

پیش‌بینی ماهانه سطح آب زیرزمینی آبخوان نهاوند با بکارگیری مدل یادگیری عمیق و ماشین بردار پشتیبان

چکیده

کاهش بارندگی و بهره‌برداری بی‌رویه از منابع آب زیرزمینی موجب افت قابل‌توجه تراز سطح آب زیرزمینی در بسیاری از مناطق جهان شده است. یکی از آبخوان‌هایی که با افت سطح آب زیرزمینی مواجه هست، آبخوان نهاوند در استان همدان و غرب ایران می‌باشد. برای مدلسازی تراز سطح آب زیرزمینی، مدل‌های حافظه کوتاه مدت طولانی (Long Short-Term Memory, LSTM) و ماشین بردار پشتیبان (Support Vector Machine, SVM) به کار گرفته شد. نقشه‌های ماهانه هم بارش، هم دما و هم تبخیر با نرم افزار ArcGIS ترسیم شد. با استفاده از اطلاعات چاه‌های مشاهداتی و تکمیل آمار چاه‌های موجود، هیدروگراف واحد آبخوان ترسیم شد. داده‌های ماهانه دوره آماری ۱۳۷۵ تا ۱۴۰۳ برای آموزش و آزمون مدل‌ها مورد استفاده قرار گرفت. آنالیز حساسیت تراز سطح آب زیرزمینی یک ماه آینده نسبت به تغییرات متغیرهای ورودی دو مدل LSTM و SVM نشان داد که مقادیر بارش، دمای هوا، تبخیر و تراز سطح آب زیرزمینی تا دو ماه قبل، متغیرهای موثر ورودی هستند. نتایج مدلسازی با مدل SVM نشان داد که تابع کرنل پایه شعاعی در مدل SVM بهترین تابع کرنل نسبت به توابع کرنل خطی و چندجمله‌ای می‌باشد. مقایسه نتایج مدل‌های هوشمند نشان داد که مدل SVM عملکرد بهتری نسبت به مدل LSTM داشته است. همچنین برای دوره زمانی آزمون مدل LSTM ضریب تعیین (Coefficient of determination, R^2) و ریشه میانگین مربعات خطا (Root Mean Square Error, RMSE) به ترتیب ۰/۸۷ و ۰/۰۴۳ بدست آمد. علاوه بر آن برای دوره زمانی آزمون مدل SVM ضریب R^2 و RMSE به ترتیب ۰/۸۹ و ۰/۰۴۲ تعیین شد. به‌طور کلی می‌توان گفت که عملکرد مدل SVM برای پیش‌بینی تراز سطح آب زیرزمینی آبخوان نهاوند بسیار مناسب بوده و قابلیت این را دارد که برای بررسی سطح آب زیرزمینی آبخوان‌های دیگر نیز استفاده شود.

واژگان کلیدی: پیش‌بینی، سطح آب زیرزمینی، آبخوان نهاوند، LSTM، SVM.

مقدمه

کاهش بارندگی و برداشت بیش از حد از منابع آب زیرزمینی در سال‌های گذشته، باعث وارد شدن فشار بیش از حد به منابع آب زیرزمینی شده است. این امر موجب افت تراز سطح آب زیرزمینی در آبخوان‌های ایران شده است. در طی دوره خشک، قابلیت تبخیر و ترقق افزایش یافته و پس از یک دوره زمانی کوتاه، مقدار ذخیره رطوبت خاک تحت تأثیر این شرایط شروع به کاهش می‌نماید که در نتیجه آن تغذیه سیستم آب زیرزمینی کم می‌شود. سطح تراز آب چاه‌ها ممکن است یک سال یا بیش‌تر پس از شروع خشک‌سالی به کمبود بارندگی واکنش نشان دهد (Moustadraf et al., 2003). این مسئله می‌تواند روی جریان سطحی تأثیرگذار باشد زیرا به آب‌های زیرزمینی منطقه وابسته است (Havril et al., 2018; Van Loon et al., 2013). آثار تهنی‌شدگی سفره‌های آب زیرزمینی پیچیده و وابسته به آبخوان بوده و مشکلات متداول بسیاری در این شرایط ایجاد می‌شود. از جمله مهمترین آنها کاهش سطح ایستابی است که افزایش هزینه پمپاژ و خشک شدن چاه‌های آب را به دنبال دارد (Mukul Fishman et al., 2011). از دیگر آثار آن می‌توان به کاهش تخلیه آب زیرزمینی به رودخانه‌ها و چشمه‌ها، تأثیر بر اکوسیستم‌های موجود و نیز فرونشست زمین اشاره کرد (Sophocleous, 2000). کاهش سطح آب زیرزمینی همچنین می‌تواند باعث افزایش شوری از طریق نفوذ آب شور دریا به آبخوان‌های ساحلی یا نشت از لایه‌های مجاور حاوی آب شور شود (Konikow and Kendy, 2005). با توجه به این که آب مصرفی برای بخش قابل توجهی از مردم ایران از منابع

سمیه عبدی^۱

حسین فتحیان^{۲*}

مهدی اسدی لور^۱

اصلان اکدرنژاد^۱

علی عصاره^۱

۱. گروه آبیاری و زهکشی، واحد اهواز، دانشگاه آزاد اسلامی،

اهواز، ایران.

۲. گروه مهندسی منابع آب، واحد اهواز، دانشگاه آزاد

اسلامی، اهواز، ایران.

* نویسنده مسئول مکاتبات

Fathian.h@iau.ac.ir

تاریخ دریافت: ۱۴۰۴/۱۰/۰۹

تاریخ پذیرش: ۱۴۰۵/۰۴/۰۵

این مقاله برگرفته از رساله دکتری است.

آب زیرزمینی تأمین می‌شود، بررسی و مطالعه این‌گونه آب‌ها، یکی از چالش‌هایی است که در چند سال گذشته در جهت برطرف نمودن مشکلات تأمین منابع آبی مطرح شده است. افت سطح آب زیرزمینی در سال‌های گذشته، سبب شده است، پژوهشگران متعددی در سراسر جهان به دنبال راه‌کارهای مختلفی برای برآورد و پیش‌بینی سطح آب زیرزمینی باشند تا با استفاده از این برآوردها بتوانند وضعیت آب زیرزمینی را به صورت دقیق‌تر و بهتری بررسی نمایند. ضرورت استفاده از تجهیزات و امکانات جدید و به روز، مقوله اجتناب‌ناپذیری است که بشر در بسیاری از علوم پیش روی خود می‌بیند. یکی از مزایای این ابزارهای جدید کاهش زمان پیش‌بینی‌ها و افزایش دقت مدل‌ها است. به منظور دستیابی به این هدف می‌توان نتایج حاصل از این ابزارها را با مقادیر مشاهداتی مقایسه نمود. در این زمینه مطالعه‌های متعددی انجام شده است که در ادامه به برخی از آن‌ها اشاره می‌گردد.

Sattari *et al.* (2017) برای پیش‌بینی سطح آب زیرزمینی در دشت اردبیل از مدل‌های رگرسیون بردار پشتیبان (Support vector regression, SVR) و مدل درختی M5 استفاده کردند. نتایج آن‌ها نشان داد که عمل کرد هر دو مدل خوب بوده ولی استفاده از مدل M5 راحت‌تر و تفسیر آن آسان‌تر است. (Mirarabi *et al.* (2019) ارزیابی مدل‌های SVR و شبکه عصبی مصنوعی (Artificial neural network, ANN) را برای پیش‌بینی سطح آب زیرزمینی انجام داده و نتایج نشان داد که عمل کرد مدل SVR بهتر از شبکه عصبی مصنوعی بوده است. (Mujahid *et al.* (2020) با استفاده از مدل شبکه عصبی مصنوعی، ارتباط بین پارامترهای هواشناسی و سطح آب زیرزمینی موجود در محدوده رودخانه راوی و سوتلج را بررسی کردند. در این تحقیق از پارامترهای هواشناسی شامل بارندگی، حداکثر دما، حداقل دما، تابش خورشیدی، سرعت باد، رطوبت نسبی، ارتفاع سطح، ناحیه چند ضلعی و عمق سطح آب استفاده شد. براساس نتایج همبستگی بین سطوح آب زیرزمینی با پارامترهای هواشناسی مشاهده شد که میزان بارندگی در مقایسه با هر پارامتر دیگری برای سایت انتخاب شده، تأثیر بیشتری بر تغییرات سطح آب زیرزمینی دارد. (Kayhomayoon *et al.* (2021) رویکرد جدید برای پیش‌بینی تلفات ذخیره‌سازی آب زیرزمینی با استفاده از یادگیری ماشین ارائه نمودند. از بین متغیرهای مورد بررسی، مصرف آب کشاورزی، سطح آبخوان، نفوذ از رودخانه و زهکش به ترتیب با همبستگی ۰/۸۴، ۰/۷۹، ۰/۷۰ و ۰/۶۹ با تلفات ذخیره‌سازی آب زیرزمینی ارتباط زیادی داشتند. همچنین دو روش سیستم استنتاج عصبی-فازی تطبیقی بهینه‌سازی هریس هاکس (HHO-ANFIS: Harris Hawks optimization adaptive neuro-fuzzy inference system) و ماشین بردار پشتیبان حداقل مربعات (Least-squares support vector machine, LS-SVM) بهترین نتایج را داشتند. به طور کلی، HHO-ANFIS برای پیش‌بینی تلفات ذخیره‌سازی آب زیرزمینی کارآمدتر بوده و روش‌های هوش مصنوعی و یادگیری ماشین می‌توانند در زمان و هزینه‌ها برای این محاسبات پیچیده صرفه‌جویی کنند. (Aderemi *et al.* (2023) از مدل‌های یادگیری ماشین مانند مدل‌های رگرسیون، مدل‌های رگرسیون خودکار عمیق (Deep auto-regressive) و شبکه‌های عصبی اتورگرسیو غیرخطی با ورودی خارجی (Nonlinear Autoregressive Neural Network with Exogenous Inputs, NARX Network) برای پیش‌بینی سطح آب زیرزمینی در کمربند کارست در آفریقای جنوبی استفاده کردند. نتایج آن‌ها نشان داد که NARX و ماشین بردار پشتیبان عمل کرد بهتری نسبت به سایر مدل‌های مورد استفاده دارند. خطیبی‌نیا (۱۳۹۴) پیش‌بینی نوسانات سطح آب زیرزمینی آبخوان بهاباد را با استفاده از مدل عددی مورد بررسی قرار داد. پیش‌بینی صورت گرفته بر اساس مدل واسنجی و صحت سنجی شده آبخوان مورد مطالعه برای سال‌های آینده (فروردین ۹۴ تا اسفند ۹۹) نشان می‌دهد که با ادامه روند کنونی برداشت آب، سطح آب زیرزمینی در انتهای دوره مورد پیش‌بینی به ویژه در نواحی با تمرکز زیاد چاه‌های بهره‌برداری، با افت شدیدی (۳/۲ متر در طول پنج سال) مواجه خواهد شد. ندیری و همکاران (۱۳۹۵) پیش‌بینی زمانی و مکانی سطح آب زیرزمینی آبخوان دوزدوزان را با استفاده از روش‌های هوش مصنوعی و زمین آمار انجام دادند. پس از نرمال‌سازی داده‌ها مدل‌سازی با ANN انجام شد. به منظور بررسی بیشتر شبیه‌سازی با مدل فازی سگنو (Sugeno Fuzzy Logic, SFL) نیز انجام شد. نتایج نشان داد که مدل کوکریجینگ با در نظر گرفتن پارامتر ثانویه توپوگرافی نسبت به مدل کریجینگ پیش‌بینی دقیق‌تری داشته است.

