهد. از هیدروگراف، واضح است که تغییرات سطح LNAPL از تغییرات سطح آب پیروی می کند. بنابراین تغییرات در سطح نفت می تواند متأثر از تغییرات در سطح آب باشد. همچنین نرخ تخلیه LNAPL به عنوان پارامتر دیگری است که می تواند روی تغییرات سطح LNAPL تاثیرگذار باشد. بنابراین متغیرهای ورودی در این مطالعه شامل تغییرات سطح آب (WL) و نرخ تخلیه LNAPL (Q)



شکل ۳. هیدروگراف ارتفاع سطح آب و LNAPL (الف) برای دوره آماری ۱۰ ساله و (ب) دوره زمانی مورد مطالعه

جدول ۱. پارامترهای آماری داده های مورد مطالعه

مجموعه داده	پارامترهای آماری					
	X_{mean}	X_{max}	X_{min}	S_x	C_v	
دوره آموزش	WL (m)	۱۰۰۴	۱۰۰۵/۳۶	۱۰۰۱/۵۹	۱/۲۶	۰/۰۰۱۲۵
	LL (m)	۱۰۰۷/۲۱	۱۰۰۸/۳۱	۱۰۰۵/۵۱	۰/۸۷	۰/۰۰۰۸
	Q (m ³ /d)	۷/۹	۱۲/۹	۲/۵	۲/۱۶	۰/۲۷۳
دوره تست	WL (m)	۱۰۰۴/۴۵	۱۰۰۵/۳۳	۱۰۰۱/۷۵	۰/۸۸	۰/۰۰۰۸۷
	LL (m)	۱۰۰۷/۲۶	۱۰۰۸/۲۹	۱۰۰۵/۶۷	۰/۸۵	۰/۰۰۰۸۴
	Q (m ³ /d)	۸/۲۳	۱۰/۳۱	۴/۷۶	۲/۲۱	۰/۲۶
کل دوره	WL (m)	۱۰۰۴/۸۱	۱۰۰۵/۳۶	۱۰۰۱/۵۹	۱/۲۱	۱۲/۰۰
	LL (m)	۱۰۰۷/۱۹	۱۰۰۸/۳۱	۱۰۰۵/۵۱	۰/۸۶	۰/۰۰۰۸۵
	Q (m ³ /d)	۷/۹۶	۱۲/۸۹	۱۲/۵	۲/۱	۰/۲۶

در GEP مشابه تکامل بیولوژیکی می باشد (Ozbek et al., 2013; Ferreira, 2001a,b). GEP دارای چندین خصوصیت اساسی می باشد که در سال های اخیر سبب جذب محققین از علوم مختلف شده است. از جمله این خصوصیات می توان به ایجاد همبستگی برای مجموعه های از داده های بزرگ بدون داشتن هر دانش مقدماتی از ارتباط

مدل برنامه نویسی بیان ژن و کاربرد آن (Gene expression programming (GEP))

برنامه نویسی بیان ژن، یک روش به نسبت جدید می باشد که نخستین بار توسط Ferreira در سال ۲۰۰۱ پیشنهاد شد. این مدل، از قوانین بنیادی الگوریتم ژنتیک (GA) و برنامه نویسی ژنتیک استفاده می کند. ارزیابی هر دانشی

جدول ۲. مقادیر اپراتورهای ژنتیکی به کاررفته در مدل GEP

توصیف پارامتر	تنظیم پارامتر
مجموعه توابع مورد استفاده	+، -، *، /، Sqrt، x ² ، x ³
تعداد کروموزوم	۳۰
اندازه سر کروموزوم	۸
تعداد ژن	۳
نرخ جهش	۰/۰۴۴
نرخ RNC	۰/۰۵
نرخ تقاطع یک نقطه ای	۰/۳
نرخ تقاطع دونقطه ای	۰/۳
نرخ تقاطع ژن	۰/۱
نرخ جابجایی ژن	۰/۱
نرخ جابجایی IS	۰/۱

سیستم استنتاج تطبیقی فازی (adaptive neuro-fuzzy inference system technique (ANFIS))

تکنیک سیستم استنتاج تطبیقی فازی (ANFIS) به صورت مقدماتی توسط Jang (1993) ارائه شد. ANFIS همچنین نوعی از شبکه عصبی می باشد که وابسته به سیستم Takagi-Sugeno می باشد (Jang, 1993; Jang, 1991). ANFIS از الگوریتم های یادگیری شبکه هوش مصنوعی و قوانین منطق فازی برای طراحی ارتباط بین فضای متغیر ورودی و خروجی استفاده می کند. به همین دلیل مزایایی هر دو روش را دارد. یعنی از مجموعه قوانین «if-then» برای تخمین توابع غیرخطی استفاده می کند (Abraham, 2005). ANFIS دارای خصوصیات متعددی می باشد مثل توضیح رفتار سیستم های پیچیده توسط قوانین «if-then»، عدم نیاز به کارشناسی مقدماتی و اولیه داده ها، اجرای راحت و سریع با دقت بالا، انتخاب توابع عضویت بزرگ برای استفاده، توانایی های عمومی قوی و اجرای عالی که سبب تسهیل قوانین فازی می شود (Al-Hmouz et al., 2012). دو نوع سیستم استنتاج تطبیقی فازی وجود دارد که شامل روش Mamdani (Mamdani and Assilian, 1975) و روش Sugeno (Takagi and Sugeno, 1985) است. در روش Mamdani، توابع عضویت خروجی باید از مجموعه فازی انتخاب شد اما در روش Sugeno توابع عضویت خروجی به دو صورت خطی و یا ثابت می باشد (Shiri et al., 2013). در این مطالعه روش Sugeno برای

بین آنها، عدم نیاز به توابع از قبل تعیین شده، اجرای راحت و سریع، تهیه یک فرمولاسیون جهت نمایش ارتباط بین متغیرهای مستقل و وابسته و اجرای فرمول توسط هر زبان برنامه نویسی اشاره کرد (Nazari, 2012). الگوریتم GEP با انتخاب تابع تناسب، مجموعه توابع مورد استفاده، پارامترهای کنترلی، مجموعه متغیرهای ورودی و خروجی، تابع اتصال و اپراتورهای ژنتیکی شروع می شود. روش کار برای پیش بینی نوسانات سطح LNAPL در مدل GEP به این صورت می باشد: مرحله اول معرفی متغیرهای ورودی و خروجی است که در این مطالعه ارتفاع سطح آب و نرخ تخلیه LNAPL به عنوان متغیرهای ورودی و ارتفاع سطح LNAPL به عنوان متغیر خروجی در نظر گرفته شده است. مرحله دوم تنظیم توابع مورد استفاده است که برای این مسئله اپراتورهای جبری پایه ای و همچنین برخی از توابع ریاضی اصلی استفاده شد (جدول ۲). مرحله سوم انتخاب تابع تناسب است که در این مسئله از تابع ریشه میانگین مربع خطاها (RMSE) استفاده شد.

