



Comparative Evaluation of Decision Tree (M5) and Least Square Support Vector Machine (LS-SVM) Models for Groundwater Level Prediction in the Mashhad Plain

Vajihe Ramezani Saani,^{1*}  Mahdi Zarei,²  Seyed Morteza Seyedian³ 

1. Hakim Sabzevari University. Faculty of Engineering. Sabzevar, Iran. Email: Ramezani.va@gmail.com

2. Associate Professor. Research Center of Social Studies & Geographical Sciences. Hakim Sabzevari University. Sabzevar, Iran. Email: m.zarei@hsu.ac.ir

3. Associate Professor of Watershed Management. Faculty of Agriculture and Natural Resources. Gonbad Kavous University, Iran. Email: s.m.seyedian@gmail.com

Article Info

Article type:
Research Article

Article history:
Received January 15 2025
Received in revised form 05 February 2025
Accepted 15 March 2025
Available online 21 March 2025

Keywords:
Groundwater level,
M5,
LS-SVM,
Mashhad plain,
Scenario.

ABSTRACT

Objective: Predicting the groundwater table level is considered one of the fundamental steps in the optimal management of water resources in arid and semi-arid regions. Nowadays, the application of intelligent models for estimating groundwater levels is increasing due to their ease of use and high accuracy in estimating complex and nonlinear mathematical equations. The aim of the present study is to estimate the groundwater table level of the Mashhad plain aquifer using the decision tree model (M5) and to compare it with the least squares support vector machine model (LS-SVM) under 10 different scenarios.

Method: For this purpose, monthly climatic data (precipitation, evaporation, and temperature) and groundwater level information from 60 piezometric wells over a 10-year statistical period were utilized, and the employed models were evaluated using statistics such as the coefficient of determination (R^2), RMSE, and MBE.

Results: The results of the LS-SVM model indicated that the highest simulation accuracy belonged to scenario 4, followed by scenario 9, while the other scenarios exhibited very low accuracy in simulating the water level. The MBE error values in scenarios 4 (-0/151) and 9 (-0/018) showed that the model simulated the groundwater level lower than reality. Based on the results of the water level simulation using the decision tree model, all scenarios were acceptable, and the scenarios 4 and 5 having the highest and lowest accuracy with coefficient of determination of 0/999 and 0/86, respectively. Overall, in both models used, scenario 4 simulated the groundwater level with almost similar accuracy. A comparison of the results of the models indicated that the LS-SVM model is more sensitive to changes in input parameters than the M5 model, such that the decision tree model, unlike the least squares support vector machine model, provided acceptable results in all scenarios.

Conclusions: In summary, the comparison of the models used suggests that the appropriate selection of climatic parameters and the examination and analysis of data have a significant impact on the accuracy of predictions.

Cite this article: Ramezani Saani, V., Zarei, M., Seyedian, S. M. (2025). Comparative Evaluation of Decision Tree (M5) and Least Square Support Vector Machine (LS-SVM) Models for Groundwater Level Prediction in the Mashhad Plain. *ECOHYDROLOGY*, 12 (1), 581-594. <https://doi.org/10.22059/IJE.2025.390138.1863>



© Vajihe Ramezani Saani,, Mahdi Zarei,, Seyed Morteza Seyedian.

Publisher: University of Tehran Press.

DOI: <https://doi.org/10.22059/IJE.2025.390138.1863>

Introduction

Groundwater resources are considered a reliable source in various regions, especially in arid and semi-arid areas, due to their sweetness, stable chemical compositions, constant temperature, lower pollution coefficients, and higher levels of reliability. On the other hand, by influencing the ecological capacity of the land, they are an important and effective phenomenon in economic development, ecological diversity, and community health. Various methods based on physical, conceptual, statistical, and numerical models have been used to predict groundwater levels, which often face limitations due to the large volume of data and input parameters, leading to reduced accuracy in predictions. Given the simplicity and acceptable results of artificial intelligence-based approaches, these tools are now widely used for assessing and modeling groundwater levels.

