

مقاله علمی-پژوهشی

مقایسه عملکرد مدل‌های LS-SVR، ANFIS، MLP و RBF در مدل‌سازی سطح آب زیرزمینی (مطالعه موردی بخش مرکزی دشت مشهد)

امه‌هانی موسوی خالدی^۱، علی فیروز زارع^{۲*}، آرش دوراندیش^۳، فرشته مدرسی^۴

تاریخ دریافت: ۱۴۰۴/۰۶/۲۰ تاریخ پذیرش: ۱۴۰۵/۰۲/۲۵

چکیده

افت سطح آب زیرزمینی در دشت مشهد، عمدتاً ناشی از عوامل انسانی، چالشی جدی در مدیریت پایدار منابع آب ایجاد کرده است. وابستگی شدید به این منابع، پیش‌بینی دقیق تغییرات را برای تصمیم‌گیری‌های مدیریتی ضروری می‌سازد. این پژوهش با بهره‌گیری از روش‌های یادگیری ماشین، تغییرات سطح آب در بخش مرکزی دشت را شبیه‌سازی و پیش‌بینی می‌کند. در این پژوهش ابتدا با استفاده از تابع خودهمبستگی جزئی (PACF) و رگرسیون لاسو فراوانی‌گرا (Frequentist Lasso Regression) متغیرهای ورودی مهم و وقفه‌های زمانی بهینه شناسایی شدند. سپس، عملکرد مدل‌های حداقل مربعات میانگین رگرسیون بردار پشتیبان (LS-SVR)، استنتاج عصبی - فازی تطبیقی (ANFIS)، شبکه عصبی پرسپترون چندلایه (MLP) و شبکه عصبی تابع پایه شعاعی (RBF) در پیش‌بینی ماهانه نوسانات سطح آب زیرزمینی طی ۳۰ سال (۱۳۷۰-۱۴۰۰) ارزیابی گردید. نتایج نشان داد که هر چهار مدل با دقت قابل قبول سطح آب زیرزمینی را شبیه‌سازی می‌کنند. مدل RBF با R^2 ، MSE، NSE و RMSE به ترتیب ۰/۱۰۰، ۰/۰۰۹ و ۰/۰۰۹ عملکرد برتری نسبت به سایر مدل‌ها دارد. یافته‌های این پژوهش پتانسیل بالای مدل‌های داده‌محور در شبیه‌سازی فرایندهای هیدرولوژیکی را تأیید کرد و می‌تواند مبنای توسعه ابزارهای هوشمند برای مدیریت منابع آب زیرزمینی در دشت مشهد و مناطق مشابه، از جمله تخصیص منابع و طراحی سامانه‌های هشداردهنده، باشد.

واژه‌های کلیدی: روش‌های یادگیری نظارت‌شده، مدل‌های یادگیری ماشین، مدل‌های داده‌محور، نوسانات سطح آب زیرزمینی

مقدمه

برای آب و اثرات تغییر اقلیم، فشار بر این منابع را تشدید کرده و مدیریت آن‌ها را پیچیده‌تر ساخته است (Kayhomayoon et al., 2021; Famiglietti, 2014). این چالش جهانی در بسیاری از دشت‌های ایران به شکلی حاد تجلی یافته است که دشت مشهد یکی از برجسته‌ترین و بحرانی‌ترین نمونه‌های آن به شمار می‌رود. این دشت استراتژیک که در وضعیت ممنوعه بحرانی قرار دارد، تحت تأثیر برداشت بیش‌ازحد، کاهش بارندگی، توسعه شهری و آبیاری ناکارآمد، با شرایط ناپایدار مواجه شده و گذار به مدیریت پایدار را به ضرورتی اجتناب‌ناپذیر بدل ساخته است (حسینی و همکاران، ۱۴۰۳; Saberifar, 2023).

تحقق چنین مدیریتی نیازمند شناخت دقیق رفتار آبخوان و پیش‌بینی تغییرات آینده سطح آب زیرزمینی در سناریوهای مختلف مدیریتی و اقلیمی است. در این راستا، مدل‌سازی و بهره‌گیری از ابزارهای هوشمند می‌تواند به عنوان یکی از ارکان اصلی سیستم‌های پشتیبان تصمیم نقش‌آفرینی کند (دوقانی و همکاران، ۱۳۹۹؛ قلی‌زاده و همکاران، ۱۴۰۲).

بهره‌برداری از منابع آب زیرزمینی در مناطق خشک و نیمه‌خشک، به دلیل عدم توازن میان نرخ برداشت و تغذیه طبیعی آبخوان‌ها، به یک چالش جهانی تبدیل شده است. این اضافه برداشت، پیامدهای متعددی نظیر افت سطح آب، شور شدن منابع آبی و خاکی، افزایش آلودگی و پدیده فرونشست زمین را در بسیاری از دشت‌ها به دنبال داشته است. علاوه بر این، عواملی چون رشد جمعیت، افزایش تقاضا

۱- دانشجوی دکتری اقتصاد کشاورزی، دانشکده کشاورزی، دانشگاه فردوسی

مشهد، مشهد، ایران

۲- دانشیار اقتصاد کشاورزی، دانشکده کشاورزی، دانشگاه فردوسی مشهد، مشهد،

ایران

۳- دانشیار گروه اقتصاد کشاورزی، دانشکده کشاورزی، دانشگاه تهران، مشهد، ایران

۴- استادیار گروه علوم و مهندسی آب، دانشکده کشاورزی، دانشگاه فردوسی مشهد،

مشهد، ایران

(*)- نویسنده مسئول: E-mail: firooz@um.ac.ir

DOI: [10.22034/ijdi.2026.546388.2628](https://doi.org/10.22034/ijdi.2026.546388.2628)

پایده‌سازی ساختارهای متنوع یادگیری ماشین، موجب ارتقای دقت و پایداری پیش‌بینی سطح آب زیرزمینی شود.

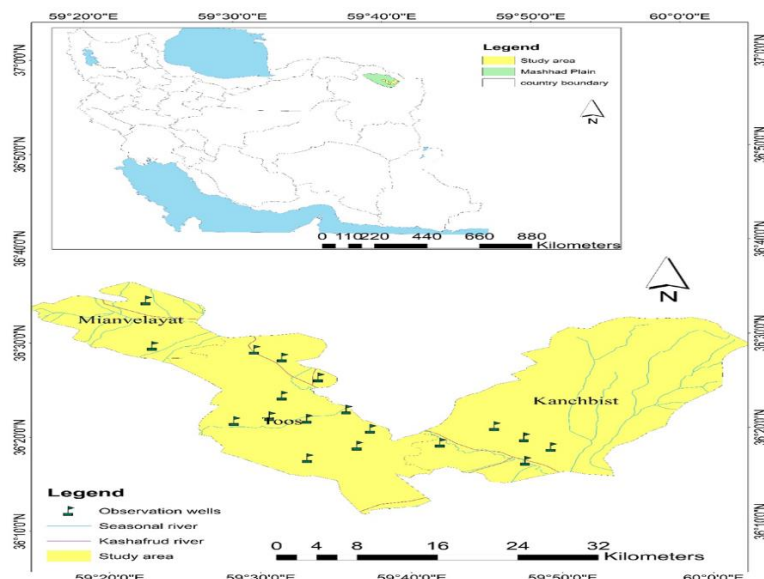
این مطالعه به بررسی توانایی چهار مدل یادگیری ماشین شامل LS-SVR، MLP، ANFIS و RBF در شبیه‌سازی و پیش‌بینی تغییرات سطح آب زیرزمینی در بخش مرکزی دشت مشهد می‌پردازد، منطقه‌ای که به دلیل افت شدید سطح آبخوان و فشارهای ناشی از برداشت بیش‌ازحد، توسعه شهری و تغییرات اقلیمی، با چالش‌های جدی زیست‌محیطی مواجه است. انتخاب این مدل‌ها بر پایه قابلیت برتر آن‌ها در شبیه‌سازی روابط غیرخطی پیچیده هیدرولوژیکی، کاهش حساسیت به نویز و عملکرد موفق در مطالعات مشابه است (احمدی و همکاران، ۱۴۰۰؛ Pourmorad et al., 2021; Ahmadi et al., 2024). به‌منظور افزایش دقت پیش‌بینی‌ها، از ترکیب روش‌های PACF و FLR برای انتخاب بهینه متغیرهای ورودی استفاده شده که به شناسایی الگوهای زمانی و تغییرات پیچیده سطح آب کمک کرده است. در این رویکرد، PACF برای استخراج تأخیرهای زمانی مؤثر و FLR برای حذف متغیرهای کم‌اثر و تعدیل اثر همخطی به کار گرفته شده‌اند؛ این ترکیب، ضمن کاهش پیچیدگی مدل و جلوگیری از بیش‌برازش، دقت و پایداری پیش‌بینی‌ها را به‌طور محسوسی افزایش می‌دهد. چارچوب ارائه شده، رویکردی کارآمد و تفسیرپذیر برای مدل‌سازی سری‌های زمانی هیدرولوژیکی فراهم می‌آورد. این پژوهش، ابزاری کلیدی برای مدیریت پایدار منابع آب در شرایط کم‌آبی و تغییرات اقلیمی ارائه می‌دهد. با حل چالش‌های انتخاب بهینه متغیرهای ورودی و مدل‌های مناسب در شبیه‌سازی فرایندهای هیدرولوژیکی پیچیده، به تدوین سیاست‌ها و استراتژی‌های مؤثر برای مقابله با افت سطح آب زیرزمینی، و همچنین تصمیم‌گیری‌های مدیریتی در تخصیص منابع و طراحی سامانه‌های هشداردهنده کمک می‌کند. این پژوهش به‌عنوان یک الگوی روش‌شناختی، می‌تواند راهگشای پژوهش‌های آتی در سایر مناطق خشک و نیمه‌خشک با چالش‌های مشابه باشد.