مرحله چهارم انتخاب اپراتورهای کنترلی می باشد. طول سر کروموزوم برابر با هشت و تعداد ژن های کروموزوم برابر با سه انتخاب شد که همواره همان مقادیر استفاده شده در اکثر مطالعات است (Ferreira 2001a, 2001b; Ferreira, 2006). مرحله پنجم شامل انتخاب تابع اتصال ژن ها می باشد که در اینجا از تابع جمع استفاده شد و مرحله ششم یا آخر شامل انتخاب اپراتورهای ژنتیکی است. پارامترهای بهینه GEP مورد استفاده در هر اجرا در جدول ۲ گزارش شده است. مقادیر ارائه شده در این جدول، همان مقادیر پیش فرض برنامه GeneXpro می باشد. در این مطالعه، GepSoft جهت اجرای الگوریتم بیان ژن استفاده شده است. در مدل مورد نظر از سه پارامتر ارتفاع سطح آب، ارتفاع سطح نفت و نرخ تخلیه LNAPL استفاده شده است. پارامترهای مختلف GEP برای یافتن بهترین توپولوژی مورد سعی و خطا قرار گرفتند. مجموعه ۱۸۰ داده برای انجام پیش بینی نوسانات سطح LNAPL مورد استفاده قرار گرفت. کل مجموعه داده ها به دو بخش شامل ۱۵۰ داده برای آموزش مدل و ۳۰ داده برای تست مدل تقسیم شد.

$$\mu_{A_i} = \frac{1}{1 + [(WL_{t-1} - c_i/a_i)]^{2b_i}} \quad (6)$$

و یا برای تابع عضویت گوسین به صورت زیر می باشد.

$$\mu_{A_i} = \exp[-(x - c_i/a_i)^2] \quad (7)$$

که a_i ، b_i و c_i پارامترهای تابع عضویت هستند.

لایه ۲: به نام گره های Rule یا قانون می باشد. گره های این لایه ثابت و دایره ای شکل هستند. لایه دوم شامل اپراتورهای فازی است و از اپراتورهای AND برای فازی سازی ورودی ها استفاده می کند و هر گره خروجی نشان دهنده firing strength آن قانون می باشد. این گره ها با علامت π نام گذاری شده اند و نشان دهنده این است که سیگنال های ورودی در هم ضرب شده و به عنوان خروجی به لایه بعدی اعمال می شوند. خروجی این لایه به صورت زیر نشان داده می شود:

$$O_i^3 = w_i = \mu_{A_i}(wl_{t-1}) * \mu_{B_i}(q_{t-1}) \quad (8)$$

$i = 1, 2$

لایه سوم: گره های میانگین نام دارند. گره ها در این لایه ثابت و دایره ای شکل هستند و با علامت N نشان داده شده اند. خروجی این لایه به نام firing strength نرمال یا وزن نرمال شده می باشد که به صورت زیر نشان داده می شود:

$$O_i^3 = \bar{w}_i = \frac{w_i}{w_1 + w_2} \quad i = 1, 2 \quad (9)$$

لایه چهارم: این گره ها تطبیقی و مربعی شکل هستند و گره های نتیجه نام دارند. تابع گره برای لایه چهارم، سهم هر قانون i را نسبت به خروجی کل محاسبه می کند و به صورت زیر تعریف می شود.

$$O_i^4 = \bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i) \quad (10)$$

$i = 1, 2$

که \bar{w}_i خروجی لایه سوم و p_i ، q_i و r_i ضرایب ترکیب خطی بوده و همچنین مجموعه پارامتر در بخش نتیجه مدل فازی Sugeno می باشد. لایه پنجم: گره های خروجی نام دارد و دایره ای شکل می باشد. در این لایه فقط یک گره ثابت وجود دارد که با علامت Σ مشخص شده است. این گره مجموع همه سیگنال های ورودی را

مدل سازی نوسانات سطح LNAPL به کار رفته است. برای نمایش معماری ANFIS، سیستم استنتاج تطبیقی فازی با دو متغیر ورودی x و y فرض می شود که در اینجا x و y می تواند ارتفاع سطح آب و نرخ تخلیه LNAPL در زمان $t-1$ باشد و متغیر خروجی z نیز می تواند ارتفاع سطح LNAPL در زمان t باشد. پس برای یک مدل فازی Takagi-Sugeno از نوع مرتبه اول، با داشتن دو قانون if-then، می توان قوانین را به شکل زیر نوشت:

$$IF = X \text{ is } A_1 \text{ AND } Y \text{ is } B_1 \text{ THEN } f_1 \\ = p_1 X + q_1 Y + r_1 \quad (2)$$

$$IF = X \text{ is } A_2 \text{ AND } Y \text{ is } B_2 \text{ THEN } f_2 : \\ = p_2 X + q_2 Y + r_2 \quad (3)$$