Methodology

In the present study, data from observation wells, rainfall, evaporation, and temperature from 10 years were used to predict the groundwater level in the Mashhad plain. The influential meteorological parameters affecting the water level, such as monthly rainfall, rainfall from one and two months prior, monthly evaporation, evaporation from the previous month, monthly temperature, temperature from one month prior, as well as the monthly water level and the water level from one month prior, were considered as inputs for the models used in the study in the form of 10 different scenarios. To utilize the aforementioned data in the Least Square support vector machine (LS-SVM) and decision tree (M5) models, the data was first divided into two categories (75% for training and 25% for testing), and then the specified algorithms were executed under the 10 desired scenarios in the Weka software.

Results

The analysis of various scenarios in the Least Square Support Vector Machine (LS-SVM) indicates that this model has only provided acceptable results in scenarios 4 and 9, while the error rates in other scenarios are significantly high. The impact of different inputs on the LS-SVM has been substantial, such that adding or removing an input led to considerable changes in the error rate. The LS-SVM model has not been able to provide a suitable simulation for the water level in the studied area with many combinations of inputs. Overall, the examination of various scenarios in the Decision Tree model (M5) shows acceptable accuracy across all input scenarios in simulating water levels. The dispersion of observed and simulated water level data in all scenarios falls within a range of $\pm 25\%$, although the accuracy of the scenarios varies. The best results from the Decision Tree model correspond to scenario 4, followed by scenario 7, while the weakest simulation is related to scenario 5.

Conclusion

Comparing different scenarios in decision tree models and support vector machine least squares models shows that both models provide the best simulation of the water level in the studied area with scenario 4. On the other hand, examining other scenarios indicates that adding or removing input parameters from the combination of parameters in scenario 4 will lead to an increase in the model's error. Overall, the sensitivity of the M5 model to changes in input parameters is much lower than that of the LS-SVM model, such that the M5 model has provided acceptable results in all combinations, while the error in some scenarios of the LS-SVM model is very high.

Author Contributions

All authors made equal contributions to the conceptualization, software development, methodology, validation, analysis, data management, drafting the original manuscript, reviewing and editing the manuscript.

Data Availability Statement

The observational rainfall data is from Regional Water Company of Khorasan Razavi <https://data.wrm.ir/>.

Acknowledgements

We would like to thank the Regional Water Company of Khorasan Razavi for providing the observational data of the study area.

Ethical considerations

The authors adhered to ethical research practices, avoiding data fabrication and falsification.

Funding

No funding was received for this research.

Conflict of interest

The authors declare no conflict of interest.

ارزیابی مقایسه‌ای مدل‌های درخت تصمیم (M5) و حداقل مربعات ماشین بردار پشتیبان (LS_SVM) در پیش‌بینی سطح آب زیرزمینی دشت مشهد

وجیهه رمضانی ثانی،^۱ مهدی زارعی،^۲ سید مرتضی سیدیان^۳ 

۱. مدرس دانشگاه حکیم سبزواری، دانشکده فنی و مهندسی، گروه عمران، ایران. رایانامه: Ramezani.va@gmail.com
۲. دانشیار، مرکز پژوهشی علوم جغرافیایی و مطالعات اجتماعی، دانشگاه حکیم سبزواری، ایران. رایانامه: m.zarei@hsu.ac.ir
۳. دانشیار، گروه مهندسی آبخیزداری، دانشکده کشاورزی دانشگاه گنبد کاووس، ایران. رایانامه: s.m.seyedian@gmail.com