مواد و روش‌ها

منطقه مورد مطالعه

منطقه مورد مطالعه به وسعت ۱۲۲۹ کیلومتر مربع در بخش مرکزی دشت مشهد، بین طول جغرافیایی ۲۰°-۵۹' تا ۶۰°-۵' شرقی و عرض ۱۵°-۳۶' تا ۳۵°-۳۶' شمالی قرار دارد. این محدوده بر اساس آمار ۱۳۹۸، دارای منابع آبی شامل ۵۱۶ حلقه چاه، ۱۶ چشمه و ۳۲ رشته قنات است که موقعیت آن در شکل (۱) نشان داده شده است.

مدل‌های آب زیرزمینی برای درک و پیش‌بینی رفتار آبخوان‌ها حیاتی هستند. روش‌های سنتی پرهزینه و زمان‌بر هستند، درحالی‌که رویکردهای داده‌محور مانند یادگیری ماشین، با نیاز کمتر به داده‌های فیزیکی و قابلیت شناسایی الگوهای پنهان، جایگزینی کارآمد ارائه می‌دهند (احمدی و همکاران، ۱۴۰۰؛ Aderemi et al., 2023; Shivaprakash et al., 2022). در مطالعات هیدرولوژیکی، مدل‌های یادگیری ماشین به‌طور گسترده‌ای برای پیش‌بینی نوسانات سطح آب زیرزمینی استفاده شده‌اند. مدل‌های کاربرد شامل شبکه‌های عصبی مصنوعی (ANN)، رگرسیون خطی چندگانه (MLR)، مدل خودرگرسیون (AR)، ماشین بردار پشتیبان (SVM) و سیستم استنتاج عصبی-فازی تطبیقی (ANFIS) هستند (Nadiri et al., 2019; Roozbahani et al., 2018; Saha et al., 2022). کارایی این مدل‌ها را در شبیه‌سازی سطح آب زیرزمینی تأیید کرده‌اند. افخمی و صراف مدل‌های ELM و ANN را با تبدیل موجک و بهینه‌سازی ازدحام ذرات کوانتومی (QPSO) ترکیب کرده و مدل هیبرید W-ELM-QPSO را معرفی کردند که بالاترین دقت پیش‌بینی را ارائه داده است (Afkhamifar & Sarraf, 2020). چید پودی و همکاران در مطالعه‌ای از مدل‌های یادگیری عمیق برای شبیه‌سازی تغییرات سطح آب زیرزمینی استفاده کردند. آن‌ها سه مدل حافظه طولانی کوتاه‌مدت (LSTM)، واحد بازگشتی دروازه‌دار (GRU) و LSTM دوطرفه (BLSTM) را ارزیابی کردند. نتایج نشان داد که GRU در اکثر موارد عملکرد نسبتاً بهتری داشت (Chidepudi et al., 2022). گونزالس و ارسنجان در مطالعه‌ای از سه مدل یادگیری ماشین، SVM، ANN و جنگل تصادفی (RF)، برای پیش‌بینی تغییرات سطح آب زیرزمینی در دانمارک استفاده کردند. آن‌ها دریافتند که مدل RF از نظر دقت از دو مدل دیگر مناسب‌تر است (Gonzalez & Arsanjani, 2021). بین و همکاران در مطالعه‌ای از سه مدل RF، شبکه عصبی مصنوعی و LSTM برای پیش‌بینی تغییرات سطح آب زیرزمینی در ویکتوریا استرالیا استفاده کردند. آن‌ها دریافتند که دقت مدل LSTM به‌طور قابل‌توجهی بالاتر از دو مدل دیگر است (Yin et al., 2021). باوجود قابلیت‌های قابل توجه مدل‌های یادگیری ماشین، انتخاب مدل بهینه همچنان یک چالش اساسی محسوب می‌شود، زیرا عواملی مانند دقت، کارایی محاسباتی، حجم داده‌ها، پارامترهای ورودی و تفسیرپذیری مدل به‌طور مستقیم بر عملکرد آن‌ها تأثیر می‌گذارند (Kenda et al., 2018; Shortridge et al., 2016). در این زمینه، انتخاب نظام‌مند متغیرهای ورودی و تعیین ساختار بهینه مدل از مسائل کلیدی محسوب می‌شود؛ با این حال، مطالعات محدودی به توسعه‌ی چارچوبی جامع و تفسیرپذیر پرداخته‌اند که علاوه بر انتخاب ورودی‌های مؤثر و



شکل ۱- منطقه مورد مطالعه بخش مرکزی دشت مشهد

داده‌های استفاده‌شده

باتوجه به تحقیقات گذشته، متغیرهای سطح آب زیرزمینی (GL)، بارندگی (PR)، دما (TE) و تبخیر (EV) به صورت ماهانه به عنوان ورودی انتخاب شدند (محمدی و همکاران، ۱۴۰۱؛ Sahoo et al., 2017).

داده‌های هواشناسی از ایستگاه سینوپتیک مشهد (سازمان

هواشناسی) و سطح آب زیرزمینی از شرکت آب منطقه‌ای خراسان رضوی، طی ۱۳۷۰-۱۴۰۰ جمع‌آوری گردید. ۱۸ چاه پیزومتریک بر اساس دوره آماری مشترک انتخاب و نواقص داده‌ها با درون‌یابی چندگانه در SPSS تکمیل شد. خصوصیات متغیرها در جدول (۱) ارائه شده است.

جدول ۱- خصوصیات آماری متغیرهای ورودی به مدل‌ها

متغیر	کمترین	میانگین	بیشترین	انحراف معیار	چولگی	کشدگی
سطح آب زیرزمینی (متر)	۵۰/۰۷	۶۱/۲۱	۱۸۷/۶۲	۸/۱۰	۹/۹۴	۱۵۶/۹۹
بارش (میلی‌متر)	۰/۰	۲۰/۰۲	۱۲۹/۶۹	۲۴/۲۹	۱/۶۶	۲/۹۵
دما	-۸/۱۳	۱۵/۹۳	۳۰/۵۵	۸/۷۵	-۰/۰۷	-۱/۳۱
تبخیر	۲۲/۳۶	۱۹۴/۶۲	۳۸۵/۰۷	۱۱۴/۷۹	۰/۲۲	-۱/۴۸

روش تعیین بهترین متغیرهای ورودی به مدل

انتخاب متغیرهای ورودی مناسب در مدل‌های یادگیری ماشین اهمیت زیادی دارد و به دلیل نبود روش استاندارد، از رویکردهای متنوع استفاده می‌شود (Ahmed et al., 2019; Sapitang et al., 2021). سطح آب زیرزمینی دارای الگوهای فصلی و خودهمبستگی است، لذا شناسایی تأخیرهای بهینه برای پیش‌بینی ضروری است (Tang et al., 2019). در این بخش، با تمرکز بر تأخیرهای زمانی، ابتدا PACF برای شناسایی تأخیرهای معنادار به کار گرفته شد، سپس با رگرسیون لاسو فراوانی‌گرا بررسی و فیلتر می‌شوند (Mouatadid et al., 2019; Roy et al., 2021).

استفاده از داده‌های خام برای ورودی به مدل‌های یادگیری ماشین باعث کاهش دقت و کارایی آن‌ها می‌شود؛ لذا داده‌ها قبل ورود به مدل توسط رابطه (۱) ابتدا استانداردسازی شده‌اند و سپس اعداد استانداردسازی شده در خروجی مدل‌ها نیز به حالت قبل برگردانده شده‌اند.

$$Z = \frac{X - \mu}{\sigma} \quad (1)$$

که در آن، Z مقدار متغیر استانداردشده، X مقدار اصلی، μ میانگین مقادیر واقعی و σ انحراف استاندارد مقادیر واقعی را نشان می‌دهد (Kellaway, 1969).

تابع خودهمبستگی جزئی

تابع خودهمبستگی جزئی (Partial Autocorrelation Function) PACF (Function) بزاری کلیدی در تحلیل سری‌های زمانی است که رابطه مستقیم بین یک متغیر و تأخیرهای آن را، با کنترل اثرات تأخیرهای میانی، اندازه‌گیری می‌کند. برخلاف تابع خودهمبستگی (ACF) که هم همبستگی‌های مستقیم و هم غیرمستقیم را ضبط می‌کند، PACF اثرات دیگر تأخیرها را حذف می‌کند که آن را به ابزاری بسیار مفید برای شناسایی پیش‌بینی‌کننده‌های مرتبط در پیش‌بینی سری‌های زمانی تبدیل می‌کند. از نظر ریاضی، PACF در تأخیر k که با ϕ_k نشان داده می‌شود، با استفاده از معادلات یول-واکر در یک مدل خودرگرسیون (AR) از مرتبه k تخمین زده می‌شود:

$$\phi_k = \frac{r_k - \sum_{j=1}^{k-1} \phi_{k-j} r_j}{1 - \sum_{j=1}^{k-1} \phi_{k-j} r_j} \quad (2)$$

که در آن r_k نشان‌دهنده خودهمبستگی در تأخیر k است و $\phi_{k-1,j}$ خودهمبستگی‌های جزئی از مرحله قبلی هستند (Box et al., 2015).