که x و y متغیرهای ورودی هستند، A_i و B_i مجموعه های فازی هستند، f_i خروجی های درون منطقه فازی می باشد که توسط قوانین فازی مشخص شده است، p_i ، q_i و r_i پارامترهای خطی یا طراحی در بخش (consequent part) مدل می باشد که در طی مرحله آموزش محاسبه می شوند. معماری ANFIS استفاده شده برای اجرای این دو قانون در شکل ۴ نشان داده شده است. دایره ها در شکل نشان دهنده گره های ثابت هستند، در حالی که مربع ها نشان دهنده گره های تطبیقی می باشند. معماری ANFIS شامل پنج لایه است (شکل ۴)، هر لایه در این شکل با جزئیات زیر بیان می شود. لایه ۱: گره های ورودی نام دارند. گره ها در این لایه تطبیقی هستند. این لایه درجه عضویت ورودی ها را نشان می دهد و می تواند به صورت زیر بیان شود:

$$O_i^1 = \mu_{A_i}(wl_{t-1}) \quad i = 1, 2 \quad (4)$$

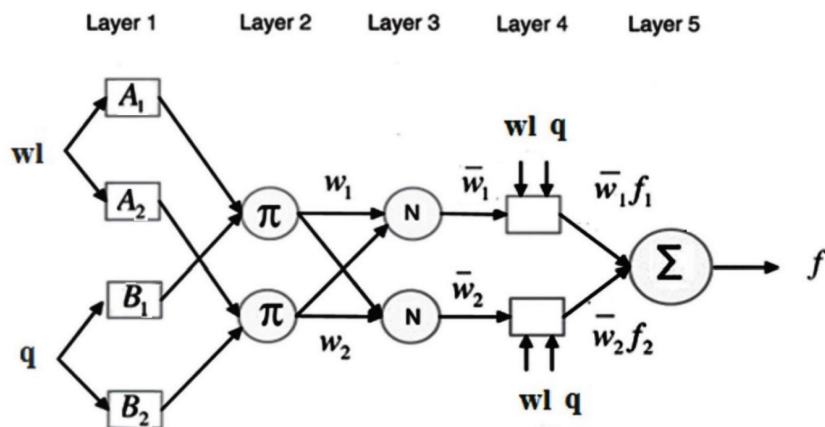
$$O_i^1 = \mu_{B_{i-2}}(q_{t-1}) \quad i = 3, 4 \quad (5)$$

که wl و q به ترتیب ارتفاع سطح آب و نرخ تخلیه LNAPL می باشند و A_i و B_i نیز کمیت های زبانی (بزرگ، کوچک و غیره) می باشند. $\mu_{A_i}(wl_{t-1})$ و $\mu_{B_{i-2}}(q_{t-1})$ توابع گره می باشند و به گونه ای انتخاب می شود که در محدوده بین ۰-۱ قرار گیرند و برای تابع عضویت "زنگی شکل" می تواند به صورت زیر نوشته شود.

در این مطالعه، برای توسعه مدل ANFIS جهت پیش بینی ارتفاع سطح LNAPL از شبکه تطبیقی و مدل Sugeno و برای ساخت مدل های ANFIS از برنامه متلب استفاده شده است.

محاسبه می کند و خروجی کلی مدل به صورت زیر می باشد (Al-Hmouz et al., 2012).

$$O_i^d = \sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (11)$$



شکل ۴. مکانیزم منطقی برای مدل Sugeno که اساس ANFIS می باشد

معادله رگرسیون خطی می باشد. اگر تنها یک متغیر مستقل وجود داشته باشد مدل را یک مدل رگرسیون خطی ساده می نامند و اگر بیش از یک متغیر مستقل وجود داشته باشد مدل را یک مدل رگرسیون خطی چند متغیره می نامند. در این مطالعه برای پیش بینی نوسانات سطح LNAPL از دو متغیر مستقل شامل ارتفاع سطح آب و نرخ تخلیه LNAPL و برای ساخت مدل رگرسیون از نرم افزار DataFit استفاده شد.

رگرسیون چند متغیره

آنالیز رگرسیون یکی از تکنیک های آماری برای مطالعه و مدل سازی ارتباط بین متغیرها می باشد. در ابتدا آنالیزور یک رابطه، بین متغیرها حدس می زند و سپس شروع به جمع آوری اطلاعات کمی می کند و آنها را روی یک نمودار دو بعدی رسم می کند، اگر داده ها در امتداد یک خط باشند بنابراین رابطه خطی است و معادله آنها به صورت زیر خواهد بود:

$$Y = ax + b \quad (12)$$

که a عرض از مبدا و b شیب خط می باشد. در اینجا اختلاف کمی بین مقادیر واقعی و مقادیر مدل وجود خواهد داشت که به نام خطای تخمین گفته می شود. این خطا می تواند برخاسته از خطای اندازه گیری، شرایط محیطی، اختلاف طبیعی و غیره باشد. بنابراین معادله اولیه خطا به صورت زیر خواهد بود.

$$y = ax + b + \epsilon \quad (13)$$

معادله بالا به نام معادله مدل رگرسیون خطی نامیده می شود. در اینجا x یک متغیر مستقل و y یک متغیر وابسته می باشد. در مدل رگرسیون، هدف پیش بینی رفتار متغیر وابسته با آگاهی از مقادیر متغیرهای مستقل با استفاده از

بحث

در این بخش، نتایج آموزش و تست مدل های ANFIS، MLP و GEP برای پیش بینی ارتفاع سطح LNAPL در محدوده مورد نظر نمایش داده و مورد بحث قرار می گیرند. مدل های مورد نظر با استفاده از پارامترهای آماری زیر مورد ارزیابی قرار می گیرند. در اینجا t مقدار اندازه گیری شده یا واقعی و 0 مقدار پیش بینی شده از مدل می باشد.

$$MAPE = \frac{1}{n} \left[\frac{\sum_{i=1}^n t_i - o_i}{\sum_{i=1}^n t_i} \times 100 \right] \quad (14)$$