اطلاعات مقاله	چکیده
نوع مقاله: مقاله پژوهشی تاریخ دریافت: ۱۴۰۳/۱۰/۲۶ تاریخ بازنگری: ۱۴۰۳/۱۱/۱۷ تاریخ پذیرش: ۱۴۰۳/۱۲/۲۵ تاریخ انتشار: ۱۴۰۴/۰۱/۰۱ کلیدواژه‌ها: سطح آب زیرزمینی، درخت تصمیم، حداقل مربعات ماشین بردار پشتیبان، دشت مشهد، سناریو.	موضوع: پیش‌بینی سطح ایستابی سفره آب زیرزمینی یکی از گام‌های اساسی مدیریت بهینه منابع آب در مناطق خشک و نیمه‌خشک محسوب می‌شود. امروزه کاربرد مدل‌های هوش مصنوعی در تخمین تراز آب زیرزمینی به دلیل سهولت استفاده و دقت بالای آن‌ها در برآورد معادله‌های پیچیده و غیرخطی ریاضی رو به افزایش است. هدف: مطالعه حاضر با هدف برآورد سطح آب آبخوان دشت مشهد با استفاده از مدل درخت تصمیم (M5) و مقایسه آن با مدل حداقل مربعات ماشین بردار پشتیبان (LS-SVM) تحت ۱۰ سناریو انجام شده است. روش تحقیق: بدین منظور از داده‌های اقلیمی ماهانه (بارش، تبخیر و دما) و اطلاعات سطح آب زیرزمینی ۶۰ چاه پیژومتری در دوره آماری ۱۰ ساله استفاده و مدل‌های به کاررفته با استفاده از آماره‌هایی همچون ضریب تعیین (R^2)، RMSE و MBE مورد ارزیابی قرار گرفت. یافته‌ها: نتایج مدل LS-SVM نشان داد بیشترین دقت شبیه‌سازی متعلق به سناریوی ۴ و سپس ۹ بوده و سایر سناریوهای مورد استفاده دقت بسیار پایینی در شبیه‌سازی سطح آب دارند. مقدار خطای MBE در سناریوهای ۴ و ۹ ($-۰/۱۵۱$) و ($-۰/۰۱۸$) نشان داد مدل سطح آب زیرزمینی را کمتر از واقعیت شبیه‌سازی نموده است. براساس نتایج شبیه‌سازی سطح آب توسط مدل درخت تصمیم در تمام سناریوها قابل قبول بوده و سناریوهای ۴ و ۵ با ضریب تعیین $۰/۹۹۹$ و $۰/۹۸۶$، به ترتیب بیشترین و کمترین دقت را دارند. به طور کلی در هر دو مدل مورد استفاده، سناریوی ۴ سطح آب زیرزمینی را با دقت تقریباً مشابهی شبیه‌سازی نموده است. مقایسه نتایج مدل‌ها حاکی از حساسیت بیشتر مدل LS-SVM نسبت به مدل M5 در برابر تغییرات پارامترهای ورودی است، به طوری که مدل درخت تصمیم برخلاف مدل حداقل مربعات ماشین بردار پشتیبان در تمام سناریوها نتایج قابل قبولی دارد. نتیجه‌گیری: در مجموع مقایسه مدل‌های مورد استفاده حاکی از این است که انتخاب مناسب پارامترهای اقلیمی و بررسی و تحلیل داده‌ها تأثیر قابل توجهی بر دقت پیش‌بینی‌ها دارد.

استناد: رمضانی ثانی، وجیهه؛ زارعی، مهدی؛ سیدیان، سید مرتضی. ارزیابی مقایسه‌ای مدل‌های درخت تصمیم (M5) و حداقل مربعات ماشین بردار پشتیبان (LS_SVM) در پیش‌بینی سطح آب زیرزمینی دشت مشهد. *اکوهیدرولوژی*، ۱۲(۱)، ۵۸۱-۵۹۴.

<https://doi.org/10.22059/IJE.2025.390138.1863>



© وجیهه رمضانی ثانی، مهدی زارعی، سید مرتضی سیدیان.

ناشر: انتشارات دانشگاه تهران.