رگرسیون لاسو فراوانی‌گرا

رگرسیون لاسو یک روش آماری برای انجام رگرسیون خطی همراه با تنظیم‌سازی و انتخاب متغیرها است. تنظیم‌سازی به جلوگیری از بیش‌برازش مدل کمک می‌کند، زیرا مقدار ضرایب رگرسیون را کاهش می‌دهد. در رگرسیون لاسو، ضرایب متغیرهای غیرمهم یا زائد دقیقاً برابر با صفر تنظیم می‌شوند که باعث می‌شود تنها متغیرهای مهم در مدل باقی بمانند. برخلاف رگرسیون خطی معمولی، هدف در رگرسیون لاسو کمینه‌سازی تابع زیر است:

$$\sum_{i=1}^N (y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j)^2 + \lambda \sum_j |\beta_j| \quad (3)$$

که در آن y_i نشان‌دهنده مقادیر مشاهده‌شده متغیر وابسته است، x_{ij} ماتریس ویژگی‌های ورودی، β بردار ضرایب رگرسیون و λ پارامتر تنظیمی است که میزان کاهش ضرایب را کنترل می‌کند (Friedman, 2009; Tibshirani, 1996). رگرسیون لاسو فراوانی‌گرا یکی از انواع رگرسیون لاسو است که به عنوان مرحله غربالگری، برای انتخاب مهم‌ترین متغیرها پیش از مدل‌سازی روابط غیرخطی مورد استفاده قرار می‌گیرد. این روش در شناسایی متغیرهای تأثیرگذار و کاهش پیچیدگی مدل کاربرد دارد (Coleman et al., 2016). برای پیاده‌سازی این مدل از نرم‌افزار MATLAB استفاده شده است.

سیستم استنتاج تطبیقی عصبی فازی

سیستم استنتاج عصبی-فازی تطبیقی ANFIS (Adaptive Neuro-Fuzzy Inference System) ترکیبی از دو رویکرد شبکه عصبی و منطق فازی است. در این مدل، شبکه عصبی پارامترهای قوانین فازی را بهینه می‌کند تا بهترین روابط بین متغیرهای ورودی و خروجی را تعیین کند. این رویکرد از مزایای هر دو رویکرد شبکه‌های عصبی و منطق فازی بهره می‌برد (Prado et al., 2020). روش استنتاج فازی شامل سه نوع سیستم پرکاربرد ممدانی، سوگنو و سوکاموتو است. این پژوهش از سیستم سوگنو بهره برده است که در مدل ANFIS دارای پنج‌لایه است که از خصوصیات ورودی برای تولید یک مقدار خروجی استفاده می‌کند. این کار را با تبدیل خصوصیات ورودی به مجموعه‌ای از شرایط، شرایط را به مجموعه‌ای از مقادیر خروجی و مقادیر خروجی را به یک عدد واحد انجام می‌دهد (سلامتیان و همکاران، ۱۴۰۲). در این پژوهش از دو روش الگوریتم خوشه‌بندی میانگین‌های c فازی (FCM) و خوشه‌بندی تفریقی (Subtractive Clustering) استفاده شده است.

حداقل مربعات رگرسیون بردار پشتیبان

مدل حداقل مربعات رگرسیون بردار پشتیبان (Least Squares Support Vector Regression) نوعی ماشین بردار پشتیبان است. تابع LS-SVR با مجموعه داده‌های آموزشی $\{x_k, y_k\}_{k=1}^N$ که داده ورودی آن شامل $x_k \in R^N$ و داده‌های خروجی $y_k \in R^N$ به صورت رابطه (۴) می‌توان بیان کرد:

$$y(x) = \omega^T \phi(x) + b \quad (4)$$

که در آن ω بردار وزن‌ها، $\phi(\cdot)$ یک تابع غیرخطی است که الگوهای ورودی را به فضای ویژگی با ابعاد بالاتر تبدیل می‌کند و b نشان‌دهنده بایاس تابع رگرسیون است (Moravej et al., 2020). این رگرسیون را می‌توان با استفاده از رابطه بهینه‌سازی (۵) حل کرد.

$$\min J(\omega, e) = \frac{1}{2} \omega^T \omega + \frac{\gamma}{2} \sum_{i=1}^N e_i^2 \quad (5)$$

با در نظر گرفتن محدودیت‌ها معادله (۶) حاصل می‌شود:

$$y_i = \omega^T \phi(x_i) + b + e_i \quad (i = 1, 2, \dots, N) \quad (6)$$

با حل معادله فوق به روش ضرایب لاگرانژ، معادله زیر حاصل می‌شود که $k(x, x_i)$ تابع کرنل است. در مدل حداقل مربعات میانگین رگرسیون بردار پشتیبان، معمولاً از سه تابع کرنل گوسی، چندجمله‌ای و خطی استفاده می‌شود. در این پژوهش از تابع کرنل گوسی و چندجمله‌ای برای حل تابع تقریب رگرسیون غیرخطی استفاده شده است.

کرل گوسی:

$$k(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right) \quad (7)$$

کرل چندجمله‌ای:

$$K(x_i, x_j) = (x_i^T x_j + c)^d \quad (8)$$

شبکه عصبی مصنوعی پرسپترون چندلایه

یکی از پرکاربردترین شبکه‌های عصبی مصنوعی، شبکه‌های عصبی MLP (Multilayer Perceptron) است. MLP ها معمولاً به گونه‌ای طراحی می‌شوند که داده‌ها از یک لایه ورودی به لایه خروجی منتقل شوند. بین هر لایه ورودی و خروجی، یک یا چند لایه مخفی وجود دارد. هر لایه مخفی از واحدهای پردازشی تشکیل شده است که از یک لایه به لایه دیگر به طور کامل به یکدیگر متصل هستند. این اتصال‌ها با وزن‌هایی مشخص می‌شوند. خروجی هر واحد پردازشی با استفاده از یک تابع فعال‌سازی محاسبه می‌شود برای اطلاعات بیشتر، به مقاله (Modaresi et al., 2018) رجوع شود.

شبکه عصبی توابع پایه شعاعی

شبکه‌های عصبی مبتنی بر تابع پایه شعاعی (RBF Radial basis function) از یک لایه پنهان تشکیل شده‌اند که ورودی‌های غیرخطی را به الگوهای جداگانه تبدیل می‌کند. لایه خروجی سپس ترکیبی خطی از این الگوها را برای تولید پیش‌بینی‌های لازم ایجاد می‌کند. تعداد واحدهای پنهان و انتخاب تابع فعال‌سازی در لایه پنهان بر توانایی مدل در انجام پیش‌بینی‌ها تأثیر می‌گذارد. با استفاده از این روش مدل‌سازی، می‌توان روابط غیرخطی بین متغیرها را به درستی شناسایی کرد. هدف لایه پنهان تبدیل ویژگی‌های ورودی غیرخطی به الگوهای قابل تفکیک است. لایه خروجی سپس ترکیبی خطی از این الگوها را برای تولید پیش‌بینی‌ها ایجاد می‌کند. عواملی مانند تعداد واحدهای پنهان و انتخاب تابع فعال‌سازی بر قابلیت پیش‌بینی مدل تأثیر می‌گذارند (Banadkooki et al., 2020).

شاخص‌های ارزیابی خطا

به منظور ارزیابی اثربخشی الگوها از معیارهای مختلف عملکرد شامل ضریب تعیین (R^2)، میانگین مربعات خطا (MSE)، ضریب کارایی نش ساتکلیف (NSE) و جذر میانگین مربعات خطا (RMSE) مورد استفاده قرار گرفت. R^2 توان مدل در تبیین واریانس داده‌ها را نشان می‌دهد. درحالی‌که RMSE و MSE شدت خطا و حساسیت مدل نسبت به خطاهای بزرگ را ارزیابی می‌کنند. همچنین، شاخص NSE که به طور گسترده در مطالعات هیدرولوژیکی به کار می‌رود، عملکرد مدل را نسبت به پیش‌بینی مبتنی بر میانگین مشاهدات مورد

سنجش قرار می‌دهد.

استفاده هم‌زمان از این شاخص‌ها امکان ارزیابی چندبعدی عملکرد مدل‌ها را فراهم می‌کند.

$$R^2 = \left[\frac{\sum_{t=1}^N (x_0 - \bar{x}_0)(x_p - \bar{x}_p)}{\sqrt{\sum_{t=1}^N (x_0 - \bar{x}_0)^2 \sum_{t=1}^N (x_p - \bar{x}_p)^2}} \right]^2 \quad (9)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (x_p - x_0)^2 \quad (10)$$