1. MLP model and application
2. Linear Regression Model (MLR)

در نخستین مرحله مطالعه، مدل‌های به‌کاررفته توسط متغیرهای ورودی ارتفاع سطح آب و نرخ تخلیه LNAPL مورد ارزیابی و آزمایش قرار گرفت. یکی از بهترین ساختارهای به‌دست‌آمده برای هر مدل، برای انجام پیش‌بینی ارتفاع سطح LNAPL انتخاب شد. خطاهای آماری برای اجرای مدل‌ها در طی دوره‌های آموزش و آزمایش محاسبه شد (جدول ۳). همان‌طور که از جدول ۳ مشاهده می‌شود مدل GEP نتایج بهتر با خطاهای کمتر و مقادیر همبستگی به نسبت بالاتری را با بقیه مدل‌ها نشان می‌دهد. شکل ۵ نتایج اجرای مدل‌ها را بر روی داده‌های آموزش و آزمایش نشان می‌دهند. همچنین تصویر ۶ مقایسه‌ای از ارتفاع سطح LNAPL پیش‌بینی شده توسط هر سه مدل، برای دوره‌های آموزش را نشان می‌دهد. در این تصاویر به‌طور کامل مشهود است که نتایج مدل GEP خیلی نزدیک به مقادیر واقعی است. همچنین نمودار پراکندگی نقاط برای مقادیر مشاهده‌ای و واقعی روی نتایج اجرای هر سه مدل رسم و در شکل ۷ نشان داده شده است. ضرایب همبستگی به‌دست‌آمده برای نتایج مدل GEP در طی دوره‌های آموزش و آزمایش به ترتیب برابر با ۰/۹۶ و ۰/۸۶۲ می‌باشد که بالاتر از ضرایب همبستگی دو مدل دیگر است. تصاویر ۶ و ۷ نشان می‌دهند که برنامه‌نویسی بیان ژن GEP قادر به مدل‌سازی و پیش‌بینی پدیده‌های غیرخطی همچون تخمین ارتفاع سطح LNAPL و آب با دقت بالاتری نسبت به سایر مدل‌ها می‌باشد. یکی از مزایای اصلی مدل GEP در مقایسه با سایر مدل‌های هوش مصنوعی، ارائه بیان ریاضی برای فرایند مورد مطالعه است. شکل ۸ معماری بیان ژن مربوط به هر ژن در این مسئله را نشان می‌دهد و معادله کلی برای پیش‌بینی سطح LNAPL برابر با مجموع معادلات به‌دست‌آمده برای هر ژن می‌باشد. این معادله به‌صورت رابطه ۱۸ می‌باشد.

$$\text{LNAPL Level} = ۶.۳ + \left(\frac{-۸.۹}{(-۷.۶۳+Q)*WL} \right) + WL + \left(۲۴۴.۴ + \frac{\sqrt{wl}}{۰.۳۸۵} \right) + \left(-\frac{(۲۵۵)}{۶.۰۹} \right) - \left(\frac{-۴.۷۹}{-۶.۱۶-Q} \right) \quad (18)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (t_i - o_i)^2}{n}} \quad (15)$$

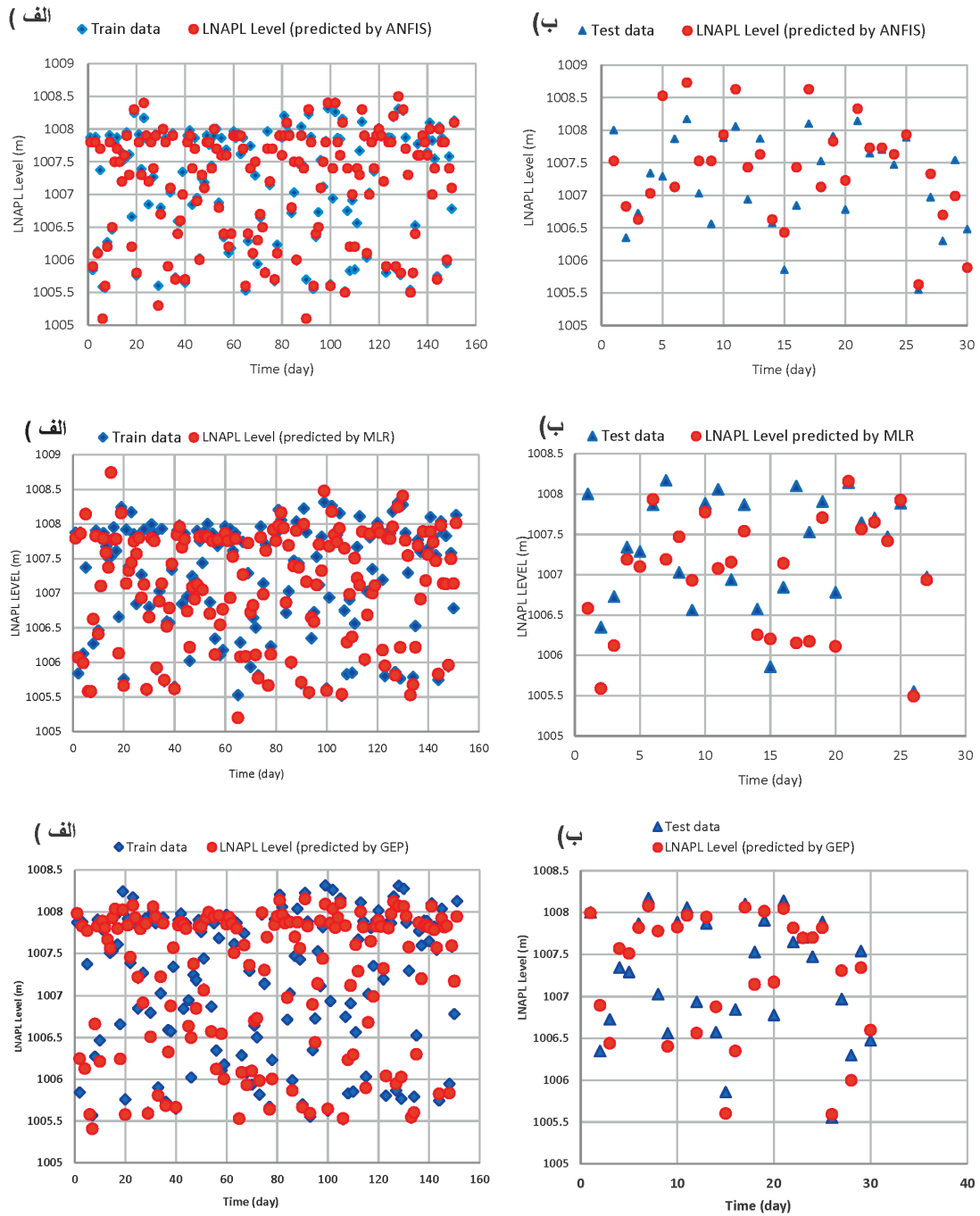
$$R = \text{Corr}(t, o) = \frac{\text{cov}(t_i, o_i)}{\sigma_t \sigma_o} \quad (16)$$

$$R^2 = \frac{(n \sum t_i o_i - \sum t_i \sum o_i)^2}{(n \sum t_i^2 - (\sum t_i)^2)(n \sum o_i^2 - (\sum o_i)^2)} \quad (17)$$