مقدمه

منابع آب زیرزمینی از یک سو به دلیل ترکیبات ثابت شیمیایی، دمای ثابت، ضریب آلودگی کمتر و سطح اطمینان بالاتر به عنوان یک منبع قابل اتکا در مناطق مختلف به ویژه مناطق خشک و نیمه خشک محسوب شده و از سوی دیگر با تأثیر بر توان اکولوژیک سرزمین، یک پدیده مهم و مؤثر در توسعه اقتصادی، تنوع اکولوژیکی و سلامت جامعه به حساب می آید (فلاحزاده و همکاران، ۱۳۹۸). در مناطق نیمه خشک با مدیریت صحیح این ذخایر ارزشمند، پیش بینی تراز آب زیرزمینی در دشت ها برای برنامه ریزی های بلندمدت و استفاده بهتر و بهینه تر از پتانسیل های آبی موجود ضرورت بیشتری دارد. بدین منظور مدل کردن سفره های آب زیرزمینی و به تبع آن تخمین سطح ایستابی از نظر ایجاد سازه های مهندسی، مطالعات هیدرولوژی و مدیریتی، شرب، مصارف کشاورزی و به دست آوردن آب های زیرزمینی با کیفیت مناسب با استفاده از شبیه سازی ریاضی و کامپیوتری به عنوان ابزاری مؤثر در بهره برداری بهینه از منابع آب زیرزمینی، اهمیت ویژه ای پیدا کرده است (میثاقی و محمدی، ۱۳۸۵). اکثر آبخوان های کشور به دلیل برداشت بی رویه از منابع آب زیرزمینی در مناطق مختلف به ویژه مناطق خشک و نیمه خشک به طور قابل توجهی با کسری مخزن مواجه شده و سطح آب زیرزمینی در آن ها با سرعت زیادی رو به کاهش است. تاکنون روش های گوناگونی مبتنی بر مدل های فیزیکی، مفهومی، آماری و عددی برای پیش بینی سطح ایستابی مورد استفاده قرار گرفته که اغلب با محدودیت حجم زیاد داده و پارامترهای ورودی روبه رو هستند که منجر به کاهش دقت در پیش بینی ها خواهد شد (رستمی نژاد دولت آباد، شهابی و مددی، ۱۴۰۲). با توجه به سادگی و نتایج قابل قبول رویکردهای مبتنی بر هوش مصنوعی، امروزه از این ابزارها به صورت گسترده برای ارزیابی و مدل سازی سطح آب زیرزمینی استفاده می شود (رجایی و ابراه، ۱۳۹۴). از جمله این روش ها می توان به شبکه عصبی مصنوعی، انواع مدل های یادگیری ماشین، رگرسیون بردار پشتیبان و ماشین بردار پشتیبان اشاره کرد که به دلیل عملکرد مناسب آن ها در مطالعات مختلف مدیریت منابع آب، تحلیل ریسک و عدم قطعیت و همچنین مطالعات هیدرولوژیکی مورد استفاده قرار می گیرند (میلان^۱ و همکاران، ۲۰۱۸).