یاری و درزی نفت چالی (۱۳۹۶) به بررسی پیش‌بینی نوسانات سطح آب زیرزمینی با استفاده از مدل MODFLOW پرداختند. بر اساس نتایج مدل، بهترین سناریو برای تعادل بخشی و احیاء آب زیرزمینی، تغذیه مصنوعی با استفاده از پساب تا افق ۱۴۰۰ می‌باشد. پناهی

و همکاران (۱۳۹۷) شبیه‌سازی و تخمین نوسانات سطح آب زیرزمینی دشت زنجان را با استفاده از GMS مورد بررسی قرار دادند. مدل با فرض نرخ کاهش آب‌های زیرزمینی و مقدار ثابت تغذیه برای فاصله زمانی ۱۵ ساله از سال ۲۰۰۷ تا ۲۰۲۲ اجرا شد که نقشه‌های به‌دست‌آمده برای تراز آب‌های زیرزمینی در این سال‌ها نشان‌دهنده افت شدید تراز آب با ادامه برداشت‌های بی‌رویه از ذخیره آب‌های زیرزمینی آبخوان زنجان است. جمالی‌زاده و همکاران (۱۳۹۹) پیش‌بینی نوسانات سطح آب زیرزمینی دشت رفسنجان را با استفاده از مدل‌های سری زمانی و GMS مورد بررسی قرار دادند. نتایج شبیه‌سازی افت آب زیرزمینی نیز نشان داد که بیشترین میزان افت آب زیرزمینی در بخش‌های جنوب غرب دشت اتفاق افتاده است و سالیانه حدود ۱۳۰ میلیون مترمکعب اضافه برداشت از منابع آب زیرزمینی صورت می‌گیرد. خیاط و همکاران (۱۴۰۰) پیش‌بینی سطح آب زیرزمینی دشت بیرجند را با استفاده از شبکه عصبی-فازی تحت تأثیر تغییرات اقلیم مورد بررسی قرار دادند. نتایج تحقیق حاکی از سیر نزولی و افت تقریبی ۳/۵ متر سطح آب زیرزمینی طی سال‌های آینده تا سال ۲۰۳۷ می‌باشد. سلامتیان و همکاران (۱۴۰۲) تراز آب زیرزمینی دشت قم را با استفاده از مدل سیستم استنتاج فازی عصبی تطبیقی پیش‌بینی نمودند. نتایج نشان داد که استفاده از داده‌های تخلیه از چاه‌ها در مدل‌سازی‌ها موجب بهبود پیش‌بینی می‌گردد. سلیمان‌پور و همکاران (۱۴۰۴) روند تغییرات دوره‌ای سطح آب زیرزمینی حوزه آبخیز بهادران در استان یزد را ارزیابی کردند. مطابق نتایج، بیش‌ترین عمق آب زیرزمینی در آبخوان‌های بهادران و شمس در سال ۱۳۹۷ به ترتیب برابر با ۶۸ و ۴۷/۷ متر است. Bahmani et al. (2020) برای مدل‌سازی سطح آب زیرزمینی، مدل‌های برنامه‌ریزی بیان ژن (Gene Expression Programming, GEP) و M5 را با تکنیک‌های تبدیل موجک و (Complete Ensemble Empirical Mode Decomposition, CEEMD) ترکیب کردند. نتایج نشان داد که مدل ترکیب شده با GEP عملکرد بهتری داشته است. Shahbazi et al. (2023) مدل‌سازی سطح آب زیرزمینی آبخوان آسپاس را با استفاده از تکنیک‌های تجزیه داده و ترکیب با مدل‌های هوش مصنوعی انجام دادند. در این مطالعه دو مدل ANN و رگرسیون بردار پشتیبان (SVR)، با دو ابزار پیش‌پردازش تبدیل موجک و CEEMD مورد استفاده قرار گرفتند. بر اساس نتایج، مدل ترکیبی CEEMD-ANN دقیق‌ترین روش برای پیش‌بینی سطح آب زیرزمینی شناخته شد.

هدف از این تحقیق پیش‌بینی ماهانه تراز سطح آب زیرزمینی آبخوان نهاوند با بکارگیری مدل یادگیری عمیق و ماشین بردار پشتیبان است. افت شدید سطح آب زیرزمینی آبخوان نهاوند در حوضه کرخه در سال‌های گذشته باعث شد که در این تحقیق به بررسی وضعیت سطح آب زیرزمینی این آبخوان پرداخته شود. علاوه بر این متغیرهای ورودی موثر در مدل یادگیری عمیق و ماشین بردار پشتیبان با روش آنالیز حساسیت تعیین می‌شود.

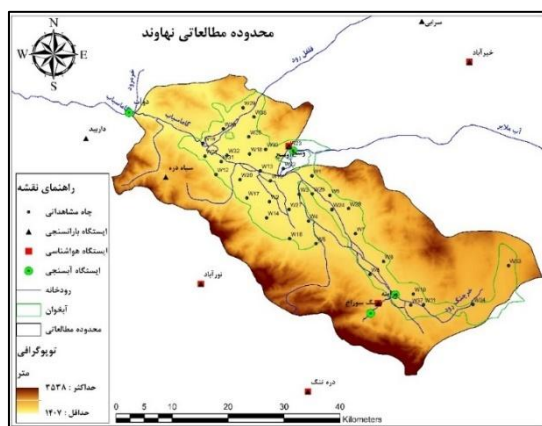
مواد و روش‌ها

منطقه مورد مطالعه

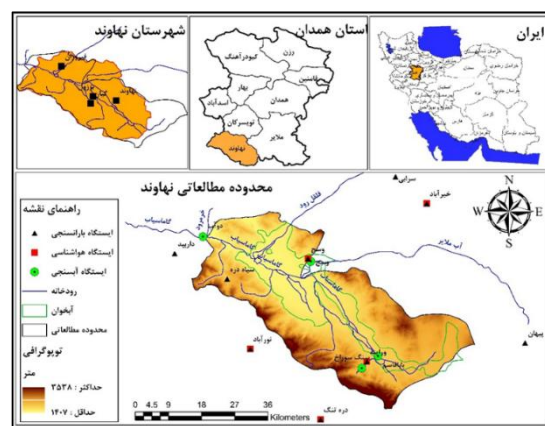
دشت نهاوند در شمال شرقی حوضه کرخه واقع شده است. این منطقه از سمت شمال به محدوده کنگاور و تویسرکان، از سمت شمال شرق، شرق و جنوب شرقی به ملایر، از جنوب به محدوده‌های اشترینان و الشتر، از جنوب غرب به نورآباد، از غرب به صحنه محدود است. این دشت با مساحت ۱۷۷۹/۲ کیلومتر مربع، زیر نظر شرکت سهامی آب منطقه‌ای همدان است. حداکثر ارتفاع در این ناحیه ۳۵۳۸ متر، در ارتفاعات رشته کوه گرین، در جنوب غربی منطقه و حداقل آن ۱۴۰۷ متر، در قسمت خروجی بر روی رودخانه گاماسیاب می‌باشد. در جدول (۱) مشخصات ایستگاه‌های آبسنجی و ایستگاه‌های هواشناسی داخل و اطراف دشت نشان داده شده است. در شکل (۱) موقعیت دشت نهاوند در حوضه آبریز کرخه، استان همدان و کشور ایران نشان داده شده است. مهمترین رودخانه‌ای که در این منطقه جریان دارد، رودخانه گاماسیاب می‌باشد. رودخانه گاماسیاب از شاخه‌های مهم رود کرخه به شمار می‌رود که در نهایت به تالاب هورالعظیم می‌ریزد. مساحت آبخوان آبرفتی نهاوند ۵۰۲/۲ کیلومتر مربع است. در شکل (۲) موقعیت چاه‌های مشاهداتی به همراه ایستگاه‌های باران سنجی، هواشناسی و آب سنجی نشان داده شده است.