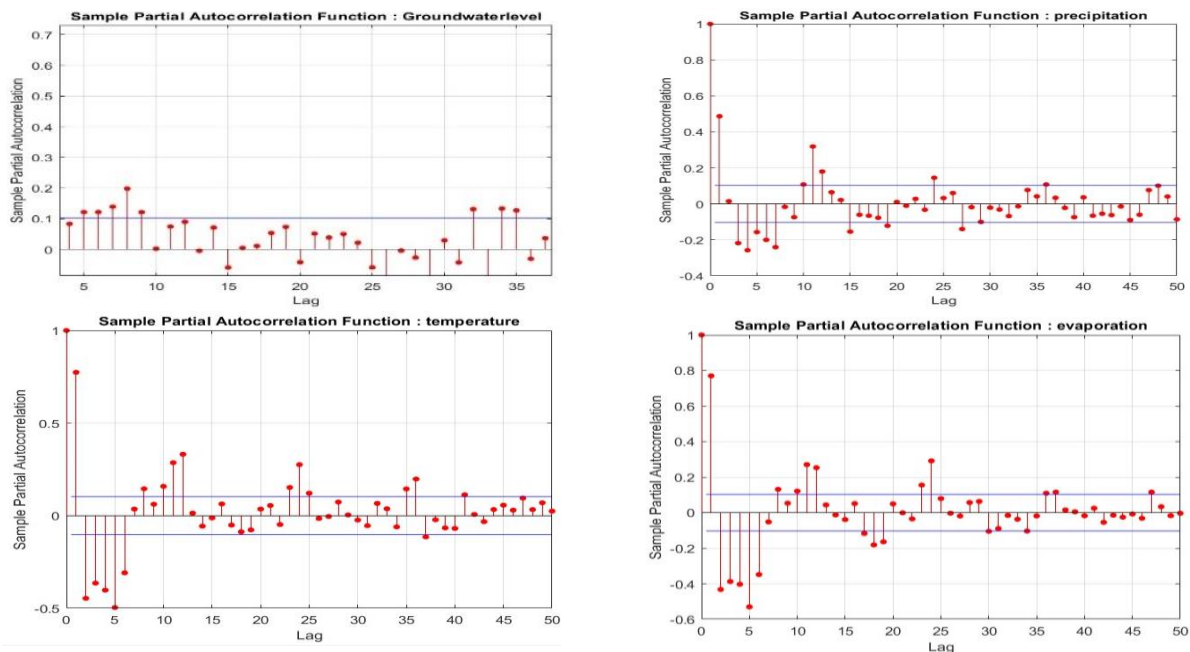
$$NSE = 1 - \left[\frac{\sum_{i=1}^n (x_0 - x_p)^2}{\sum_{i=1}^n (x_0 - \bar{x}_0)^2} \right] \quad (11)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_p - x_0)^2}{n}} \quad (12)$$

که در آن x_0 و x_p به ترتیب به مقادیر داده‌های مشاهداتی و شبیه‌سازی شده هستند. ضریب تعیین و ضریب نش همواره در بازه (۰ تا ۱) قرار دارند. هرچه مقدار ضریب تعیین و ضریب نش به یک نزدیک‌تر باشد، مدل بهتر می‌تواند تغییرات متغیر وابسته را توضیح دهد. همچنین، هرچه مقدار RMSE و MSE به صفر نزدیک‌تر باشد، مدل دقیق‌تر است (Mobayen et al., 2023).

نتایج و بحث

در این پژوهش از چهار مدل مختلف ANFIS، LS-SVR، MLP و RBF برای شبیه‌سازی سطح آب زیرزمینی در بخش مرکزی دشت مشهد استفاده شد. میانگین وزنی سطح آب زیرزمینی ۱۸ چاه پیژومتری موجود در محدوده مورد مطالعه با استفاده از روش تیسن بندی استخراج شد و سپس مورد شبیه‌سازی قرار گرفت. برای ارزیابی تأثیر وقفه‌های زمانی بر سطح آب زیرزمینی، PACF تا ۵۰ وقفه محاسبه شد. حدود اطمینان ۹۵٪ ترسیم گردید و تنها وقفه‌هایی که مقدار PACF آن‌ها از این نوار فراتر رفت، به عنوان وقفه‌های معنادار انتخاب شدند. انتخاب سقف ۵۰ وقفه باهدف پوشش حافظه‌ی چند فصلی و چندساله‌ی آبخوان صورت گرفت. نتایج این تحلیل در شکل (۲) نشان داده شده است. لازم به ذکر است که ۱۸ چاه پیژومتری استفاده شده در این پژوهش، تمامی چاه‌های دارای داده معتبر و پیوسته در محدوده مطالعاتی را شامل می‌شود و از نظر توزیع مکانی، نماینده‌ی مناسبی از آبخوان مرکزی به شمار می‌آیند.



شکل ۲- نمودارهای خودهمبستگی جزئی سری‌های زمانی سطح آب زیرزمینی، بارش، دما و تبخیر با ۵۰ تأخیر (ماهانه).

بر اساس نتایج PACF تعداد وقفه‌های مناسب برای پارامترهای ورودی برای پیش‌بینی سطح آب زیرزمینی به صورت زیر است:

$$GL_{t-1}, GL_{t-2}, GL_{t-5}, GL_{t-6}, GL_{t-7}, GL_{t-8}, GL_{t-9}, GL_{t-26}, GL_{t-29}, GL_{t-32}, GL_{t-34}, GL_{t-35}, GL_{t-40}, \\ GL_{t-43}, GL_{t-45}, GL_{t-46}, GL_{t-48}, PR_t, PR_{t-1}, PR_{t-3}, PR_{t-4}, PR_{t-5}, PR_{t-6}, PR_{t-7}, PR_{t-11}, PR_{t-12}, \\ PR_{t-15}, PR_{t-19}, PR_{t-24}, PR_{t-37}, TE_t, TE_{t-1}, TE_{t-1}, TE_{t-2}, TE_{t-3}, TE_{t-4}, TE_{t-5}, TE_{t-6}, TE_{t-8}, TE_{t-10}, \\ TE_{t-11}, TE_{t-12}, TE_{t-23}, TE_{t-24}, TE_{t-25}, TE_{t-35}, TE_{t-36}, TE_{t-37}, TE_{t-41}, EV_{t-1}, EV_{t-2}, EV_{t-3}, EV_{t-4}, \\ EV_{t-5}, EV_{t-6}, EV_{t-8}, EV_{t-10}, EV_{t-11}, EV_{t-12}, EV_{t-17}, EV_{t-18}, EV_{t-19}, EV_{t-24}, EV_{t-36}, EV_{t-37}, \text{and } EV_{t-47}$$

گردیدند و دقت با میانگین مربعات خطاهای پیش‌بینی (FMSE) محاسبه شد. جدول (۲) ترکیب بهینه وقفه‌های زمانی را نشان می‌دهد؛ سطح آب زیرزمینی وابستگی قوی به نه گام گذشته دارد که نشان از رفتار خودرگرسیو و اینرسی‌ای آبخوان در شرایط نیمه‌خشک دارد. بارش و دما تنها در زمان t مؤثرند، حاکی از تأثیر فوری بارش بر تغذیه و دما بر تبخیر-تعرق. تبخیر تا سه گام زمانی تأثیر دارد که بیانگر نقش تجمعی آن در کاهش ذخایر آبی است. یافته‌ها با نتایج مطالعات پیشین هم‌سو است و تأیید می‌کند که تأخیرهای چندماهه در متغیرهای تأثیرگذار، نقشی کلیدی در پیش‌بینی سطح آب زیرزمینی با استفاده از مدل‌های یادگیری ماشین دارند (جعفری و همکاران، ۱۳۹۹؛ Solgi et al., 2019; Zare & Koch, 2018).

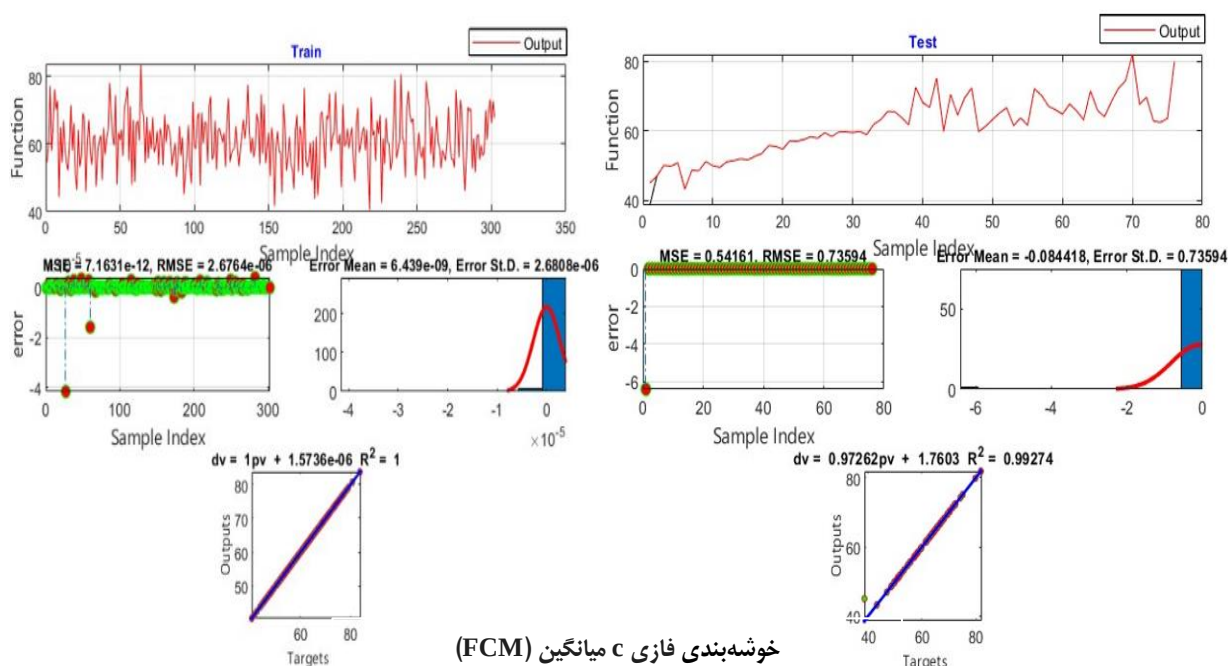
تحلیل PACF دینامیک پیچیده سیستم آب زیرزمینی دشت مشهد را آشکار می‌سازد. سطح آب دارای حافظه دوگانه است: اینرسی قوی کوتاه‌مدت (واکنش آنی) و حافظه بلندمدت ناشی از چرخه‌های اقلیمی. واکنش آبخوان به بارش با تأخیر معنادار (ناشی از نفوذ) همراه است، درحالی‌که دما و تبخیر به طور غیرمستقیم از طریق کنترل تقاضای کشاورزی و پمپاژ نقش تعیین‌کننده دارند. این پیچیدگی، ناکارآمدی مدل‌های خطی را اثبات و ضرورت استفاده از مدل‌های یادگیری ماشین پیشرفته را توجیه می‌کند و این تحلیل را اساس مهندسی ویژگی برای توسعه مدل پیش‌بینی قابل اتکا قرار می‌دهد. در مرحله بعد، از FLR به منظور حذف متغیرهای کم‌اهمیت استفاده شد. این تکنیک با منظم‌سازی ضرایب رگرسیون، از برآزش بیش از حد جلوگیری می‌کند. داده‌ها به ۸۰٪ آموزش و ۲۰٪ صحت‌سنجی تقسیم