بررسی مطالعات متعدد حاکی از عملکرد مناسب روش های مبتنی بر هوش مصنوعی در شبیه سازی سطح آب زیرزمینی است که با پیشرفت روزافزون علوم مرتبط سعی بر برطرف نمودن موانع و ضعف های احتمالی روش های مذکور و ارائه الگوریتم های جدید در مطالعات مختلف شده است. ابارشی و همکاران (۱۳۹۳) طی پژوهشی در دشت زرین گل به مقایسه عملکرد سه روش هوش مصنوعی برای پیش بینی نوسانات سطح ایستابی پرداختند. نتایج نشان داد هر سه روش شبکه های عصبی، فازی-عصبی و برنامه ریزی بیان ژن می توانند به طور موفقیت آمیزی برای پیش بینی نوسانات سطح ایستابی مورد استفاده قرار گیرند. از طرفی عملکرد مدل برنامه ریزی بیان ژن در مقایسه با دو روش دیگر از جنبه های مختلف بهتر بود. پورصالحی و همکاران (۱۴۰۰) به مقایسه کارایی الگوریتم جنگل تصادفی در پیش بینی تراز سطح ایستابی آبخوان دشت بیرجند با دو مدل درخت تصمیم و شبکه عصبی مصنوعی پرداختند. نتایج پژوهش نشان داد که الگوریتم جنگل تصادفی توانایی بیشتری در شبیه سازی تراز سطح ایستابی آبخوان دارد و پارامترهای بارندگی و رطوبت نقش مؤثرتری نسبت به سایر پارامترها ایفا می کنند. در پژوهش دیگری، پیری و همکاران (۱۴۰۱) به تخمین تراز آب زیرزمینی دشت دشتستان در استان بوشهر با استفاده از شبکه عصبی مصنوعی، مدل ماشین بردار پشتیبان و درخت تصمیم پرداختند. نتایج تحقیق حاکی از دقت بالای هر سه روش مورد استفاده در شبیه سازی سطح ایستابی منطقه مورد مطالعه بوده، اما به طور کلی روش شبکه عصبی مصنوعی براساس نتایج روش برتر برای پیش بینی سطح ایستابی تشخیص داده شده است. رستمی نژاد و همکاران (۱۴۰۲) به ارزیابی کارایی مدل درخت تصمیم در ترکیب با تبدیل موجک (WM5P) برای پیش بینی نوسانات سطح ایستابی دشت کرمان-باغین پرداختند. نتایج نشان داد مدل ترکیبی موجک-درخت تصمیم دقت بالاتری در شبیه سازی تغییرات سطح آب زیرزمینی نسبت به مدل منفرد دارد. جبالبارزی و ملکیان (۱۳۹۸) کارایی دو روش شبکه عصبی و برنامه ریزی بیان ژن را در پیش بینی سطح ایستابی منطقه خشک و نیمه خشک دشت جیرفت بررسی کردند که براساس نتایج روش شبکه عصبی مصنوعی دقت بهتری در پیش بینی تراز سطح آب زیرزمینی منطقه مورد مطالعه دارا بوده است. در تحقیق دیگری از ترکیب پیش پردازنده موجک با مدل های یادگیری ماشین، رویکرد دسته بندی گروهی داده ها و روش حداقل مربعات ماشین بردار پشتیبان به منظور پیش بینی تغییرات تراز سطح آب زیرزمینی دشت زرنده-ساوه استفاده گردید. نتایج نشان داد

مدل ترکیبی پیش‌پردازنده موجه با حداقل مربعات ماشین بردار پشتیبان نسبت به مدل‌های منفرد و سایر مدل‌های ترکیبی مورد بررسی بهترین عملکرد را داشته است (سامانی، طهماسبی و طهماسبی سروستانی، ۱۳۹۹).