جدول ۱: مشخصات ایستگاه‌های مورد استفاده

نام ایستگاه	متغیر	کد ایستگاه	طول جغرافیایی (متر)	عرض جغرافیایی (متر)	ارتفاع از سطح دریا (متر)
سیاه دره		۲۱-۷۱۲	۷۷۴۹۳۷	۳۷۹۵۹۴۵	۱۷۵۷
سرابی		۲۱-۸۸۲	۲۶۸۲۵۲	۳۸۳۳۷۸۹	۱۹۲۵
پیهان	بارش (میلیمتر)	۲۱-۹۸۱	۳۰۳۴۵۵	۳۷۷۸۶۳۵	۱۸۷۰
داربید		۲۱-۵۴۶	۷۶۰۳۰۲	۳۸۰۲۰۰۳	۱۸۶۰
وراینه		۲۱-۵۳۴	۲۶۰۴۶۲	۳۷۷۳۵۷۵	۱۷۹۵
وسج	بارش (میلیمتر)،	۲۱-۳۷۵	۲۴۴۴۹۵	۳۸۰۱۵۳۱	۱۵۶۶
دره تنگ الشتر	دما (سانتی‌گراد)،	۲۱-۹۵۴	۲۴۷۹۰۷	۳۷۵۷۸۳۷	۱۷۳۰
نورآباد	تبخیر (میلی‌متر)	۲۱-۱۴۵	۲۲۸۷۹۳	۳۷۷۷۰۲۳	۱۸۲۰
خیرآباد		۲۱-۰۳۲	۲۷۶۷۵۳	۳۸۱۶۵۲۰	۱۷۷۰



شکل ۲: موقعیت چاه‌های مشاهداتی در منطقه مورد مطالعه



شکل ۱: موقعیت محدوده مطالعاتی نهاوند در استان همدان

آماده‌سازی داده‌ها

اطلاعات مورد نیاز شامل مقادیر ماهانه بارش، دمای هوا و تبخیر در ایستگاه‌های هواشناسی و داده‌های چاه‌های مشاهداتی در دوره آماری ۱۳۷۵ تا ۱۴۰۳ از شرکت آب منطقه‌ای استان همدان تهیه گردید. با توجه به اطلاعات چاه‌های مشاهده‌ای در آبخوان نهاوند، ۳۷ چاه که دارای دوره آماری بهتری بودند، انتخاب گردید. بعد از تکمیل آمار سطح آب زیرزمینی با استفاده از نرم افزار HEC-4، دوره آماری ۲۸ ساله (۱۳۷۵-۱۴۰۳) برای ترسیم هیدروگراف واحد انتخاب شد. به منظور بررسی داده‌های پرت از روش نمودار جعبه‌ای (Box plot) استفاده شد. در صورت مواجهه با داده‌های پرت، آن داده‌ها حذف و دوباره با نرم‌افزار HEC-4 بازسازی شدند. برای ترسیم نقشه‌ها از نرم‌افزار GIS و روش کریجینگ استفاده شد.

نقشه‌های هم‌بارش ماهانه برای دشت ترسیم شد و سپس مقادیر بارش ماهانه استخراج گردید. برای بررسی داده‌های دما و افزایش طول دوره آماری آن‌ها از روش تفاضلات استفاده شد. بعد از بررسی داده‌های پرت، نقشه‌های هم‌دمای ماهانه ترسیم شد و سپس مقادیر دمای ماهانه استخراج گردید. برای بررسی داده‌های تبخیر و افزایش طول دوره آماری آن‌ها از رابطه بین دما و تبخیر برای هر ایستگاه استفاده شد. سپس نقشه‌های هم تبخیر ماهانه ترسیم و مقادیر تبخیر ماهانه استخراج شد.

استاندارد سازی داده‌ها

برای استاندارد سازی داده‌ها از رابطه (۱) استفاده شد (سلگی، ۱۳۹۳).

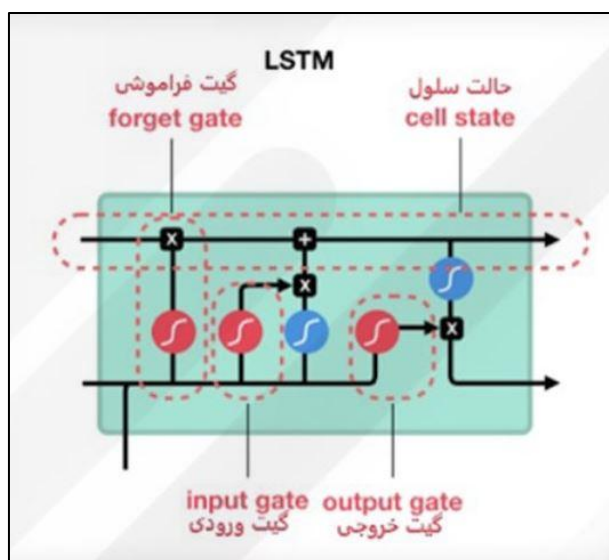
$$y = 0.5 + \left[0.5 \times \left(\frac{x - \bar{x}}{x_{\max} - x_{\min}} \right) \right] \quad (1)$$

که X داده مورد نظر، $\bar{X}$ میانگین داده‌ها، $X_{\max}$ حداکثر داده‌ها، $X_{\min}$ حداقل داده‌ها و y داده استاندارد شده است.

مدل حافظه کوتاه مدت طولانی (LSTM)

شبکه LSTM، نوعی خاص از شبکه عصبی بازگشتی (Recurrent Neural Network, RNN) محسوب می‌شود. شبکه RNN نوعی شبکه عصبی مصنوعی است که حافظه‌ی داخلی دارد. به عبارت دیگر، این شبکه یک شبکه عصبی معمولی است که در ساختارش حلقه‌ای دارد که از طریق آن در هر گام، خروجی گام قبلی، به همراه ورودی جدید، به شبکه وارد می‌شود. این حلقه به شبکه کمک می‌کند تا اطلاعات قبلی را در کنار اطلاعات جدید داشته باشد و بتواند براساس این اطلاعات خروجی مدنظر را به ما بدهد. شبکه RNN در مسائلی که نیاز باشد حافظه بلندمدت داشته باشد نمی‌تواند خیلی خوب عمل کند. دلیل این موضوع مفهومی به نام محوشدگی گرادیان (Vanishing Gradient) است. این مسئله باعث به وجود آمدن شبکه LSTM شده است. شبکه LSTM مشکل حافظه‌ی بلندمدت شبکه RNN را حل می‌کند. شبکه LSTM سازوکارهایی داخلی به اسم گیت (Gate) دارد. این گیت‌ها جریان اطلاعات را کنترل می‌کنند، همین‌طور مشخص می‌کنند چه داده‌هایی در توالی مهم هستند و باید هم‌چنان حفظ بشوند و چه داده‌هایی باید حذف شوند. به این شکل، شبکه اطلاعات مهم را در طول زنجیره‌ی توالی عبور می‌دهد تا خروجی مدنظر بدست آید. در شکل (۳) ساختار یک شبکه LSTM مشاهده می‌شود. مفهوم اصلی شبکه LSTM همان cell state و گیت‌های همراهش است، گیت‌ها اطلاعات را در cell state به روز نگه می‌دارند. این گیت‌ها شبکه‌های عصبی مختلفی هستند که تصمیم می‌گیرند چه اطلاعاتی به cell state وارد بشوند. این گیت‌ها در طول آموزش شبکه یاد می‌گیرند چه اطلاعاتی باید حفظ یا فراموش شوند. برای اجرای مدل LSTM، با کدنویسی در محیط نرم افزار MATLAB و استفاده از تمامی پارامترهای موثر در اجرای مدل، اقدام به مدل‌سازی شد.

با مقایسه و ارزیابی شاخص‌های آماری مانند ریشه میانگین مربعات خطا برای الگوریتم‌های آموزش مختلف، می‌توان ارزیابی کرد که الگوریتم کاهش گرادیان منجر به ایجاد حداقل خطای مدل شده است. در شبکه‌ی LSTM تعمیم‌پذیری از طریق کنترل پیچیدگی معماری شبکه (تعداد لایه‌ها و نورون‌ها)، استفاده از تکنیک‌های منظم‌سازی و به کارگیری راهبرد توقف زود هنگام (Early Stopping) تضمین می‌شود. این اقدامات با کاهش واریانس مدل و جلوگیری از بیش‌برازش، تعادل مناسبی میان بایاس و واریانس برقرار می‌کنند. علاوه بر این، نرمال‌سازی داده‌های ورودی و انتخاب بهینه طول پنجره زمانی نقش مهمی در پایداری گرادیان و بهبود عملکرد مدل بر روی مجموعه آزمون دارند.



شکل ۳: ساختار شماتیکی از مدل LSTM (Hochreiter and Schmidhuber, 1997)

مدل SVM

ماشین بردار پشتیبان، یکی از روش‌های یادگیری با نظارت (Supervised learning) است که از آن برای طبقه‌بندی و رگرسیون استفاده می‌کنند. ماشین بردار پشتیبان یک طبقه‌بندی کننده دودویی غیرآماري است که در سال‌های اخیر بسیار مورد توجه قرار گرفته است. خصوصیات SVM عبارتند از: (۱) ایجاد حدکثر تعمیم در طراحی طبقه‌بندی کننده‌ها، (۲) دارا بودن توانایی یافتن جواب بهینه کلی تابع، (۳) تهیه خودکار ساختار و سازوکار بهینه در حل مسائل طبقه‌بندی، (۴) مدل‌سازی توابع غیرخطی با استفاده از کرنل‌های غیرخطی و شیوه

حاصل ضرب داخلی در فضاهای هیلبرت. ماشین بردار پشتیبان الگوریتمی است که مدل‌های خطی منحصر به فرد را یافته و با استفاده از آنها حداکثر حاشیه ابر صفحه را بدست می‌آورد. حداکثر شدن حاشیه ابر صفحه، حداکثر شدن تفکیک بین طبقات را حاصل نموده و موجب بهبود دقت در فرآیند مدلسازی می‌شود. به نزدیکترین نقاط آموزشی موجود در حاشیه ابر صفحه، بردارهای پشتیبان اطلاق گردیده و از آن‌ها برای مشخص کردن مرز بین طبقات استفاده می‌شود (Shin et al., 2005).

تفاوت اساسی این روش با طبقه‌بندی‌کننده‌های مرسوم آماری در قابلیت پردازش و طبقه‌بندی داده‌های ابرطیفی بوده و کاهش تعداد باندها در فرآیندهای مختلف مدلسازی نیاز نمی‌باشد. همچنین در این روش با استفاده از یک الگوریتم بهینه‌سازی تعداد نمونه‌های مورد نیاز برای تشکیل مرز کلاس‌ها به دست آمده و یک مرز تصمیم‌گیری خطی بهینه برای جداسازی کلاس‌ها محاسبه می‌شود. اگر داده‌ها به صورت خطی و جدا از هم باشند، SVM با استفاده از ماشین‌های خطی یک سطح بهینه را با کمترین خطا و حداکثر فاصله میان صفحه و نزدیکترین نقاط آموزشی (بردارهای پشتیبان) تفکیک نموده و آموزش می‌دهد (Shin et al., 2005).