جدول ۲- وقفه‌های زمانی متغیرهای منتخب

متغیر	مراحل زمانی منتخب
سطح آب زیرزمینی	$GL_{t-1}, GL_{t-2}, GL_{t-6}, GL_{t-9}$
بارش	PR_t
دما	TE_t
تبخیر	$EV_t, EV_{t-1}, EV_{t-2}, EV_{t-3}$

جدول ۳- مقادیر شاخص‌های آماری محاسبه‌شده برای مدل ANFIS

مدل	الگوریتم خوشه‌بندی	تقسیم‌بندی داده‌ها	ریشه میانگین مربعات خطا	میانگین مربعات خطا	ضریب تعیین	ضریب کارایی نش‌سا تکلیف
ANFIS	خوشه‌بندی تفریقی	آموزش	۰/۰۰۰	۰/۰۰۰	۱/۰۰	۱/۰۰
		آزمون	۰/۸۵۲	۰/۹۹	۰/۹۹	۰/۷۰
	خوشه‌بندی میانگین‌های c فازی	آموزش	۰/۰۰۰	۰/۰۰۰	۱/۰۰	۱/۰۰
		آزمون	۰/۷۴۰	۰/۷۴۰	۰/۹۹۲	۰/۹۹۰



شکل ۳- نمودار داده‌های شبیه‌سازی شده با مدل ANFIS و پارامترهای الگو

نتایج مدل ANFIS

مدل ANFIS بر پایه‌ی سیستم استنتاج فازی نوع سوگنو توسعه یافت. در این مدل از توابع عضویت گوسی با ساختار (۴،۳،۳) برای ورودی‌ها و خروجی‌های خطی استفاده شد و آموزش مدل با الگوریتم

یادگیری ترکیبی (Hybrid: LSE + BP) انجام گرفت. تنظیم پارامترها با جست‌وجوی شبکه‌ای^۱ و اعتبارسنجی نگه‌داشته^۲ انجام

- ۱ . Grid Search
- ۲ . Hold-out Validation

تابع هسته‌ای مورد بررسی قرار گرفت. پارامترهای تنظیمی با استفاده از جست‌وجوی شبکه‌ای تعیین شدند. در پیکربندی نخست، از هسته چندجمله‌ای با درجه ۳، $\text{KernelScale}=3$ و $\text{BoxConstraint}=1000$ استفاده شد. در پیکربندی دوم، هسته گوسی با مقادیر $\text{KernelScale}=30$ و $\text{BoxConstraint}=300$ به کار گرفته شد که بر اساس کمینه‌سازی RMSE در داده‌های اعتبارسنجی انتخاب شدند. این تنظیمات باهدف مقایسه‌ی توان مدل در شناسایی روابط غیرخطی و بهبود دقت پیش‌بینی سطح آب زیرزمینی انتخاب شدند. عملکرد هر کدام با شاخص‌های آماری در جدول (۴) و شکل (۴) بررسی شده است. بر اساس جدول (۴) مدل LS-SVR با گوسی با مقادیر ضریب تعیین 0.998 نسبت به مدل LS-SVR با کرنل چندجمله‌ای کارایی بهتری را داشته است و تقریباً از دقت یکسانی در تخمین رفتار سطح آب زیرزمینی برخوردار دارد. مدل LS-SVR با کرنل چندجمله‌ای نسبت به مدل دیگر عملکرد ضعیف‌تری داشته است. یافته‌ها با نتایج تانگ و همکاران (Tang et al., 2019) و احمدی و همکاران (۱۴۰۰) همخوانی دارد.

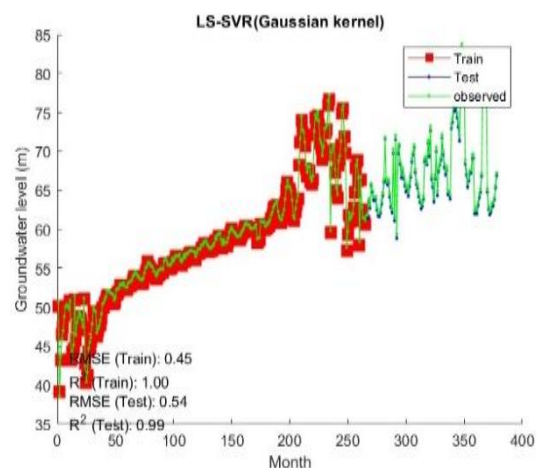
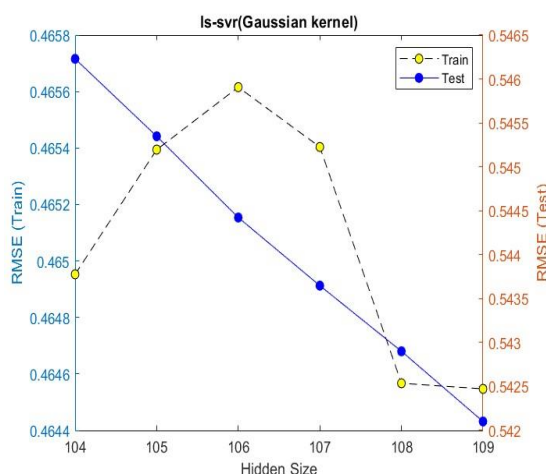
شد. در این فرایند، دو ساختار توابع عضویت گوسی $(3,3,3)$ و $(4,3,3,3)$ همراه با دو روش خوشه‌بندی تفریقی و c فازی مقایسه گردیدند. انتخاب پیکربندی بهینه بر اساس کمینه‌سازی RMSE و بیشینه‌سازی NSE انجام گرفت. همچنین، برای جلوگیری از بیش‌برازش، توقف زودهنگام طبق روش (Prechelt ۲۰۰۲) اعمال شد. نتایج مدل‌سازی ANFIS با داده‌های مشاهداتی در شکل (۳) و شاخص‌های آماری در جدول (۳) نشان می‌دهد که ANFIS در مرحله آموزش عملکرد ایده‌آلی دارد ($R^2=1.00$, $\text{RMSE}=0$)، اما در آزمون، خوشه‌بندی c فازی ($\text{NSE}=0.990$, $\text{RMSE}=0.740$) به مراتب پایدارتر از روش تفریقی ($\text{NSE}=0.70$, $\text{RMSE}=0.852$) است. این تفاوت، برتری ساختار انعطاف‌پذیر c فازی در تعمیم‌پذیری و مقاومت در برابر نویز داده‌های هیدرولوژیکی را نشان می‌دهد؛ بنابراین، انتخاب الگوریتم خوشه‌بندی نقش کلیدی در عملکرد نهایی مدل دارد.

نتایج مدل LS-SVR

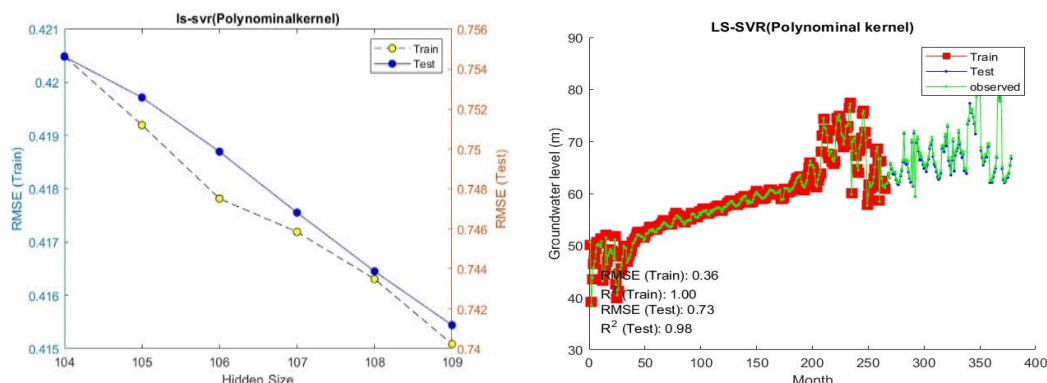
مدل LS-SVR به منظور ارزیابی کارایی ساختارهای مختلف، با دو

جدول ۴- مقادیر شاخص‌های آماری محاسبه‌شده برای مدل LS-SVR

مدل	نوع تابع کرنل	تقسیم‌بندی داده‌ها	ریشه میانگین مربعات خطا	میانگین مربعات خطا	ضریب تعیین	ضریب کارایی نش ساتکلیف
LS-SVR	کرنل گوسی	آموزش	۰/۴۵۵	۰/۲۰۷	۰/۹۹۶	۰/۹۹۶
		آزمون	۰/۵۳۷	۰/۲۸۸	۰/۹۸۸	۰/۹۸۸
	کرنل چندجمله‌ای	آموزش	۰/۳۶۴	۰/۱۳۲	۰/۹۹۸	۰/۹۹۸
		آزمون	۰/۷۳۱	۰/۵۳۵	۰/۹۷۸	۰/۹۷۸



الف) مدل حداقل مربعات میانگین رگرسیون بردار پشتیبان (کرنل گوسی)



ب) مدل حداقل مربعات میانگین رگرسیون بردار پشتیبان (کرنل چندجمله‌ای)

شکل ۴- نمودار داده‌های شبیه‌سازی شده با مدل LS-SVR بر اساس توابع کرنل (الف) گوسی و (ب) چندجمله‌ای

RMSE کمتر از ۵٪ بود. شاخص‌های آماری مدل MLP در جدول (۵) و شبیه‌سازی آن برای کل دوره در شکل (۵) ارائه شده است. بر اساس شاخص‌های آماری و در مقایسه با مطالعات پیشین ساهو و همکاران و سونی و همکاران (Sahu et al., 2020; Soni et al., 2024)، مدل در مرحله آزمون توانسته است رفتار سطح آب زیرزمینی را به خوبی برآورد کند.