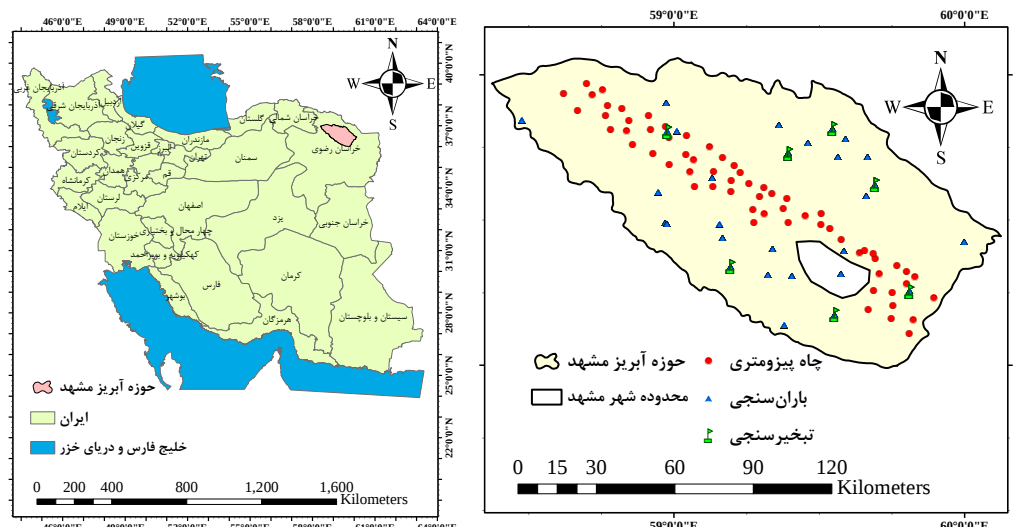
در پژوهشی از روش رگرسیون بهبودیافته جنگل تصادفی با ترکیبی از ویژگی‌های تصادفی به‌منظور پیش‌بینی کوتاه‌مدت تراز سطح ایستابی در چینگدائو کشور چین استفاده شد که نتایج نشان‌دهنده دقت بالای روش مورد استفاده در پیش‌بینی تراز سطح ایستابی منطقه است (وانگ^۱ و همکاران، ۲۰۱۸). پیش‌بینی سطح آب زیرزمینی دشت تهران-کرج با استفاده از مدل‌های منطق فازی، ANN، ANFIS و LSSVM نشان‌دهنده برتری نسبی روش ANFIS نسبت به سایر روش‌ها به‌خصوص در پیش‌بینی یک‌ماهه و دوماهه سطح آب زیرزمینی بود (ودیعتی^۲ و همکاران، ۱۴۰۰). فام^۳ و همکاران کارایی هفت مدل مختلف یادگیری ماشین (جنگل تصادفی، درخت تصادفی، مدل درخت M5T، الگوریتم یادگیری درخت، SVR، هرس خطای کاهش‌یافته درخت و رگرسیون خطی وزنی محلی) در پیش‌بینی تراز سطح ایستابی در شمال بنگلادش مورد ارزیابی قرار دادند. نتایج آن‌ها نشان داد مدل‌های مبتنی بر درخت تصادفی و جنگل تصادفی از عملکرد بهتری نسبت به سایر مدل‌ها برخوردارند (فام و همکاران، ۲۰۲۲) در مطالعه دیگری وئی^۴ و همکاران (۲۰۲۲) از ترکیب مدل تبدیل موجه با الگوریتم‌های یادگیری ماشین به‌منظور پیش‌بینی سطح آب زیرزمینی بهره بردند و نشان دادند که مدل‌های ترکیبی استفاده‌شده می‌توانند با دقت زیادی تغییرات سطح ایستابی را پیش‌بینی نمایند. اخیراً رواتی و لکشمی^۵ (۲۰۲۳) به بررسی سطح آب زیرزمینی در یک منطقه شهری با روش‌های ماشین بردار پشتیبان (SVM) و درخت تصمیم (M5) پرداخته و نشان دادند به‌رغم اینکه براساس پارامترهای آماری هر دو مدل دقت مناسبی دارند، اما مدل ماشین بردار پشتیبان سطح آب زیرزمینی را با دقت بهتری شبیه‌سازی نموده است. آبخوان دشت مشهد از زیرحوزه‌های حوزه آبخیز قره‌قروم در زمره دشت‌های بحرانی کشور محسوب می‌شود که با چالش‌های افت سطح ایستابی و کاهش کیفیت مواجه است، لذا تحقیق حاضر به‌منظور پیش‌بینی سطح ایستابی و شناسایی پارامترهای مؤثر هواشناسی در تغییرات سطح ایستابی با استفاده از روش درخت تصمیم و ماشین بردار پشتیبان و استفاده از ترکیبات متنوع پارامترهای اقلیمی به‌عنوان ورودی مدل‌ها هدف‌گذاری شده است. بررسی شرایط فعلی این منابع و شبیه‌سازی رفتار هیدرولوژیکی آن‌ها به‌منظور مدیریت بهینه و ازطرف دیگر شناخت صحیح و بهره‌برداری اصولی از منابع آب زیرزمینی می‌تواند در توسعه پایدار فعالیت‌های اجتماعی و اقتصادی منطقه نقش مؤثری ایفا کند.