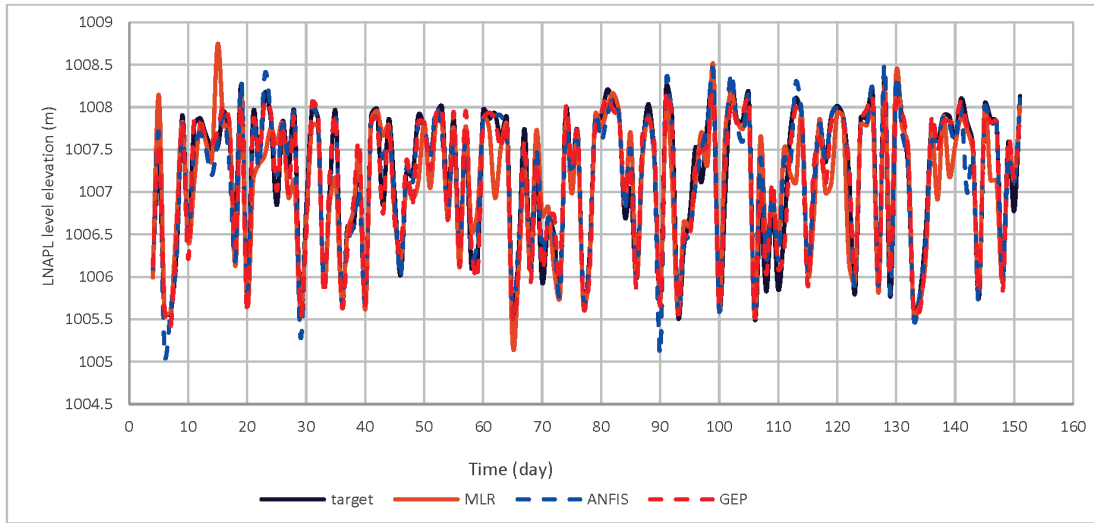
پارامترهای آماری برای نتایج مدل‌ها در جدول ۳ آورده شده است. MAPE یا درصد میانگین مطلق خطا، یک مقیاسی از دقت پیش‌بینی روش مورد استفاده است که همواره به‌صورت درصد بیان می‌شود. RMSE یا جذر میانگین مربع خطا، میانگین مقدار بزرگی خطاها را با اختصاص وزن بیشتر به خطاهای بزرگ‌تر نشان می‌دهد و مقدار آن بین صفر تا بینهایت می‌باشد. مقادیر کمتر RMSE نشان‌دهنده دقت بالای مدل می‌باشد. پارامتر R شامل یک ارتباط آماری بین دو متغیر می‌باشد و حدود آن بین صفر و یک است. R‌های بزرگ‌تر نشان‌دهنده دقت بالاتر مدل می‌باشد. پارامتر R² نشان‌دهنده این است که چند درصد از رفتار متغیر وابسته، می‌تواند توسط متغیرهای مستقل پیش‌بینی شود. حدود تغییرات R² نیز بین صفر و یک می‌باشد که مقادیر بزرگ‌تر R² نشان‌دهنده اجرا بهتر و دقت بالاتر مدل می‌باشد. این مطالعه به کاربرد مدل‌های مختلف مثل GEP، ANFIS و MLR در پیش‌بینی ارتفاع سطح LNAPL کمک می‌کند.

جدول ۳. پارامترهای آماری پیش‌بینی نوسانات سطح LNAPL در هر سه روش

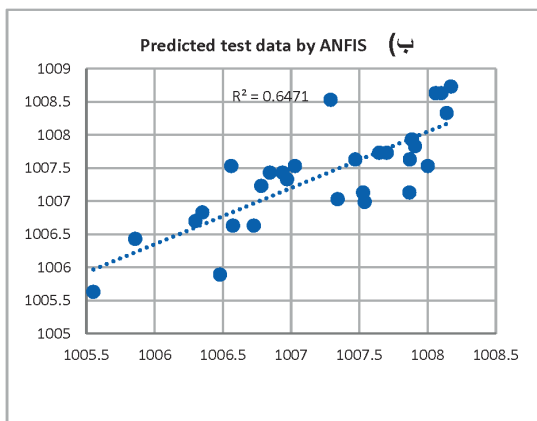
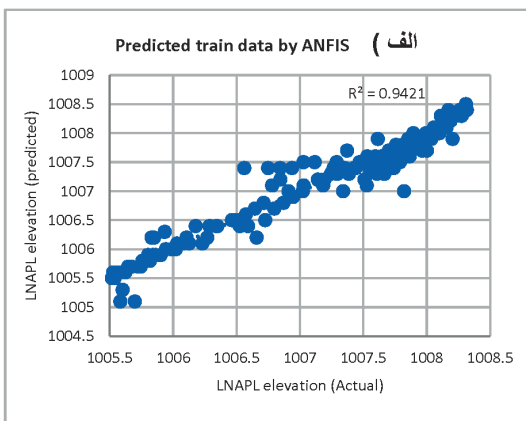
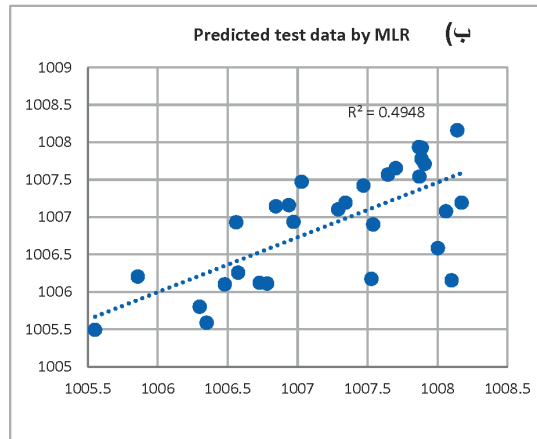
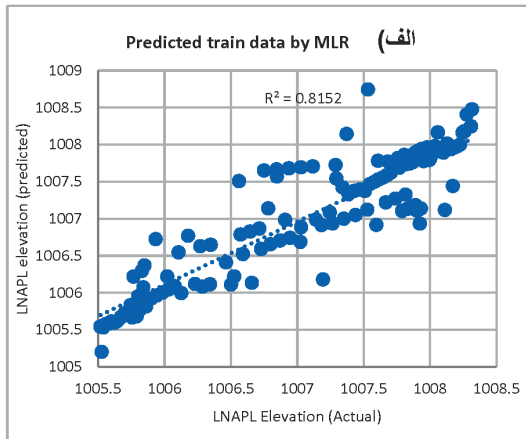
پارامترهای آماری	نوع داده	GEP	ANFIS	MLR
MAPE	داده‌های تست	۰/۰۰۰۰۶	-۰/۰۰۰۵۴	۰/۰۰۱۱
	داده‌های آموزش	-۰/۰۰۰۰۵۸	۰/۰۰۰۰۱	۰/۰۰۰۰۴۲
RMSE	داده‌های تست	۰/۰۹۹	۰/۸۹	۱/۸۳
	داده‌های آموزش	۰/۱۶۲۷	۰/۳۷۹	۰/۲۰۹
R	داده‌های تست	۰/۹۲۸	۰/۸۰۴	۰/۷۰۳
	داده‌های آموزش	۰/۹۸۲	۰/۹۷	۰/۹۰۲
R ²	داده‌های تست	۰/۸۶۲	۰/۶۵	۰/۵
	داده‌های آموزش	۰/۹۶۴	۰/۹۴۲	۰/۸۱۵