اگر نقاط آموزشی را به صورت $[x_i, y_i]$ و بردار ورودی حائز شرط $x_i \in R^n$ باشد، در این صورت ارزش هر طبقه به صورت $y_i \in \{-1, 1\}$ $i = 1, \dots, i$ تعریف می‌گردد. آن‌گاه قواعد تصمیم‌گیری به نحوی تعریف شده توسط یک صفحه بهینه که طبقات تصمیم‌گیری باینری را تفکیک می‌کند که به صورت رابطه (۲) قابل بیان می‌باشد:

$$Y = \text{sgn}[\sum_{i=1}^N y_i a_i (X \times X_i) + b] \quad (2)$$

که در آن Y خروجی، y_i ارزش طبقه نمونه X_i و a_i و b پارامترهای تعیین کننده ابر صفحه می‌باشند. اگر امکان تفکیک داده‌ها بصورت خطی فراهم نباشد، در این صورت رابطه (۲) به شرح زیر تغییر می‌یابد:

$$Y = \text{sgn}[\sum_{i=1}^N y_i a_i K(X \times X_i) + b] \quad (3)$$

در این رابطه $K(X \times X_i)$ تابع کرنلی است که برای ایجاد ماشین‌های بردار پشتیبان با حالت‌های مختلف از سطوح تصمیم‌گیری غیرخطی در فضای داده‌ها، ضرب‌های داخلی تولید کرده و بدین منظور نیاز به تعریف معادله خط می‌باشد. معادله خط در فضای دوبعدی با استفاده از رابطه (۴)، معادله صفحه از رابطه (۵) و معادله صفحه نمایش از رابطه (۶) محاسبه می‌گردد.

$$w_1 x_1 + w_2 x_2 + b = 0 \quad (4)$$

$$w_1 x_1 + w_2 x_2 + w_3 x_3 + b = 0 \quad (5)$$

$$w^T x + b = 0 \quad (6)$$

در مسائلی که داده‌ها به طور خطی تفکیک‌پذیر نباشند، می‌توان با نگاشت داده‌ها به یک فضای ویژگی، آن‌ها را به صورت خطی جداپذیر نمود. در حالت کلی ابعاد این فضا بینهایت بوده، بنابراین انجام محاسبات در فضای ویژگی می‌تواند زمان‌بر باشد. برای غلبه بر این مشکل از توابع کرنل استفاده می‌شود، در نتیجه معادله صفحه تفکیک‌کننده برای حالت غیرخطی با دخالت تابع کرنل $K(X)$ به صورت رابطه زیر می‌باشد.

$$w^T K(X) + b = 0 \quad (7)$$

که در آن تابع $K(X)$ با انجام نگاشت، داده‌ها را از فضای غیرخطی به یک فضای خطی منتقل می‌نماید. می‌توان این تابع را به صورت $K(X_i, X_j)$ نیز تعریف نموده و به وسیله آن فضاهای پیچیده‌تری را مورد بررسی قرار داد. به این فرآیند ترفند کرنل (Kernel trick) گفته می‌شود. در مدل SVM انتخاب تابع کرنل بسیار مهم بوده و در مسایل مختلف با توجه به ماهیت مسئله می‌توان موارد متفاوتی را در نظر گرفت. بنابراین نمی‌توان تابعی را به طور قطعی به عنوان تابع مناسب برای SVM معرفی کرد. انواع مختلف تابع در جدول ۲ ارائه شده است. تابع کرنل خطی، حالت خاصی از تابع کرنل چندجمله‌ای بوده و در مسایل ساده‌تر کارایی خوبی از خود نشان می‌دهد، ولی در مسایل پیچیده تابع کرنل چندجمله‌ای می‌تواند بسیار مفیدتر باشد. در تابع کرنل چندجمله‌ای پارامتر d باید از طرق آزمون و خطا یا از طریق الگوریتم‌های بهینه‌سازی محاسبه شود. تابع کرنل گوسی یا شعاعی (RBF) نوع دیگری از توابع کرنل بوده و معمولاً در مسائلی به کار برده می‌شود که هیچگونه شناختی از نوع و ماهیت داده‌ها در دسترس نباشد. برای اجرای مدل SVM در محیط نرم‌افزار MATLAB کدنویسی شد.

جدول ۲: توابع کرنل مورد استفاده در ماشین‌های بردار پشتیبان

نوع تابع	تابع کرنل
خطی	$K(X_i, X_j) = X_i^T \times X_j$
چند جمله‌ای	$K(X_i, X_j) = (\gamma X_i^T \times X_j + C)^d$
RBF	$K(X_i, X_j) = -\exp(-\gamma X_i - X_j ^2)$

معیارهای ارزیابی روش‌ها

به منظور ارزیابی عملکرد و دقت دو مدل LSTM و SVM برای پیش‌بینی سطح آب زیرزمینی از شاخص‌های آماری شامل $RMSE$ ، R^2 و معیار اطلاعاتی آکائیک (Akaike information criterion, AIC) استفاده شد که در ادامه روابط آن‌ها ارائه شده است.

$$R^2 = \frac{(\sum_{i=1}^n (G_{iobs} - \bar{G}_{iobs})(G_{ipre} - \bar{G}_{ipre}))^2}{\sum_{i=1}^n (G_{iobs} - \bar{G}_{iobs})^2 \sum_{i=1}^n (G_{ipre} - \bar{G}_{ipre})^2} \quad (8)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (G_{iobs} - G_{ipre})^2}{n}} \quad (9)$$

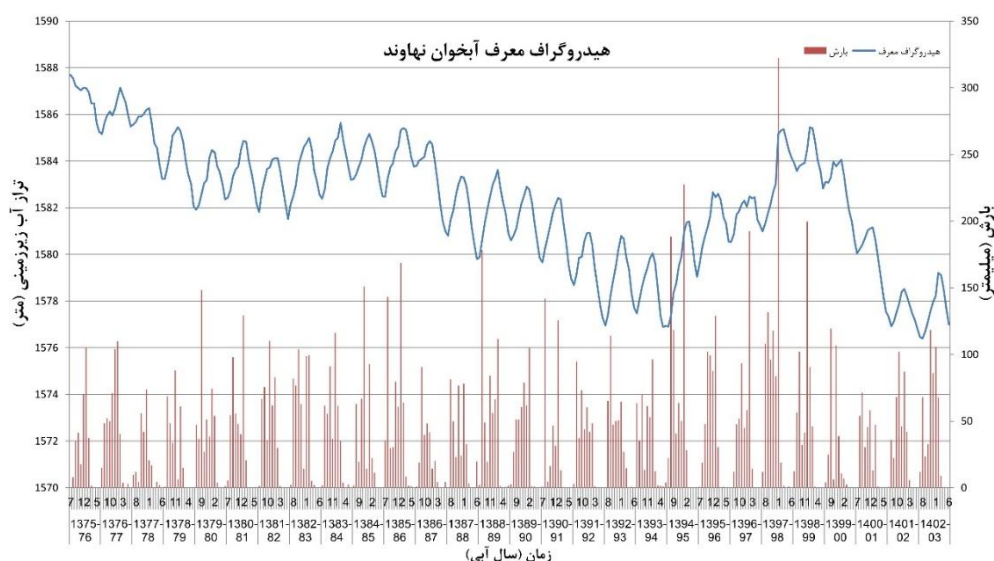
$$AIC = m \times \ln(RMSE) + 2(Npar) \quad (10)$$

که n تعداد داده‌ها، G_{iobs} : داده‌های مشاهداتی، $\bar{G}_{iobs}$: میانگین داده‌های مشاهداتی، G_{ipre} : داده‌های محاسباتی، $\bar{G}_{ipre}$: میانگین داده‌های محاسباتی، m : تعداد پارامترهای مدل و $Npar$: تعداد داده‌های آموزش دیده هستند. ضریب تعیین (R^2) میزان انطباق داده‌هایی که توسط مدل به وجود می‌آید و داده‌های واقعی را نشان می‌دهد و هر چقدر این ضریب به عدد یک نزدیکتر باشد، بیانگر انطباق بهتر و میزان خطای کمتر است؛ لذا در همه حالت‌ها برای بررسی میزان تأثیرگذاری هر پارامتر بر روی عملکرد مدل‌ها از این عامل استفاده شد. همچنین پارامتر $RMSE$ خطای داده‌های محاسباتی و مشاهداتی را بیان می‌کند. هر چه مقدار این عدد کمتر باشد، آموزش و شبیه‌سازی داده‌ها بهتر صورت گرفته است. در مورد معیار AIC نیز می‌توان گفت که هر چه معیار اطلاعاتی آکائیک مربوط به مدل کمتر باشد، آن مدل عملکرد بهتری دارد. مقدار کمتر این معیار، ناشی از دو عامل شامل خطای مدل و تعداد پارامترهاست، پس معیار خوبی برای بررسی مدل‌ها است.