مواد و روش‌ها

منطقه مورد مطالعه

دشت مشهد با وسعت ۹۹۰۹ کیلومتر مربع در شمال شرق استان خراسان رضوی بین طول‌های جغرافیایی ۵۸ درجه و ۲۰ دقیقه تا ۶۰ درجه و ۸ دقیقه و عرض‌های ۳۵ درجه و ۵۸ دقیقه تا ۳۷ درجه و ۳ دقیقه ایران قرار دارد (شکل ۱)، که ارتفاعات و دشت‌های آن به‌ترتیب ۶۵۵۸ و ۳۳۵۱ کیلومتر مربع وسعت داشته و از سمت شمال به هزار مسجد (کپه داغ) و از جنوب به کوه‌های بینالود، همچنین از شمال غرب توسط حوضه رودخانه اترک و جنوب شرق با رودخانه جامرود محدود شده و برای کشیدگی هر دو رشته‌کوه از شمال غرب به جنوب شرق است. قله بینالود، مرتفع‌ترین نقطه با ارتفاع ۳۲۴۸ متر از سطح دریاست که در جنوب شرق واقع شده و تنگل شور به‌عنوان خروجی حوضه، با ارتفاع ۹۰۴ متر از سطح دریا پایین‌ترین نقطه است (زارعی و همکاران، ۲۰۲۴). میانگین بارندگی درازمدت منطقه ۲۶۵ میلی‌متر در سال است که بیشترین بارندگی در زمستان و بهار اتفاق می‌افتد، کل آب تجدیدپذیر دشت مشهد ۹۳۵ میلیون متر مکعب و کل تخلیه آبخوان ۱۰۷۱ میلیون متر مکعب است (دوگانی، دوراندیش و قربانی، ۱۳۹۹). به‌دلیل برداشت بیش از حد از آب‌های زیرزمینی فرونشست زمین در برخی از نقاط منطقه مورد مطالعه رخ داده است که باعث شده از سال ۱۹۶۸ فعالیت‌های جدید صنعتی و کشاورزی در این آبخوان ممنوع باشد (عالم، فلاحتی و فرمانیه، ۱۳۹۸).

1. Wang
2. Vadiati
3. Pham
4. Wei
5. Revathy & Lekshmy



شکل ۱. موقعیت منطقه مطالعاتی، چاه‌های پیزومتري و ایستگاه‌های هواشناسی مورد استفاده در استان خراسان رضوی و ایران

روش پژوهش

امروزه کاربرد مدل‌های درختی در داده‌کاوی برای ساخت درخت و به دست آوردن قوانین برای پیش‌بینی‌ها بسیار رایج است. انواع مختلفی از الگوریتم‌های درختی از قبیل الگوریتم‌های CART، CHAID، QUEST، C5 و M5 استفاده می‌شوند که مدل M5 برای آموزش مدل‌های پیش‌بینی مقادیر عددی به کار می‌رود. تفاوت اصلی این الگوریتم با CART از آنجایی است که برگ‌های مدل M5 برخلاف درختان رگرسیونی CART که شامل مقادیر عددی هستند، حاوی مدل‌های خطی چندمتغیری هستند و از این دیدگاه، می‌توان مدل M5 را متشابه با توابع خطی چندضابطه‌ای دانست. علاوه بر این M5، در مقایسه با درختان رگرسیونی مانند CART، در یک کاربرد مشخص، کوچک‌تر، قابل فهم‌تر، دقیق‌تر و از کارایی بالاتری برخوردار است. در تحقیق حاضر از داده‌های سطح آب چاه‌های مشاهده‌ای، بارندگی، تبخیر، دما طی ۱۰ سال آبی برای پیش‌بینی سطح آب آبخوان دشت مشهد استفاده گردید. بدین صورت که پارامترهای هواشناسی تأثیرگذار بر سطح آب همچون بارندگی ماهانه، بارندگی یک ماه و دو ماه قبل، تبخیر ماهانه، تبخیر ماه قبل، دمای ماهانه، دمای یک ماه قبل و همچنین تراز سطح آب ماهانه و یک ماه قبل در قالب ۱۰ سناریوی متفاوت به‌عنوان ورودی مدل‌های مورد استفاده در تحقیق در نظر گرفته شدند (جدول ۱). برای استفاده از داده‌های مذکور در مدل‌های حداقل مربعات ماشین بردار پشتیبان و درخت تصمیم ابتدا داده‌ها به دو دسته (۷۵ درصد برای آموزش و ۲۵ درصد برای آزمون) تقسیم و سپس الگوریتم‌های مذکور تحت ۱۰ سناریوی مورد نظر در نرم‌افزار وکا (Weka) اجرا گردید.