نتایج مدل MLP

مدل MLP با استفاده از الگوریتم پس انتشار لونیگ-مارکوارت آموزش داده شد. در لایه پنهان از تابع فعال‌سازی tansig و در لایه خروجی از تابع purelin استفاده شد. تعداد نورون‌های لایه پنهان مدل MLP با جستجوی شبکه‌ای در بازه ۱۰ تا ۸۰ بهینه‌سازی شد و ۴۰ نورون به عنوان مقدار بهینه تعیین گردید. این نتیجه با اعتبارسنجی متقابل ده تایی تأیید شد، به طوری که انحراف معیار

جدول ۵- مقادیر شاخص‌های آماری محاسبه شده برای مدل MLP

ضریب کارایی نش	ضریب تعیین	میانگین مربعات خطا	ریشه میانگین مربعات خطا	تقسیم‌بندی داده‌ها	مدل
۰/۹۹۲	۰/۹۹۲	۰/۴۴۵	۰/۶۷۷	آموزش	MLP
۱/۰۰	۱/۰۰	۰/۰۰۰	۰/۰۱۳	آزمون	

لذا بر سایر مدل‌ها جهت پیش‌بینی سطح آب زیرزمینی برتری دارد. این نتایج در راستای مطالعات پیشین است (Di Nunno et al., 2023; Yan & Ma, 2016; Zhang et al., 2021).

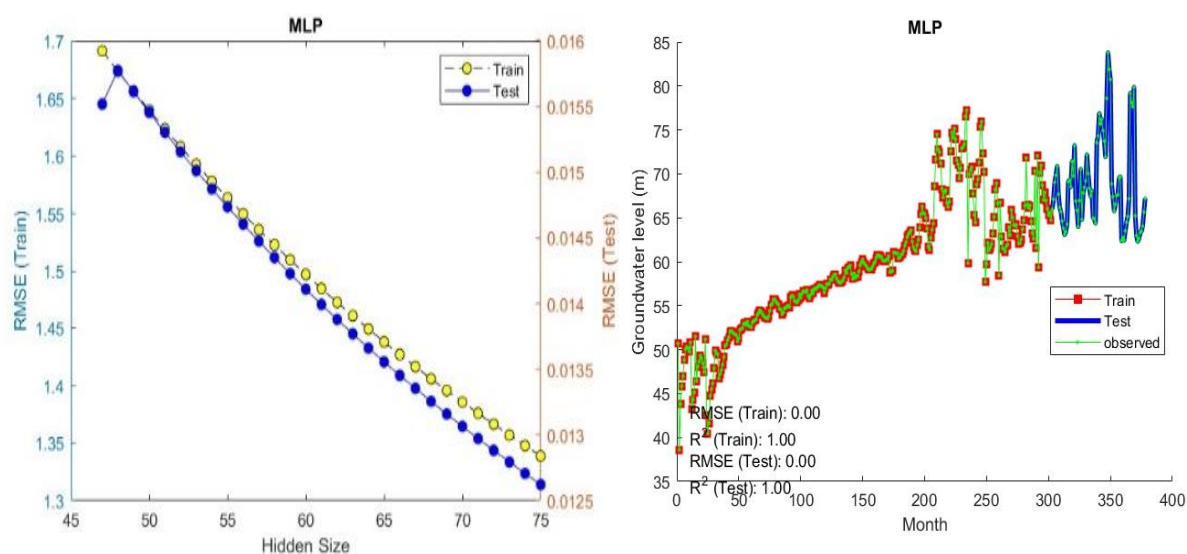
شکل (۶) مقادیر شبیه‌سازی شده و واقعی بخش آموزش و آزمون را برای مدل RBF نشان می‌دهد. مشخص است که اختلاف بین مقادیر شبیه‌سازی شده و مشاهداتی بسیار ناچیز است که این نکته توانایی بالای مدل در شبیه‌سازی سطح آب زیرزمینی را نشان می‌دهد.

نتایج مدل RBF

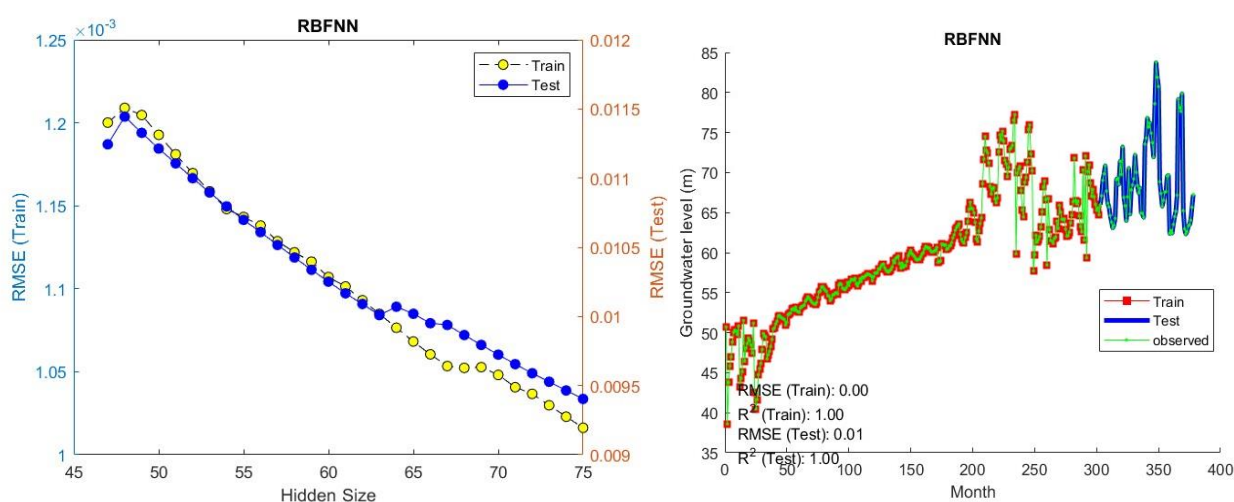
در مدل RBF مقادیر پارامتر spread در بازه {۱۰، ۱۵، ۲۰، ۲۵} و تعداد نورون‌ها در بازه {۳۰، ۴۰، ۶۰} به روش جستجوی شبکه‌ای بررسی گردید. پارامترهای تنظیمی شامل مقدار Spread = ۲۰، Goal = و حداکثر تعداد نورون‌ها برابر ۴۰ بودند. مقادیر مربوط به شاخص‌های آماری حاصل از مدل RBF در جدول (۶) گزارش شده است. مدل RBF برخلاف سه مدل دیگر بر اساس شاخص‌های آماری میانگین مربعات خطا و ضریب تعیین در مرحله آموزش و آزمون توانسته است به خوبی رفتار سطح آب زیرزمینی را برآورد کند.

جدول ۶- مقادیر شاخص‌های آماری محاسبه شده برای مدل RBF

ضریب کارایی نش	ضریب تعیین	میانگین مربعات خطا	ریشه میانگین مربعات خطا	تقسیم‌بندی داده‌ها	مدل
۱/۰۰	۱/۰۰	۰/۰۰۰	۰/۰۰۱	آموزش	RBF
۱/۰۰	۱/۰۰	۰/۰۰۰	۰/۰۰۹	آزمون	



شکل ۵- نمودار داده‌های شبیه‌سازی شده با مدل MLP بر اساس تعداد نورون‌های مختلف در لایه میانی



شکل ۶- نمودار داده‌های شبیه‌سازی شده با مدل RBF و پارامترهای الگو

شاخص‌های ارزیابی در مجموعه آموزشی و توقف فرایند آموزش بر مبنای عملکرد مجموعه اعتبارسنجی، نشان می‌دهد که بهبود این شاخص‌ها صرفاً حاصل تداوم یادگیری بر روی داده‌های آموزشی نبوده، بلکه ناشی از استخراج ساختارهای زمانی غالب و پایدار موجود در داده‌ها است. با این حال، از آنجاکه رفتار شاخص‌هایی نظیر R^2 و RMSE به مقیاس زمانی تحلیل و دامنه تغییرات متغیر هدف حساس است، تفسیر این مقادیر و به‌کارگیری نتایج حاصل از آن‌ها باید در چارچوب ویژگی‌های زمانی داده‌ها و شرایط هیدروژئولوژیکی سامانه

مدل‌ها به سبب معماری منحصربه‌فرد و نقاط قوت متمایز، رفتارهای عملکردی گوناگونی از خود بروز می‌دهند. هرچند در برخی از این مدل‌ها ضریب تعیین (R^2) در مجموعه آموزشی به عدد ۱ رسید، اما با اعمال مکانیسم توقف زودهنگام مطابق Prechelt (2002)، آموزش زمانی متوقف می‌گردد که بهبود خطای RMSE در مجموعه اعتبارسنجی طی ۵ تا ۱۰ دوره‌ی متوالی کمتر از 10^{-4} باشد. این آستانه از یادگیری نوین جلوگیری کرده و توان تعمیم‌پذیری مدل‌ها را تضمین می‌کند. هم‌زمانی دستیابی به مقادیر بسیار بالا برای

استفاده از داده‌های فضایی، تغییرات اقلیمی، و داده‌های چند منبع به همراه مقایسه مدل‌های مختلف یادگیری ماشین و روش‌های زمین‌آمار می‌تواند پیاده‌سازی مدل دقیق‌تر و جامع‌تر را تسهیل کند. بدون اعتبارسنجی محلی مکرر و به‌روزرسانی منظم مدل‌ها، تعمیم‌پذیری نتایج محدود خواهد بود؛ لذا ضروری است این موارد به طور مستمر در طرح‌های مدیریتی لحاظ شوند.