نتایج

ترسیم هیدروگراف واحد آب زیرزمینی

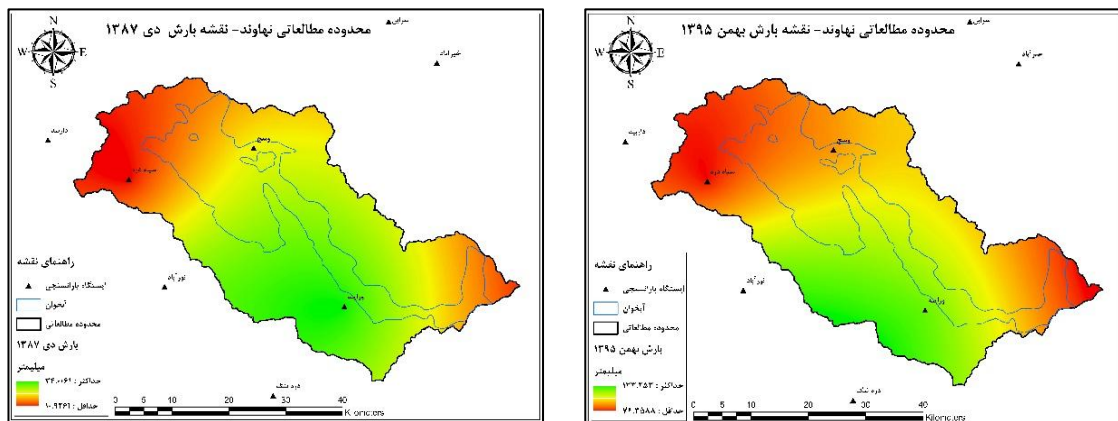
با ترسیم چندضلعی‌های تیسن برای چاه‌های مشاهداتی در محیط نرم‌افزار GIS و محاسبه مساحت هر چاه مشاهداتی، در نهایت نمودار هیدروگراف واحد آب زیرزمینی در مقیاس ماهانه مطابق شکل (۴) ترسیم شد. با توجه به شکل ۴ در زمان‌هایی که بارش در منطقه رخ داده، با تغذیه آبخوان سطح آب زیرزمینی افزایش یافته است. بعد از ترسیم هیدروگراف واحد آب زیرزمینی، مقادیر تراز سطح آب زیرزمینی به عنوان متغیر خروجی در دو مدل LSTM و SVM برای مدل سازی و پیش‌بینی مورد استفاده قرار می‌گیرند.



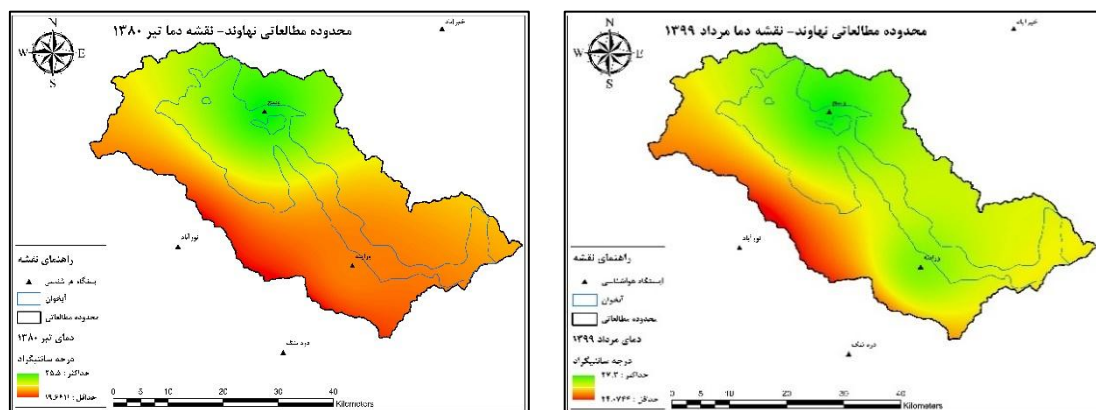
شکل ۴: هیدروگراف واحد آبخوان نهاوند و مقادیر بارش متناظر

ترسیم نقشه‌های هم‌بارش، هم‌دما و هم‌تبخیر

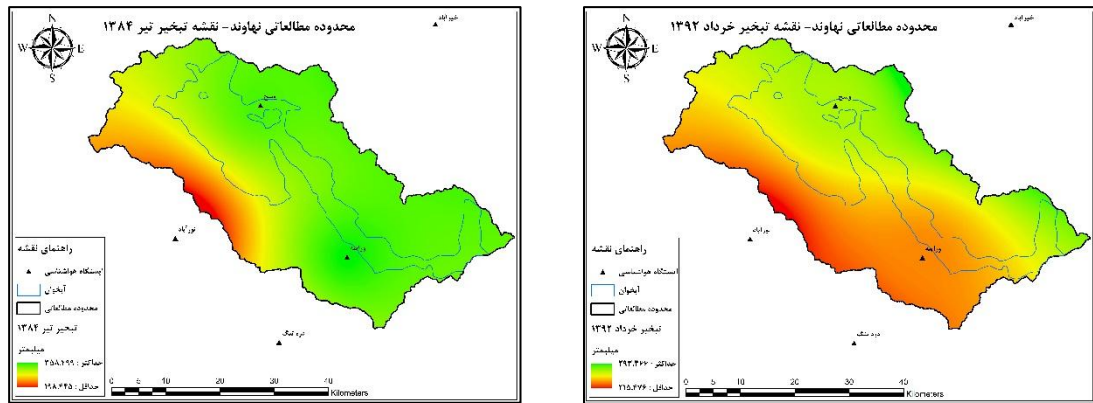
هدف از ترسیم نقشه‌های هم‌بارش، هم‌دما و هم‌تبخیر، برآورد مقادیر متوسط ماهانه بارش، دمای هوا و تبخیر در سطح آبخوان است که به عنوان متغیرهای ورودی در مدل‌های LSTM و SVM مورد استفاده قرار می‌گیرند. با توجه به این نقشه‌ها، مقدار متوسط بارش روی آبخوان در طی دوره آماری ۲۸ ساله (۱۳۷۵-۱۴۰۳) برابر ۴۶۳ میلی‌متر، در دوره ده ساله (۱۳۹۳-۱۴۰۳) برابر ۵۲۱ میلی‌متر و در سال آبی (۱۴۰۳-۱۴۰۲) برابر ۵۳۵ میلی‌متر می‌باشد. دلیل افزایش بارش در دهه اخیر مربوط به افزایش میزان بارش در سال‌های ۱۳۹۷ و ۱۳۹۸ نسبت به دیگر سال‌ها است. در ادامه برخی از نقشه‌ها به عنوان نمونه برای دو سال ارائه می‌شوند. در شکل (۵) نقشه هم‌بارش در ماه‌های مختلف در سال‌های آماری نشان داده شده است. در شکل ۵ بخش‌هایی از آبخوان که رنگ مشابه دارند، بارش یکسان دارند. با توجه به نقشه هم‌بارش بهمن ۱۳۹۵ و دی ۱۳۸۷، مقادیر بارش در بخش جنوبی آبخوان بیشتر از سایر مناطق است. همچنین مقادیر بارش در بخش شمال غربی و جنوب شرقی آبخوان کمتر از سایر مناطق است. براساس نقشه‌های هم‌دما ماهانه ترسیم شده، متوسط دمای سالانه در دوره ۲۸ ساله برابر ۱۳/۰۴ درجه سلسیوس، در دوره ده ساله (۱۳۹۳-۱۴۰۳) برابر ۱۴/۱۰ درجه سلسیوس و در سال آبی (۱۴۰۳-۱۴۰۲) برابر ۱۴/۵۴ درجه سلسیوس بدست آمد. نتایج نشان از افزایش دما در طی چند سال اخیر نسبت به دوره بلندمدت دارد. در شکل (۶) نقشه هم‌دما برای ماه‌های مختلف برای دو سال ارائه شده است. با توجه به نقشه هم‌دما آبخوان در تیرماه ۱۳۸۰ و مردادماه ۱۳۹۹ در شکل ۶، بخش‌های شمالی و جنوبی آبخوان به ترتیب از بیش‌ترین و کمترین دمای هوا برخوردار هستند. با توجه به نقشه‌های هم‌تبخیر ترسیم شده، متوسط مقدار سالانه تبخیر در دوره ۲۸ ساله (۱۳۷۵-۱۴۰۳) برابر ۲۰۴۷ میلی‌متر، در دوره ده ساله (۱۳۹۳-۱۴۰۳) برابر ۲۰۱۳ میلی‌متر و در سال آبی (۱۴۰۳-۱۴۰۲) برابر ۲۰۳۶ میلی‌متر می‌باشد. در شکل (۷) نقشه هم‌تبخیر برای ماه‌های مختلف برای دو سال ارائه شده است. با توجه به نقشه هم‌تبخیر در تیرماه ۱۳۸۴ و خردادماه ۱۳۹۲ در شکل ۷، بخش‌های شمالی و شرقی آبخوان بیشترین تبخیر دارند و در بخش جنوب غربی آبخوان، کمترین تبخیر صورت می‌گیرد.



شکل ۵: نقشه‌های هم‌بارش در محدوده مطالعاتی و آبخوان نهاوند



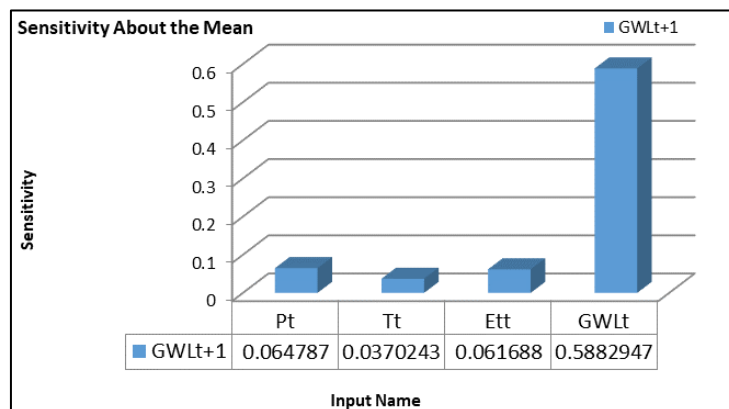
شکل ۶: نقشه‌های هم‌دما در محدوده مطالعاتی و آبخوان نهاوند



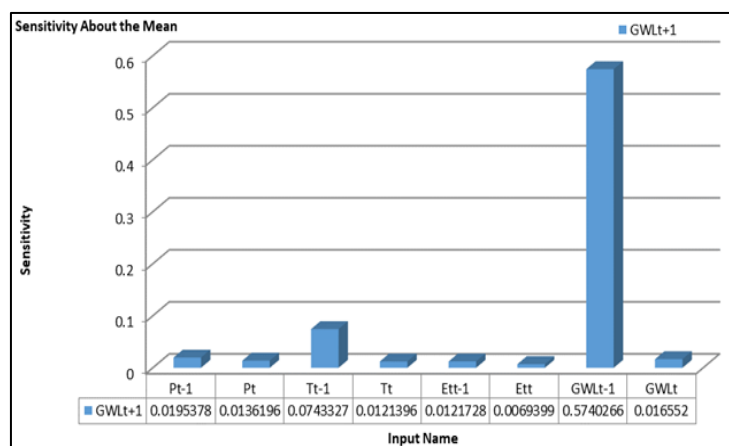
شکل ۷: نقشه‌های هم تبخیر در محدوده مطالعاتی و آبخوان نهاوند