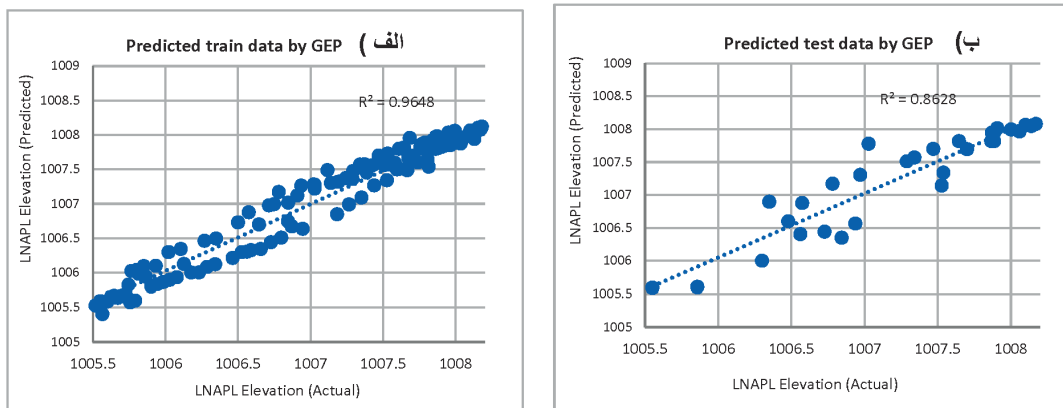
شکل ۵. نتایج پیش بینی ارتفاع سطح LNAPL توسط مدل های ANFIS، MLR و GEP برای داده های آموزش و (ب) داده های تست



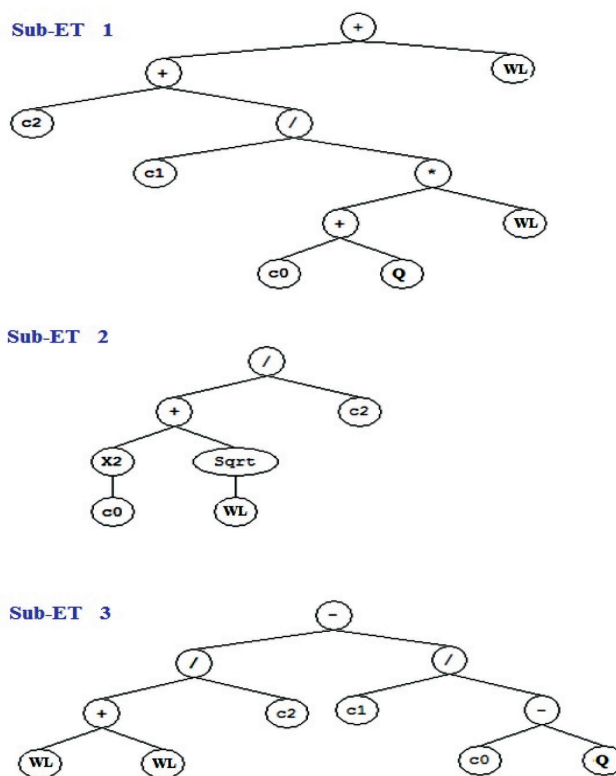
شکل ۶. مقادیر ارتفاع سطح LNAPL پیش‌بینی شده و مشاهده شده با استفاده از مدل‌های GEP، ANFIS و MLR بر روی داده‌های آموزش



شکل ۷. نمودار پراکندگی ارتفاع واقعی سطح LNAPL در مقابل ارتفاع پیش‌بینی شده، (الف) برای داده‌های آموزش و (ب) برای داده‌های تست، به ترتیب برای مدل‌های GEP، ANFIS و MLR



ادامه شکل ۷ .



شکل ۸. معماری بیان ژن و معادله به دست آمده برای هر ژن

نتیجه گیری

آب زیرزمینی به همراه نرخ برداشت از لایه LNAPL به عنوان متغیرهای تاثیرگذار بر ارتفاع LNAPL و به عنوان متغیرهای ورودی و ارتفاع سطح LNAPL نیز به عنوان متغیر خروجی در نظر گرفته شدند. ۱۸۰ داده برای مطالعه مورد نظر جمع آوری و آماده شد و به دو گروه شامل ۱۵۰

در این مطالعه توانایی GEP، ANFIS و MLR در پیش بینی ارتفاع سطح LNAPL مورد ارزیابی قرار گرفت. با رسم هیدروگراف های ۱۰ ساله و یک ساله مشخص شد که تغییرات ارتفاع سطح LNAPL به طور کامل از تغییرات سطح آب زیرزمینی پیروی می کند. بنابراین ارتفاع سطح

tice, *Studies in Fuzziness and Soft Computing*, 181, Germany: Springer Verlag, 53-83, doi: 10.1007/11339366, 3.