منابع

احمدی، زهرا؛ صفوی، ح... و زکری، م. (۱۴۰۰). پیش‌بینی سطح آب‌های زیرزمینی تحت سناریوهای تغییر اقلیم و مدیریت منابع آب. *مجله آب و فاضلاب*، ۳۱(۶)، ۳۴-۴۷.

دوقانی، ا.، دوراندیش، ا.، و قربانی، م. (۱۳۹۹). رتبه‌بندی شاخص‌های تاب‌آوری دشت مشهد در برابر کاهش منابع آب زیرزمینی با استفاده از روش بیزی بهترین-بدترین (BWM). *مدیریت آب و آبیاری*، ۱۰(۲)، ۳۰۱-۳۱۶.

قلی‌زاده، وحید؛ صفاری، امیر؛ احمدآبادی، علی؛ و کرم، اکبر. (۱۴۰۲). ارزیابی آسیب‌پذیری کیفی آبخوان دشت مشهد با استفاده از ترکیب مدل‌های SINTACS و GODS. *پژوهش‌های ژئومورفولوژی کمی*، ۱۲(۳)، ۵۹-۸۰.

حسینی، س. ز.؛ شاهدهی، ا.؛ و حنفی، ح. (۱۴۰۳). بررسی تغییرات زمانی و مکانی پارامترهای کمی منابع آب زیرزمینی با استفاده از روش میان‌بایی وزن تفکیک معکوس (مطالعه موردی: دشت مشهد). *مجله آب و قنات*، ۴(۲)، ۱۰۳-۱۱۴.

جعفری، م.؛ اوجاقلو، ح.؛ و زارع، م. (۱۳۹۹). شبیه‌سازی نوسانات سطح آب زیرزمینی با استفاده از ماشین بردار پشتیبان و سیستم استنتاج عصبی-فازی تطبیقی (مطالعه موردی: دشت مراغه). *نشریه آبیاری و زهکشی ایران*، ۱۴(۳)، ۹۴۲-۹۵۶.

محمدی، ف.؛ محمدی، م.؛ بهمنش، ج.؛ و جلیلی، س. (۱۴۰۱). بررسی نوسانات سطح آب‌های زیرزمینی با استفاده از تکنیک‌های محاسبات نرم (مطالعه موردی: دشت‌های ارومیه و میاندوآب). *آبیاری و مهندسی آب*، ۱۲(۴۷)، ۲۰۰-۲۱۹.

سلامتیان، س. ا.؛ ابراری، ح.؛ و نظری، ا. (۱۴۰۲). پیش‌بینی سطح آب‌های زیرزمینی با استفاده از سیستم استنتاج عصبی-فازی تطبیقی (مطالعه موردی: دشت قم). *آبیاری و مهندسی آب*، ۱۳(۵۱)، ۲۸۵-۳۰۴.

مورد مطالعه انجام پذیرد. در این میان مدل RBF در پیش‌بینی سطح آب زیرزمینی عملکرد برتری از خود نشان داد؛ این برتری ناشی از ساختار فنی خاص آن است. توابع فعال‌سازی گاوسی با خاصیت محلی‌سازی امکان شناسایی دقیق الگوهای غیرخطی را در فضای چندبعدی فراهم می‌کنند. علاوه بر این، آموزش هیبریدی موجب همگرایی سریع‌تر و پایداری عددی بیشتر نسبت به شبکه‌های چندلایه می‌شود. درنهایت، تنظیم پارامتر σ (spread) نقش مؤثری در کنترل تعمیم‌پذیری و جلوگیری از بیش‌برازش دارد.

نتیجه‌گیری

این پژوهش باهدف پاسخ به چالش اساسی مدیریت منابع آب در دشت‌های خشک از جمله دشت مشهد (ارتقای دقت پیش‌بینی نوسانات سطح آب زیرزمینی در مواجهه با پیچیدگی‌های هیدروژئولوژیک) چارچوبی ترکیبی ارائه کرد که ابتدا با استفاده از FLR و PACF، متغیرهای مؤثر و تأخیرهای کلیدی زمانی به‌صورت ساخت‌یافته انتخاب شدند و سپس عملکرد چهار مدل یادگیری ماشین (LS-SVR, ANFIS, MLP و RBF) مقایسه گردید؛ یافته‌ها نشان دادند که گزینش سیستماتیک تأخیرها می‌تواند موجب بهبود معنادار دقت پیش‌بینی می‌شود و مدل RBF با کمترین خطای پیش‌بینی و پایداری تعمیم‌پذیر نسبت به داده‌های آزمون عملکرد برتر را ارائه داد، این نتایج می‌تواند مؤید این باشد که ادغام دانش ساختاری (در انتخاب ویژگی‌های زمانی) با مدل‌های غیرخطی داده‌محور، رویکردی اثربخش برای کاهش عدم قطعیت در پیش‌بینی سطح آب زیرزمینی است.

توصیه می‌شود چارچوب مدل‌سازی موجود با استفاده از مدل‌های هیبریدی فیزیکی-داده‌محور توسعه یابد تا ضمن حفظ دانش مکانیزمی جریان‌های زیرزمینی، قابلیت بهره‌گیری از پردازش داده‌های حجیم و غیرخطی مدل‌های یادگیری ماشین نیز تأمین شود. احتمالاً این رویکرد می‌تواند دقت پیش‌بینی و تعمیم‌پذیری مدل را به‌ویژه در دشت‌های خشک بهبود بخشد. باتوجه به اهمیت عدم قطعیت در داده‌ها و مدل‌ها، افزودن بخش تحلیل عدم قطعیت به چارچوب پیشنهادی و اجرای اعتبارسنجی در مناطق مشابه با ویژگی‌های هیدروژئولوژیک مشابه، می‌تواند باعث افزایش اعتماد به مدل‌ها و نتایج شود. پیاده‌سازی سازوکارهای بازآموزی پویا با استفاده از داده‌های به‌روز در طول زمان، می‌تواند موجب حفظ و افزایش دقت مدل‌ها و قابلیت کاربرد آن‌ها در شرایط متغیر محیطی خواهد شد. برای بهره‌برداری عملی از مدل‌های پیش‌بینی، پیشنهاد می‌شود خروجی مدل‌ها در سیستم‌های پشتیبان تصمیم‌گیری تطبیقی به کار گرفته شود تا پایش مستمر، تنظیم برداشت و برنامه‌ریزی مدیریت منابع آبی به‌صورت پیشگیرانه و منطبق بر شرایط واقعی انجام گیرد.

Aderemi, B. A., Olwal, T. O., Ndambuki, J. M., & Rwanga, S. S. 2023. Groundwater levels forecasting using machine learning models: a case study of the groundwater region 10 at Karst Belt, South Africa. *Systems and Soft Computing*, 5, 200049.