آنالیز حساسیت تراز سطح آب زیرزمینی نسبت به تغییرات متغیرهای ورودی دو مدل LSTM و SVM

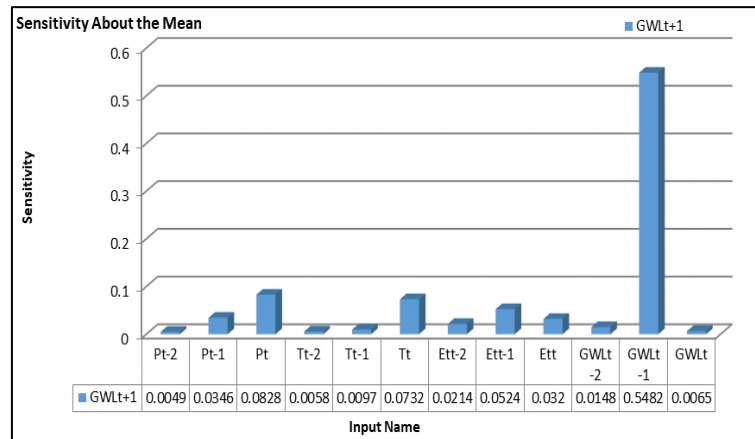
در شکل (۸) نتایج آنالیز حساسیت متغیرهای ورودی ارائه شده است. همان طوری که در این شکل مشاهده می‌شود، متغیر تراز سطح آب زیرزمینی در زمان فعلی (GWL_t) دارای بیشترین تأثیر بر روی متغیر خروجی است. بعد از این، متغیر مقدار بارش در زمان فعلی (P_t) دارای تأثیر بیشتری بر روی خروجی است. متغیر دمای هوا در زمان فعلی (T_t) کمترین تأثیر را روی خروجی داشته است. به دلیل اینکه بتوان از همه ظرفیت مربوطه استفاده نمود و مدل سازی با دقت بیشتری صورت گیرد، متغیرهای ورودی با گام‌های زمانی مختلف نیز در نظر گرفته شد. در شکل‌های (۹) و (۱۰) نتایج آنالیز حساسیت برای متغیرهای ورودی با یک و دو گام زمانی تأخیر در نظر گرفته شد.



شکل ۸: نتایج آنالیز حساسیت متغیرهای ورودی



شکل ۹: نتایج آنالیز حساسیت متغیرهای ورودی با یک گام زمانی تأخیر



شکل ۱۰: نتایج آنالیز حساسیت متغیرهای ورودی با دو گام زمانی تأخیر

براساس نتایج آنالیز حساسیت متغیرهای ورودی، ۵ ترکیب برای مدل‌سازی با مدل‌های هوشمند مطابق جدول (۳) در نظر گرفته شد. ابتدا از همه متغیرهای ورودی استفاده شد (ترکیب شماره یک). در ترکیب شماره دو، متغیرهای ورودی با یک گام زمانی تأخیر مورد استفاده قرار گرفت. در ترکیب شماره سه، متغیرهای ورودی با دو گام زمانی تأخیر در نظر گرفته شد، به دلیل اینکه در آنالیز حساسیت مولفه‌های دارای دو گام زمانی تأخیر، عمل‌کرد خوبی نداشته‌اند، از افزایش تأخیر به متغیرهای ورودی خودداری شد. در ترکیب‌های شماره چهار و پنج، از بین متغیرهای موثر در متغیر خروجی، مطابق آنالیز حساسیت، استفاده شد. در این جدول T_{t-1} ، T_{t-2} به ترتیب دمای متوسط با دو، یک گام تأخیر زمانی و T_t مقدار دمای متوسط در ماه جاری را نشان می‌دهد. ET_{t-1} ، ET_{t-2} به ترتیب تبخیر با دو، یک گام تأخیر زمانی و ET_t تبخیر در ماه جاری را نشان می‌دهد. GWL_{t-1} ، GWL_{t-2} به ترتیب تراز سطح آب زیرزمینی با دو، یک گام تأخیر زمانی و GWL_t تراز سطح آب زیرزمینی در ماه جاری را نشان می‌دهد. P_{t-1} ، P_{t-2} به ترتیب بارش با دو، یک گام تأخیر زمانی و P_t مقدار بارش در ماه جاری را نشان می‌دهد. GWL_{t+1} متغیر خروجی و تراز سطح آب زیرزمینی در ماه آتی می‌باشد.

جدول ۳: جزئیات ترکیبات ورودی‌های مختلف مدل LSTM و SVM

ترکیب	ورودی	خروجی
۱	T_t, ET_t, GWL_t, P_t	
۲	$T_{t-1}, T_t, ET_{t-1}, ET_t, GWL_{t-1}, GWL_t, P_{t-1}, P_t$	
۳	$T_{t-2}, T_{t-1}, T_t, ET_{t-2}, ET_{t-1}, ET_t, GWL_{t-2}, GWL_{t-1}, GWL_t, P_{t-2}, P_{t-1}, P_t$	GWL_{t+1}
۴	$T_{t-1}, P_t, P_{t-1}, GWL_t, GWL_{t-1}$	
۵	$T_t, ET_{t-1}, ET_t, GWL_{t-1}, P_{t-1}, P_t$	

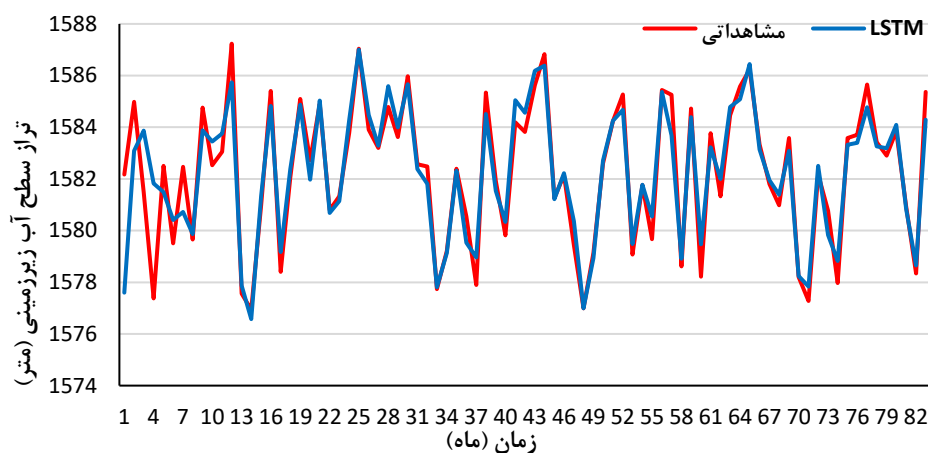
نتایج مدل حافظه‌ی کوتاه‌مدت طولانی (LSTM)

در این مطالعه، ۵ ترکیب، برای اجرای مدل LSTM مطابق جدول (۳) در نظر گرفته شد. برای این منظور در هر ترکیب، ساختارهای مختلفی مورد بررسی قرار گرفت. نتایج بهترین ساختار مدل LSTM در هر ترکیب در جدول (۴) ارائه شده است. همان‌طوری‌که در این جدول مشاهده می‌شود، ترکیب شماره سه با ۱۲ متغیر ورودی بهترین عمل‌کرد را داشته است. با توجه به جدول ۴ ضریب R^2 و RMSE برای دوره آموزش مدل LSTM به ترتیب ۰/۸۷ و ۰/۰۴۳ بدست آمد. در ترکیب شماره سه، حداکثر تکرار برابر ۱۰۰۰، Mini Batch برابر ۰/۰۳، Loss برابر ۰/۰۰۳۳ و Base Learning rate برابر ۰/۰۰۰۱ بوده است. خطای مدل از محاسبه اختلاف مقادیر تراز سطح آب زیرزمینی پیش‌بینی شده و مشاهداتی بدست آمد. اعتبارسنجی مدل با مقایسه نتایج تراز سطح آب زیرزمینی پیش‌بینی شده با مدل‌ها و تراز سطح آب زیرزمینی مشاهداتی برای دوره زمانی به غیر از آموزش مدل‌ها (دوره زمانی آزمون) انجام شده است. در شکل (۱۱) مقایسه بین عمل‌کرد مدل

LSTM با مقادیر مشاهداتی در مرحله آزمون صورت گرفته است. همان طوری که در این شکل دیده می شود، مدل LSTM در برخی از نقاط حداکثری و حداقلی عمل کرد مناسبی نداشته است.