- Adamowski, J. and Chan, H. F., 2011. A wavelet neural network conjunction model for groundwater level forecasting. *Journal of Hydrology*, 407, 28-40.

- Al-Hmouz, A., Shen, J., Al-Hmouz, R. and Yan, J., 2012. Modeling and Simulation of an Adaptive Neuro-Fuzzy Inference System (ANFIS) for Mobile Learning. *IEEE Transactions on Learning Technologies* 5, 3, 226-237.

- Aytek, A. and Alp, M., 2008. An application of artificial intelligence for rainfall runoff modeling. *Journal of Earth System Science*. 117,2, 145-155.

- Azari, T. and Samani, N., 2018. Modeling the Neuman's well function by an artificial neural network for the determination of unconfined aquifer parameters. 22, 4, 1135-1148.

- Cimen, M. and Kisi, O., 2009. Comparison of two different data-driven techniques in modeling lake level fluctuations in Turkey. *Journal of Hydrology*, 378, 253-262.

- Coppola, E., Szidarovszky, F., Davis, D., Spayad, S., Poulton, M. and Roman, E., 2007. Multi objective analysis of a public wellfield using artificial neural networks. *Groundwater* 45,1, 53-61.

- Coppola, Jr., Emery, A., Rana, Anthony, J., Poulton, Mary. M., Szidarovszky, F. and Uhi, V. W., 2005. A neural network model for predicting aquifer water level elevations. *Groundwater*. 43, 2, 231-241.

- Danandeh Mehr, A., Kahya, E. and Yerdelen, C., 2014. Linear genetic programming application for successive-station monthly streamflow prediction. *Computers and Geosciences*, 70, 63-72.

- Elzwayie, A., El-shafie, A., Yaseen,

داده برای آموزش و ۳۰ داده برای آزمایش تقسیم شد. نتایج اجرای سه مدل توسط پارامترهای آماری جذر میانگین مربع خطا، درصد میانگین خطای مطلق و ضریب همبستگی و R^2 مورد تحلیل و بررسی قرار گرفت و مشخص شد که میزان پارامترهای آماری گفته شده برای مدل GEP بهتر از سایر مدل‌ها می‌باشد. همچنین روش GEP در شبیه‌سازی نقاط اکسترمم، دقت قابل توجهی را از خود نشان داد که کارآمد بودن این روش را در مسائل هیدرولیکی اثبات می‌کند درحالی‌که دقت روش ANFIS در شبیه‌سازی این نقاط کمتر از GEP بود که می‌تواند به دلیل محدود بودن قوانین فازی به کاررفته در این روش باشد زیرا احتمال وقوع نقاط اکسترمم در سطح آب و LNAPL کم است و این مقادیر جزء قوانین تاثیرگذار فازی در نظر گرفته نشده است و می‌تواند یکی از نقاط ضعف این روش باشد. لازم به ذکر است همواره نمی‌توان گفت که در تمامی موارد عملکرد GEP بهتر خواهد بود، چه بسا روی یک سری از داده‌های دیگر روش ANFIS عملکرد مناسب‌تری داشته باشد. همچنین در این مطالعه با کمک برنامه GEP یک معادله برای پیش‌بینی سطح LNAPL ارائه و توسط مقادیر واقعی آزمایش و مورد تحلیل قرار گرفت. مشخص شد نتایج دارای خطای پایینی است و می‌توان از این معادله در سرچاه برای پیش‌بینی ارتفاع سطح LNAPL استفاده کرد.

منابع

- ناصری، ح.، عسگری، ف.، خدایی، ک. و علیجانی، ف.، ۱۳۹۹. تأثیر آبیاری غرقایی و قطره‌ای هوشمند بر نوسانات تراز سطح آب زیرزمینی با استفاده از مدل فیزیکی. فصلنامه زمین‌شناسی ایران، ۵۳، ۱۴.

- میرعربی، ع.، ناصری، ح.، نخعی، م. و علیجانی، ف.، ۱۳۹۸. بررسی کارایی مدل هیبریدی هالت-وینترز موجکی (WHW) در شبیه‌سازی تراز سطح ایستابی آبخوان ساحلی ارومیه. فصلنامه زمین‌شناسی ایران، ۴۹، ۱۸.

- Abraham, A., 2005. Adaptation of Fuzzy Inference System Using Neural Learning, in Nedjah, Nadia; de Macedo Mourelle, Luiza, *Fuzzy Systems Engineering: Theory and Prac-*