Afkhamifar, S., & Sarraf, A. 2020. Prediction of

- Mathematics. By Erwin Kreyszig. Pp. xx, 899. 68s.(Wiley.). The Mathematical Gazette, 53(386), 444-444 .
- Kenda, K., Čerin, M., Bogataj, M., Senožetnik, M., Klemen, K., Pergar, P., Laspidou, C., & Mladenčić, D. 2018. Groundwater modeling with machine learning techniques: Ljubljana polje aquifer. *Proceedings*.
- Mobayen, R., Najafzadeh, M., & Farrahi-Moghaddam, K. 2023. Evaluation of regression-based soft computing techniques for estimating energy loss in gabion spillways. *Environment and Water Engineering*, 9(2), 241-255 .
- Modaresi, F., Araghinejad, S., & Ebrahimi, K. 2018. A comparative assessment of artificial neural network, generalized regression neural network, least-square support vector regression, and K-nearest neighbor regression for monthly streamflow forecasting in linear and nonlinear conditions. *Water resources management*, 32, 243-258 .
- Moravej, M., Amani, P., & Hosseini-Moghari, S.-M. 2020. Groundwater level simulation and forecasting using interior search algorithm-least square support vector regression (ISA-LSSVR). *Groundwater for Sustainable Development*, 11, 100447 .
- Mouatadid, S., Adamowski, J. F., Tiwari, M. K., & Quilty, J. M. 2019. Coupling the maximum overlap discrete wavelet transform and long short-term memory networks for irrigation flow forecasting. *Agricultural Water Management*, 219, 72-85 .
- Nadiri, A. A., Naderi, K., Khatibi, R., & Gharekhani, M. 2019. Modelling groundwater level variations by learning from multiple models using fuzzy logic. *Hydrological sciences journal*, 64(2), 210-226 .
- Pourmorad, S., Kabolizade, M., & Dimuccio, L. A. 2024. Artificial intelligence advancements for accurate groundwater level modelling: an updated synthesis and review. *Applied Sciences*, 14(16), 7358 .
- Prado, F., Minutolo, M. C., & Kristjanpoller, W. 2020. Forecasting based on an ensemble autoregressive moving average-adaptive neuro-fuzzy inference system-neural network-genetic algorithm framework. *Energy*, 197, 117159 .
- Prechelt, L. 2002. Early stopping-but when? In *Neural Networks: Tricks of the trade* (pp. 55-69). Springer .
- Roosbahani, A., Ebrahimi, E., & Banihabib, M. E. 2018. A framework for ground water management based on bayesian network and MCDM techniques. *Water resources management*, 32(15), 4985-5005 .
- Roy, D. K., Biswas, S. K., Mattar, M. A., El-Shafei, A. A., Murad, K. F. I., Saha, K. K., Datta, B., & Dewidar, A. Z. 2021. Groundwater level prediction using a multiple objective genetic algorithm-grey groundwater level in Urmia Plain aquifer using hybrid model of wavelet Transform-Extreme Learning Machine based on quantum particle swarm optimization. *Journal of Watershed Engineering and Management*, 12(2), 351-364 .
- Ahmadi, A., Olyaei, M., Heydari, Z., Emami, M., Zeynolabedin, A., Ghomlaghi, A., Daccache, A., Fogg, G. E., & Sadegh, M. 2022. Groundwater level modeling with machine learning: a systematic review and meta-analysis. *Water*, 14(6), 949 .
- Ahmed, A. N., Othman, F. B., Afan, H. A., Ibrahim, R. K., Fai, C. M., Hossain, M. S., Ehteram, M., & Elshafie, A. 2019. Machine learning methods for better water quality prediction. *Journal of Hydrology*, 578, 124084 .
- Banadkooki, F. B., Ehteram, M., Ahmed, A. N., Teo, F. Y., Fai, C. M., Afan, H. A., Sapitang, M., & El-Shafie, A. 2020. Enhancement of groundwater-level prediction using an integrated machine learning model optimized by whale algorithm. *Natural resources research*, 29, 3233-3252 .
- Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. 2015. *Time series analysis: forecasting and control*. John Wiley & Sons .
- Chidepudi, S. K. R., Massei, N., Jardani, A., Henriot, A., Allier, D., & Baulon, L. 2022. Groundwater level simulations using wavelet transform-assisted deep learning models. *AGU Fall Meeting Abstracts* ,
- Coleman, K. D., Schmidt, K., & Smith, R. C. 2016. Frequentist and Bayesian Lasso Techniques for Parameter Selection in Nonlinearly Parameterized Models. *IFAC-PapersOnLine*, 49(18), 416-421 .
- Di Nunno, F., Giudicianni, C., Creaco, E. & Granata, F. 2023. Multi-step ahead groundwater level forecasting in Grand Est, France: Comparison between stacked machine learning model and radial basis function neural network. *Groundwater for Sustainable Development*, 23, 101042 .
- Famiglietti, J. S. 2014. The global groundwater crisis. *Nature Climate Change*, 4(11), 945-948 .
- Friedman, J. 2009. *The elements of statistical learning: Data mining, inference, and prediction*.
- Gonzalez, R. Q., & Arsanjani, J. J. 2021. Prediction of groundwater level variations in a changing climate: a Danish case study. *ISPRS International Journal of Geo-Information*, 10(11), 792 .
- Kayhomayoon, Z., Azar, N. A., Milan, S. G., Moghaddam, H. K., & Berndtsson, R. 2021. Novel approach for predicting groundwater storage loss using machine learning. *Journal of Environmental Management*, 296, 113237 .
- Kellaway, F. 1969. *Advanced Engineering*

- Sciences, 20(7), 2611–2628 .
- Soni, A., Barma, S. D., & Mahesha, A. 2024. Optimizing Groundwater Forecasting: Comparative Analysis of MLP Models Using Global and Regional Precipitation Data. EGU General Assembly Conference Abstracts.
- Solgi, A., Dr, H. Z., yasuj, F. F., & Eskandari, A. 2019. Evaluation of Combined ANFIS with Wavelet Transform to Modeling and Forecasting Groundwater Level. *Journal of Watershed Management Research*, 9(18), 56-69.
- Tang, Y., Zang, C., Wei, Y., & Jiang, M. 2019. Data-driven modeling of groundwater level with least-square support vector machine and spatial-temporal analysis. *Geotechnical and Geological Engineering*, 37(3), 1661–1670 .
- Tibshirani, R. 1996. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 5,8(1),267-288 .
- Yan, Q., & Ma, C. 2016. Application of integrated ARIMA and RBF network for groundwater level forecasting. *Environmental Earth Sciences*, 75(5), 396 .
- Yin, W., Fan, Z., Tangdamrongsub, N., Hu, L., & Zhang, M. 2021. Comparison of physical and data-driven models to forecast groundwater level changes with the inclusion of GRACE–A case study over the state of Victoria, Australia. *Journal of Hydrology*, 602, 126735.
- Zare, M., & Koch, M. (2018). Groundwater level fluctuations simulation and prediction by ANFIS-and hybrid Wavelet-ANFIS/Fuzzy C-Means (FCM) clustering models :Application to the Miandarband plain. *Journal of Hydro-Environment Research*, 18, 63-76 .
- Zhang, H., Zhao, J., & Chen, C. 2021. Groundwater level prediction based on neural networks: a case study in Linze, northwestern China. *E3S Web of Conferences*.
- relational analysis based weighted ensemble of ANFIS models. *Water*, 13(21), 3130 .
- Saberifar, R. 2023. Climate Change and Water Crisis (Case Study, Mashhad in Northeastern Iran). *Polish Journal of Environmental Studies*, 3.(1) .
- Saha, S., Mallik, S., & Mishra, U .2022. Groundwater depth forecasting using machine learning and artificial intelligence techniques: a survey of the literature. *Recent Developments in Sustainable Infrastructure (ICRDSI-2020)—GEO-TRA-ENV-WRM: Conference Proceedings from ICRDSI-2020 Vol, 2*.
- Sahoo, M., Das, T., Kumari, K., & Dhar, A. 2017. Space-time forecasting of groundwater level using a hybrid soft computing model. *Hydrological sciences journal*, 62(4), 561–574 .
- Sahu, R. K., Müller, J., Park, J., Varadharajan, C., Arora, B., Faybishenko, B., & Agarwal, D. 2020. Impact of input feature selection on groundwater level prediction from a multi-layer perceptron neural network. *Frontiers in Water*, 2, 573034 .
- Sapitang, M., Ridwan, W. M., Ahmed, A. N., Fai, C. M., & El-Shafie, A. 2021. Groundwater level as an input to monthly predicting of water level using various machine learning algorithms. *Earth Science Informatics*, 14(3), 1269–1283 .
- Shivaprakash, K. N., Swami, N., Mysorekar, S., Arora, R., Gangadharan, A., Vohra, K., Jadeyegowda, M., & Kiesecker, J. M. 2022. Potential for artificial intelligence (AI) and machine learning (ML) applications in biodiversity conservation, managing forests, and related services in India. *Sustainability*, 14(12), 7154 .
- Shortridge, J. E., Guikema, S. D., & Zaitchik, B. F. 2016. Machine learning methods for empirical streamflow simulation: a comparison of model accuracy, interpretability, and uncertainty in seasonal watersheds. *Hydrology and Earth System*

Comparison of LS-SVR, ANFIS, MLP, and RBF Model Performance in Groundwater Level Modeling (Case Study: Central Part of Mashhad Plain)

O. Mousavikhaledi¹, A. Firoozzare^{*2}, A. Dourandish³, F. Modaresi⁴

Recived: Sep.11, 2025

Accepted: May.15, 2026

Abstract

The decline in groundwater levels in Mashhad Plain, primarily driven by anthropogenic factors, presents a significant challenge to sustainable water resource management. The heavy reliance on these resources necessitates accurate prediction of changes for effective managerial decision-making. This study employs machine learning methods to simulate and predict groundwater level variations in the central part of the plain. Initially, key input variables and optimal time lags were identified using the Partial Autocorrelation Function (PACF) and Frequency Lasso Regression (Frequentist Lasso Regression). Subsequently, the performance of Least Squares Support Vector Regression (LS-SVR), Adaptive Neuro-Fuzzy Inference System (ANFIS), Multi-Layer Perceptron Neural Network (MLP), and Radial Basis Function Neural Network (RBF) was evaluated for monthly prediction of groundwater level fluctuations over 30 years (1991-2021). Results demonstrated that all four models simulate groundwater levels with acceptable accuracy. The RBF model exhibited superior performance with R^2 , MSE, NSE, and RMSE values of 1.00, 0.00, 1.00, and 0.009, respectively. The findings affirm the high potential of data-driven models in simulating hydrological processes and can serve as a foundation for developing intelligent tools for groundwater resource management in Mashhad Plain and similar regions, including resource allocation and warning system design.

Keywords: Data-Driven Models, Groundwater Level Fluctuations, Machine Learning Models, Supervised Learning Methods

1- Ph.D. student in Agricultural Economics, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran.

2- Associate Professor of Agricultural Economics, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran

3- Associate Professor of Agricultural Economics, Faculty of Agriculture, University of Tehran, Tehran, Iran

4- Assistant Professor of Water Science and Engineering, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran

(* - Corresponding Author Email: firooz@um.ac.ir)