جدول ۴: بهترین عمل کرد در ترکیب های مختلف مدل LSTM در مقیاس ماهانه

ترکیب	input size	max Epochs	پارامترهای مدل			R ²	RMSE
			Batch Mini RMSE	Mini Batch Loss	Base Learning rate		
۱	۴	۱۰۰۰	۰/۰۳	۰/۰۰۰۶۱	۰/۰۰۰۱	۰/۹۱	۰/۰۵۲
۲	۸	۱۰۰۰	۰/۰۴	۰/۰۰۰۸۴	۰/۰۰۰۱	۰/۸۷	۰/۰۵۳
۳	۱۲	۱۰۰۰	۰/۰۳	۰/۰۰۰۳۳	۰/۰۰۰۱	۰/۹۵	۰/۰۴۳
۴	۵	۱۰۰۰	۰/۰۳	۰/۰۰۰۵۷	۰/۰۰۰۱	۰/۹۰	۰/۰۴۷
۵	۶	۱۰۰۰	۰/۰۳	۰/۰۰۰۳۸	۰/۰۰۰۱	۰/۹۴	۰/۰۴۶



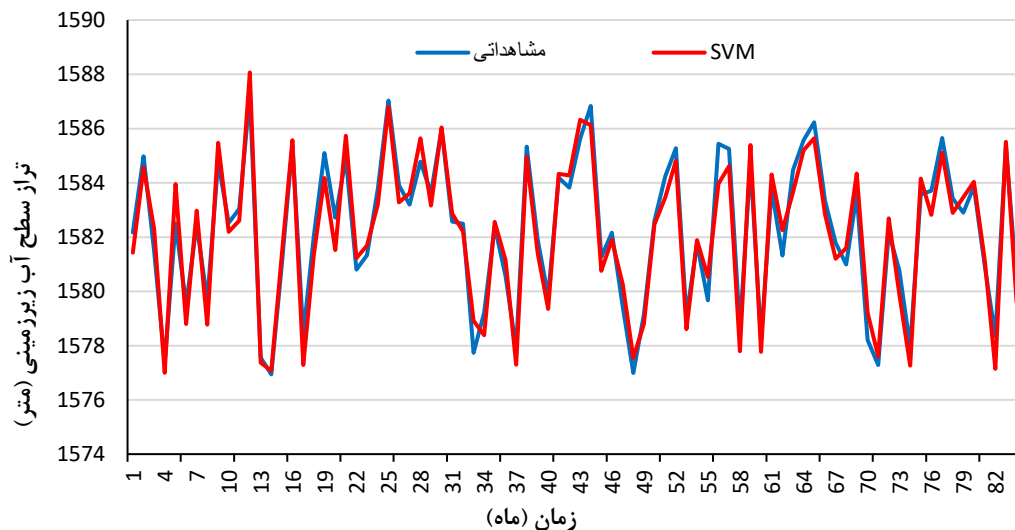
شکل ۱۱: مقایسه نتایج مدل LSTM با مقادیر مشاهداتی در مرحله آزمون مدل

نتایج مدل SVM

برای اجرای مدل SVM، با استفاده از تمامی پارامترهای موثر در اجرای مدل، اقدام به مدل سازی شد. برای این منظور مطابق جدول (۳)، پنج ترکیب مختلف مورد بررسی قرار گرفت. در هر ترکیب، توابع مختلف کرنل در مدل SVM در نظر گرفته شد. با بررسی نتایج ترکیب شماره سه با ۱۲ پارامتر بهترین عمل کرد را داشته است. بر این اساس با در نظر گرفتن ترکیب شماره سه به عنوان متغیرهای ورودی در مدل SVM، ساختار بهینه مدل به ازاء توابع کرنل مختلف حاصل شد. مقادیر شاخص های آماری و مقادیر بهینه پارامترها مربوط به توابع کرنل مختلف در مدل SVM برای پیش بینی تراز سطح آب زیرزمینی به ازاء دوره زمانی آموزش و آزمون در جدول ۵ نشان داده شده است. با توجه به جدول ۵، تابع کرنل پایه شعاعی در دوره زمانی آزمون مدل SVM دارای کمترین RMSE و بیشترین R² می باشد. در دوره زمانی آزمون مدل SVM برای تابع کرنل پایه شعاعی، مقادیر ضریب R² و RMSE به ترتیب ۰/۹۶ و ۰/۲۳ بدست آمد. به عبارتی تابع کرنل پایه شعاعی کارایی بهتری نسبت به تابع چندجمله ای و خطی در پیش بینی تراز سطح آب زیرزمینی دارد. همچنین با توجه به جدول ۵ ضریب R² و RMSE برای دوره آموزش مدل SVM به ترتیب ۰/۹۶ و ۰/۲۳ بدست آمد. شکل ۱۲ مقایسه تغییرات تراز سطح آب زیرزمینی مشاهداتی و پیش بینی شده با مدل SVM با تابع کرنل پایه شعاعی برای دوره زمانی آزمون مدل را نشان می دهد. با توجه به شکل ۱۲ می توان دریافت که مدل SVM از لحاظ پیش بینی تراز سطح آب زیرزمینی بخوبی آموزش دیده است.

جدول ۵: مقادیر شاخص‌های آماری توابع کرنل در مدل SVM به ازاء ترکیب شماره سه برای دوره‌های زمانی آموزش و آزمون مدل

تابع کرنل	آموزش		آزمون		مقادیر بهینه پارامترها			
	RMSE	R ²	RMSE	R ²	γ	σ^2	d	c
چند جمله‌ای	۰/۰۳۸	۰/۸۴	۰/۰۵۷	۰/۷۲	۶۸	-	۳	-۱
پایه شعاعی	۰/۰۲۳	۰/۹۶	۰/۰۴۲	۰/۸۹	۹	۰/۷	-	-
خطی	۰/۰۵۵	۰/۶۲	۰/۰۷۳	۰/۵۷	۷	-	-	-



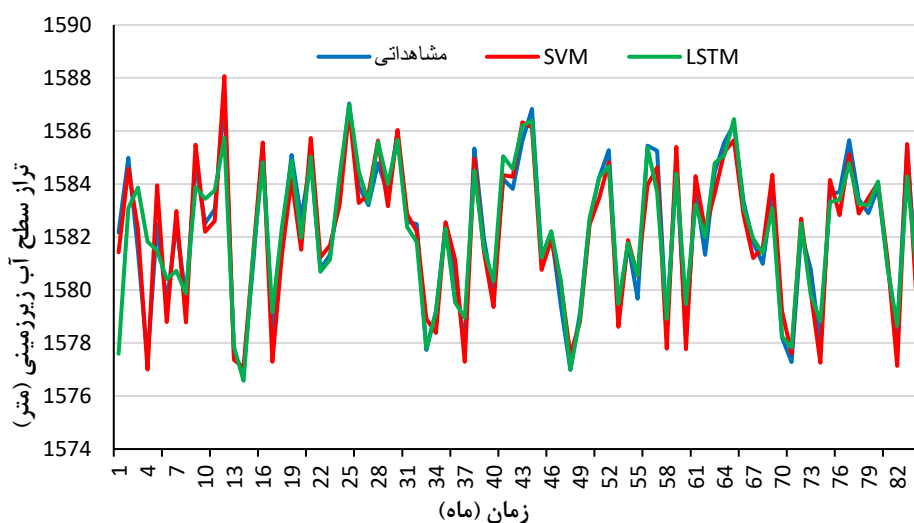
شکل ۱۲: مقایسه تغییرات تراز سطح آب زیرزمینی مشاهداتی و شبیه‌سازی‌شده با بهترین مدل SVM برای دوره زمانی آزمون

مقایسه نتایج مدل‌ها

در این بخش به مقایسه نتایج حاصل از اجرای دو مدل پرداخته می‌شود. در جدول (۶) بهترین عمل کرد دو مدل ارائه شده است. با مقایسه نتایج مشاهده می‌شود که مدل SVM با ضریب تعیین ۰/۸۹ و خطای ۰/۰۴۲ در مرحله آزمون مدل، دارای بهترین عمل کرد بوده است. این مدل دارای معیار اطلاعاتی آکائیک ۱۵۸/۵۷ می‌باشد که نسبت به مدل دیگر کمتر است. در شکل (۱۳) نتایج پیش‌بینی سطح آب زیرزمینی دو مدل در دوره آزمون براساس هیدروگراف‌های واحد آبخوان نهاوند مورد مقایسه قرار گرفتند. همانطوری که مشاهده می‌شود مدل SVM عملکرد بهتری داشته و مقادیر پیش‌بینی سطح آب زیرزمینی به مقادیر مشاهداتی نزدیک‌تر است. لذا می‌توان از این مدل برای پیش‌بینی تراز سطح آب زیرزمینی آبخوان نهاوند استفاده نمود. نتایج این پژوهش نشان داد که مدل SVM با استفاده از تابع کرنل پایه شعاعی (RBF) در پیش‌بینی تراز سطح آب زیرزمینی عمل کرد بهتری نسبت به مدل LSTM دارد. این نتیجه با یافته‌های مطالعات متعددی مانند Behzad و همکاران (2010) و Zhou و همکاران (2017) همخوانی دارد.

جدول ۶: مقایسه شاخص‌های آماری دو مدل SVM و LSTM

مدل	R ²		RMSE		AIC	آزمون
	آموزش	آزمون	آموزش	آزمون	آموزش	
LSTM	۰/۹۵	۰/۸۷	۰/۰۲۶	۰/۰۴۴	۴۹۱	۱۵۸/۵۹
SVM	۰/۹۶	۰/۸۹	۰/۰۲۳	۰/۰۴۲	۴۹۰/۸	۱۵۸/۵۷



شکل ۱۳: مقایسه مدل LSTM و SVM برای پیش بینی سطح آب زیرزمینی در مرحله آزمون

بحث و نتیجه گیری

در این تحقیق با استفاده از مدل های هوش مصنوعی شامل LSTM و SVM، تراز سطح آب زیرزمینی آبخوان نهاوند پیش بینی شد. آنالیز حساسیت تراز سطح آب زیرزمینی یک ماه آینده نسبت به تغییرات متغیرهای ورودی دو مدل LSTM و SVM نشان داد که مقادیر بارش، دمای هوا، تبخیر و تعرق و تراز سطح آب زیرزمینی تا دو ماه قبل، متغیرهای موثر ورودی هستند. نتایج مدلسازی با مدل SVM نشان داد که تابع کرنل پایه شعاعی در مدل SVM، بهترین تابع کرنل است. در دوره زمانی آزمون برای تابع کرنل پایه شعاعی، RMSE و R^2 به ترتیب ۰/۰۴۲ و ۰/۸۹ بدست آمد. به عبارتی تابع کرنل پایه شعاعی در پیش بینی تراز سطح آب زیرزمینی، کارایی بهتری نسبت به تابع چندجمله ای و خطی دارد. نتایج مدلسازی نشان داد که مدل LSTM در برخی از نقاط حداکثری و حداقلی عمل کرد مناسبی نداشته است. مقایسه نتایج دو مدل نشان می دهد که مدل SVM دارای عملکرد بهتری نسبت به مدل LSTM برخوردار است. مدل SVM با ضریب تعیین ۰/۸۹ و خطای ۰/۰۴۲ در مرحله آزمون مدل، دارای بهترین عملکرد بوده است. این مدل دارای معیار AIC برابر با ۱۵۸/۵۷ بود و نسبت به مدل دیگر مقدار کمتر دارد. به طور کلی می توان گفت که عمل کرد مدل SVM برای مدلسازی تراز سطح آب زیرزمینی مناسب بوده و قابلیت این را دارد که از این مدل برای پیش بینی تراز آب زیرزمینی در دوره های آتی استفاده کرد. برتری مدل SVM نسبت به LSTM را می توان به تولدایی بالای SVM در مدلسازی روابط غیرخطی با حجم داده های محدود نسبت داد. از آنجا که مدل LSTM برای استخراج الگوهای زمانی پیچیده به داده های آموزشی فراوان نیاز دارد، در مجموعه داده های نسبتاً کوچک هیدرولوژیکی ممکن است نتواند تمام ظرفیت خود را نشان دهد. در مقابل، مدل SVM با استفاده از کرنل پایه شعاعی توانسته است روابط غیرخطی بین متغیرهای ورودی و تراز سطح آب زیرزمینی را با دقت بیشتری شناسایی کرده و به دلیل مقاومت بیشتر در برابر بیش برزش و نویز داده ها، عملکرد بهتری نسبت به مدل LSTM ارائه دهد.