- Zaher. M., Afan, H. A. and Falah Allawi, M., 2016. RBFNN-based model for heavy metal prediction for different climatic and pollution conditions. *Neural Computing and Applications*. 28, 8, 1991-2003.
- Emamgholizadeh, S., Moslemi, Kh., Karami, Gh.H., 2014. Prediction the Groundwater Level of Bastam Plain (Iran) by Artificial Neural Network (ANN) and Adaptive Neuro-Fuzzy Inference System (ANFIS) *Water Resources Management*. DOI 10.1007/s11269-014-0810-0.
 - Ferreira, C., 2001a. Gene expression programming in problem solving. In: *Sixth Online World Conference on Soft Computing in Industrial Applications (invited tutorial)*, Springer, London, 635-653, <https://doi.org/10.1007/978-1-4471-0123-9-54>.
 - Ferreira, C., 2001b. Gene expression programming: a new adaptive algorithm for solving problems. *Complex Systems* 13,2, 87-129.
 - Ferreira, C., 2006. *Gene Expression Programming: Mathematical Modeling by an Artificial Intelligence*. Springer, Berlin, Heidelberg New York, 478.
 - Ghorbani, M. A, Deo, Ravinesh, C., Karimi, V., Yaseen, Zaher. M. and Terzi, Ozlem., 2017. Implementation of a hybrid MLP-FFA model for water level prediction of Lake Egirdir, Turkey. *Stochastic Environmental Research and Risk Assessment* . 32, 6, 1683-1697.
 - Guldal, V. and Tongal, H., 2010. Comparison of Recurrent Neural Network, Adaptive Neuro-Fuzzy Inference System and Stochastic Models in E girdir Lake Level Forecasting. *Water Resources Management*, 24, 105-128.
 - Hawthorne, J.M., 2011. Diagnostic Gauge Plots, *Applied NAPL Science Review*, 1 <http://www.icontact-archive.com/IXYNsGudxSsIUd-6HuogSpblft2mtIAJM> .
 - Jang, J.S.R., 1991. Fuzzy Modeling Using Generalized Neural Networks and Kalman Filter Algorithm. *Proceedings of the 9th National Conference on Artificial Intelligence*, Anaheim, CA, USA, July, 14-19, 2, 762-767.
 - Jang, J.S.R., 1993. ANFIS: adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man and Cybernetics* 23,3, 665-685.
 - Karimi, S., Shiri, J., Kisi, O. and Makarynsky, O., 2012. Forecasting water level fluctuations of Urmieh Lake using Gene Expression Programming and Adaptive Neuro-Fuzzy Inference System. *International Journal of Ocean and Climate Systems* 3,109-125.
 - Kisi, O., Shiri, J. and Nikoofar, B., 2012. Forecasting daily lake levels using artificial intelligence approaches. *Computers & Geosciences* 41,169-180.
 - Mamdani, E. H. and Assilian, S., 1975. An experiment in linguistic synthesis with a fuzzy logic controller. *International Journal of Man-Machine Studies* 7,1, 1-13.
 - Mpallas, L., Tzimopoulos, C. and Evangelides, C., 2011. Comparison between neural networks and adaptive neuro-fuzzy inference system in modeling Lake Kerkini water level fluctuation lake management using Artificial Intelligence. *Journal of Environmental Science and Technology* 4, 366-376.
 - Nadim, F., Hoag George, E., Liu, Sh., Carley Robert, J. and Zack, P., 2000. Detection and remediation of soil and aquifer systems contaminated with petroleum products: an overview. *Journal of Petroleum Science and Engineering*. 26, 1-4, 169-178.
 - Nayak. Purna, C., Satyaji Rao, Y.R. and Sudheer, K. P., 2006. Groundwater Level Forecasting in a Shallow Aquifer Using Artificial Neural Network Approach. *Water Resources Management*. 20, 1,77-90.

- Nazari, A., 2012. Prediction performance of PEM fuel cells by gene expression programming. *International Journal Hydrogen Energy* 37, 18972-18980.
- Noori, R., Khakpour, A., Omidvar, B. and Farokhnia, A., 2010. Comparison of ANN and principal component analysis-multivariate linear regression models for predicting the river flow based on developed discrepancy ratio statistic. *Expert Systems with Applications* 37, 5856-5862.
- Ozbek, A., Unsal, M. and Dikec, A., 2013. Estimating uniaxial compressive strength of rocks using genetic expression programming. *Journal of Rock Mechanics Geotechnical Engineering*, 5, 325-329.
- Sanikhani, H., Deo, Ravinesh, C., Yaseen, Zaher .M., Eray, O. and Kisi, O., 2018. Non-tuned data intelligent model for soil temperature estimation: A new approach. *Geoderma*. 330, 52-64.
- Shiri, J., Kisi, O., Yoon, H., Lee, K. K. and Hossein Nazemi, A., 2013. Predicting groundwater level fluctuations with meteorological effect implications-A comparative study among soft computing techniques. *Computers and Geosciences*. 56, 32-44.
- Solomatine, D., See, L. and Abraham, R., 2009. *Data-Driven Modelling: Concepts, Approaches and Experiences*. Practical Hydroinformatics. Water Science and Technology Library, 68, 17-30 Springer, Berlin, Heidelberg, <https://doi.org/10.1007/978-3-540-79881-1-2>.
- Takagi, T. and Sugeno, M., 1985. Fuzzy identification of systems and its application to modeling and control. *IEEE Transactions on System, Man and Cybernetics* 15,1, 116-132.
- Tiri, A., Belkhir, L. and Mouni, L., 2018. Evaluation of surface water quality for drinking purposes using fuzzy inference system. *Groundwater for Sustainable Development*. 6, 235-244.
- Wang, W. C., Chau, K. W., Cheng, C. T. and Qiu, L., 2009. A comparison of Performance of several artificial intelligence methods for forecasting monthly discharge time series. *Journal of Hydrology*, 374, 294-306.
- Yaseen, Zaher .M., Ebtahaj, Isa., Bonakdari, H., Deo, R. C., Danandeh Mehr, A., Wan Mohtar, W. H. M., Diopf, L., El-shafie, A., Singhi, Vijay, P., 2017. Novel approach for stream-flow forecasting using a hybrid ANFIS-FFA model. 554, 263-276.
- Yaseen, Zaher. M., El-shafie, A., J, O., H.A. and Sayl, K.N., 2015. Artificial Intelligence based models for stream-flow forecasting. *Journal of Hydrology*, doi: <http://dx.doi.org/10.1016/j.jhydrol.2015.10.038>.
- Yaseen, Zaher. M., Kisi, O. and Demir, V., 2016b. Enhancing Long-Term Streamflow Forecasting and Predicting using Periodicity Data Component: Application of Artificial Intelligence. *Water Resources Management*. 30: 4125. <https://doi.org/10.1007/s11269-016-1408-5>.
- Yaseen, Zaher. Mundher., Falah A, M., Yousif, A. A., Jaafar, O., Mohamad Hamzah, F. and El-Shafie, A., 2016a. Non-tuned machine learning approach for Hydrological time series forecasting. *Neural Computing and Applications*. 30, 5, 1479-1491.
- Yoon, H., Jun, S., Hyun, Y., Bae, G. and Lee, K., 2011. A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. *Journal of Hydrology*. 396, 1-2, 128-138.
- Zaqoot, Hossam. A., Hamada, M. and Migdad, Sh., 2018. A Comparative Study of Ann for Predicting Nitrate Concentration in Groundwater Wells in the Southern Area of Gaza Strip. *Applied Artificial Intelligence*. 32, 7-8, 727-744.