سپاسگزاری

نویسندگان این تحقیق، از شرکت آب منطقه ای همدان به دلیل در اختیار قرار دادن اطلاعات مورد نیاز، تقدیر و تشکر می کنند.

منابع

- پناهی، م.، میثاقی، ف. و عسگری، پ.، ۱۳۹۷. شبیه‌سازی و تخمین نوسانات سطح آب زیرزمینی با استفاده از GMS. فصلنامه علمی و پژوهشی علوم محیطی، ۱۶(۱)، ۱-۱۴.
- جمالی‌زاده، م.ع.، بذرافشان، ا.ا.، مهدوی نجف‌آبادی، ر.، آذره، ع. و رفیعی ساردوئی، ا.، ۱۳۹۹. پیش‌بینی نوسانات سطح آب زیرزمینی با استفاده از مدل‌های سری زمانی و GMS (مطالعه موردی: دشت رفسنجان). اکوهیدرولوژی، ۷(۱)، ۹۷-۱۰۹.
- خطیبی‌نیا، ک.، ۱۳۹۴. پیش‌بینی نوسانات سطح آب زیرزمینی در آبخوان بهاباد با استفاده از مدل عددی. پایان‌نامه کارشناسی ارشد، گروه علوم زمین، دانشگاه صنعتی شاهرود.
- سلامتیان، س.ا.، ابراری، ح. و نظری، ع.، ۱۴۰۲. پیش‌بینی تراز آب زیرزمینی با استفاده از مدل استنتاج تطبیقی عصبی فازی (مطالعه موردی: دشت قم). نشریه مهندسی آبیاری و آب ایران، ۱۳(۵۱)، ۳۰۴-۳۸۵.
- سلگی، ا.، ۱۳۹۳. پیش‌بینی جریان رودخانه با مدل ترکیبی شبکه عصبی مصنوعی-موجک و مقایسه آن با روش‌های سیستم استنتاج فازی عصبی تطبیقی و شبکه عصبی مصنوعی (مطالعه موردی: رودخانه گاماسیاب نهاوند). پایان‌نامه کارشناسی ارشد رشته مهندسی منابع آب، دانشکده مهندسی علوم آب، دانشگاه شهید چمران.
- سلیمان‌پور، س.م.، زندی‌فر، س.، رحمتی، ا. و معتمدنیا، م.، ۱۴۰۴. ارزیابی روند تغییرات دوره‌ای سطح آب زیرزمینی حوزه آبخیز بهادران استان یزد. مدل‌سازی و مدیریت آب و خاک، ۱۵(۱)، ۲۹-۴۴.
- ندیری، ع.ا.، نادری، ک.، اصغری‌مقدم، ا. و حبیبی، م.ح.، ۱۳۹۵. پیش‌بینی زمانی و مکانی سطح آب زیرزمینی با استفاده از روش‌های هوش مصنوعی و زمین‌آمار (مطالعه موردی: آبخوان دشت دوزدوزان). جغرافیا و برنامه‌ریزی، ۲۰(۵۸)، ۲۸۱-۳۰۱.
- یاری، ر.ا. و درزی‌نفت‌چالی، ع.، ۱۳۹۶. پیش‌بینی نوسانات سطح آب زیرزمینی تحت سناریوهای مختلف مدیریتی با استفاده از مدل MODFLOW. آبیاری و آب ایران، ۸(۲)، ۱۰۳-۱۱۵.

Aderemi, B.A., Olwal, T.O., Julius, M., Ndambuki, J.M. and 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).

Bahmani, R. and Ouarda, T.B.M.J., 2020. Groundwater level modeling with hybrid artificial intelligence techniques. *Journal of Hydrology*, 595, 1-12.

Behzad, M., Asghari, K., Coppola, E.A. Jr., 2010. Comparative study of SVMs and ANNs in aquifer water level prediction. *Journal of Computing in Civil Engineering*, 24(5), 408-413.

Havril, T., Tóth, Á., Molson, J.W., Galsa, A. and Mádl-Szőnyi, J., 2018. Impacts of predicted climate change on groundwater flow systems: Can wetlands disappear due to recharge reduction? *Journal of Hydrology*, 563, 1169-1180.

Hochreiter, S. and Schmidhuber, J., 1997. Long short-term memory. *Neural Computation*, 9(8), 1735-1780.

Kayhomayoon, Z., Azar, N.A., Milan, S.G., Moghaddam, H.K. and Berndtsson, R., 2021. Novel approach for predicting groundwater storage loss using machine learning. *Journal of Environmental Management*, 296 (15), 113237.

Konikow, L.F. and Kendy, E., 2005. Groundwater Depletion: A Global Problem. *Hydrogeology Journal*, 13, 317-320.

Mirajabi, A., Nassery, H.R., Nakhaei, M., Adamowski, J., Akbarzadeh, A.H. and Alijani, F., 2019. Evaluation of data-driven models (SVR and ANN) for groundwater-level prediction in confined and unconfined systems. *Environmental Earth Sciences*, 478(489).

Moustadraf, J., Razack, M. and Sinan, M., 2003. Evaluation of the impacts of climate changes on the coastal Chaouia aquifer, Morocco, using numerical modeling. *Hydrogeology Journal*, 16(7), 1411-1426.

Mujahid, I., Usman, A.N., Afaq, A., Habib-ur, R., Usman, G. and Tallat, F., 2020. Relating groundwater levels with meteorological parameters using ANN technique. *Measurement*, 166(108163).

Mukul Fishman, R., Siegfried, R.T., Raj, P., Modi, V. and Lall, U., 2011. Over-extraction from shallow bedrock versus deep alluvial aquifers: Reliability versus sustainability considerations for India's groundwater irrigation. *Water Resources Research*, 47(6).

Sattari, M.T., Mirabbasi, R., Shamsi Sushab, R. and Abraham, J., 2017. Prediction of Groundwater Level in Ardebil Plain Using Support Vector Regression and M5 Tree Model. *National GroundWater Association*, 56(4), 636-646.

Shahbazi, M., Zarei, H. and Solgi, A., 2023. De-noising groundwater level modeling using data decomposition techniques in combination with artificial intelligence (case study Aspas aquifer). *Applied Water Science*, 13(88), 1-18.

Shin, K.S., Lee, T.S. and Kim, H.J., 2005. An application of support vector machines in bankruptcy prediction model. *Expert Systems with Applications*, 28(1), 127-135.

Sophocleous, M., 2000. From Safe Yield to Sustainable Development of Water Resources-The Kansas Experience. *Journal of Hydrology*, 235, 27-43.

Van Loon, A.F. and Van Lanen, H.A.J., 2013. Making the distinction between water scarcity and drought using an observation-modeling framework. *Water Resources Research*, 49, 1483-1502.

Zhou, T., Wang, F. and Yang, Z., 2017. Comparative analysis of ANN and SVM models combined with wavelet preprocess for groundwater depth prediction. *Water*, 9(10), 781.

Monthly Groundwater Level Prediction in the Nahavand Aquifer Using Deep Learning and Support Vector Machine (SVM) Models

Somayeh Abdi¹
Hossein Fathian^{2*}
Mehdi Asadi Lour¹
Aslan Egdernezhad¹
Ali Asareh¹

1. Department of Irrigation and Drainage, Ahv.C., Islamic Azad University, Ahvaz, Iran.

2. Department of Water Resources Engineering, Ahv.C., Islamic Azad University, Ahvaz, Iran.

*Corresponding author:
Fathian.h@iau.ac.ir

Received date: December/30/2025

Accepted date: June/26/2026

Abstract

Reduced rainfall and excessive groundwater exploitation have caused a significant decline in groundwater levels in many parts of the world. One of the aquifers affected by a drop in groundwater level is the Nahavand aquifer in Hamedan Province, western Iran. In this study, the groundwater level was modeled using Long Short-Term Memory (LSTM) and Support Vector Machine (SVM) methods. Monthly precipitation, temperature, and evaporation maps were prepared using ArcGIS software. Based on the information from observation wells and the completion of existing well data, the unit hydrograph of the aquifer was drawn. Monthly data from the statistical period 1996 to 2024 were used for training and testing the models. Sensitivity analysis of the groundwater level one month ahead with respect to changes in the input variables of the two models, LSTM and SVM, showed that precipitation, air temperature, groundwater and evapotranspiration level up to two months earlier were the effective input variables. The modeling results with the SVM model showed that the radial basis function kernel was the best kernel compared with the linear and polynomial kernels. Comparison of the results of the intelligent models showed that the SVM model performed better than the LSTM model. The results showed that, for the test period, the LSTM model achieved an R^2 of 0.87 and an RMSE of 0.043. Also, for the test period, the SVM model achieved an R^2 of 0.89 and an RMSE of 0.042. Overall, it can be said that the SVM model performs very well for predicting the groundwater level of the Nahavand aquifer and can also be used to investigate groundwater levels in other aquifers.

Keywords: Prediction, Groundwater level, Nahavand aquifer, LSTM, SVM.