جدول ۱. پارامترهای ورودی به مدل‌ها در سناریوهای مختلف

سناریو	ورودی‌ها
۱	$G_{t-1} - T_{t-1} - T_t - Et - Pt - P_{t-2} - P_{t-1}$
۲	$G_{t-1} - T_t - Et - Pt - P_{t-1}$
۳	$Pt - Et - T_{t-2} - T_{t-1} - T_t - G_{t-1}$
۴	$P_{t-1} - Pt - Et - T_{t-1} - T_t - G_{t-1}$
۵	$Pt - Et - T_{t-3} - T_{t-2} - T_{t-1} - T_t - G_{t-1}$
۶	$P_{t-1} - Et - T_{t-3} - T_{t-2} - T_{t-1} - T_t - G_{t-1}$
۷	$P_{t-2} - P_{t-1} - Pt - Et - T_{t-2} - T_{t-1} - T_t - G_{t-1}$
۸	$P_{t-2} - Pt - Et - T_{t-2} - T_{t-1} - T_t - G_{t-1}$
۹	$Et - T_{t-1} - T_t - G_{t-1} - Pt$
۱۰	$Pt - Et - T_t - G_{t-1}$

۱. مدل درخت تصمیم M5

مدل درخت تصمیم معمولاً از چهار بخش ریشه، گره، شاخه و برگ‌ها تشکیل شده و اساس روش، تقسیم و غلبه صفات برای نمونه‌هایی است که به یک گره می‌رسند. در ابتدا با تقسیم کردن فضای مسئله به صورت برگشتی یک درخت رگرسیونی خطی

چندمتغیره برازش داده می‌شود؛ بدین صورت که ریشه گره از بالا به سمت پایین آغاز شده، به برگ‌ها می‌رسد و برای ایجاد شاخه در یک گره تقسیم از پارامتر انحراف معیار مقادیر متغیر هدف به عنوان یک معیار اندازه‌گیری خطا در آن گره استفاده می‌شود (فلاحی، واریانی و گلیان، ۱۳۹۰). فرایند انشعاب در هر گره بارها تکرار می‌شود و آزمون برای انجام عملیات تقسیم در گره‌های مذکور انجام شده، سپس صفتی که موجب کاهش بیشتر انحراف معیار گردد، انتخاب می‌شود و در نهایت با رسیدن به برگ‌ها مجموع مجذور انحراف میانگین داده‌ها به صفر خواهند رسید و ترکیب نتایج آن را تحلیل می‌کند (نهرین، ستاری و بیگزالی، ۱۳۹۲). کاهش انحراف معیار استاندارد با استفاده از رابطه (۱) محاسبه می‌شود.

$$SDR = sd(t) - \sum \frac{|T_i|}{|T|} \times sd(T_i) \quad \text{رابطه (۱)}$$

در رابطه فوق، T شامل نمونه‌هایی است که به گره مورد نظر رسیده‌اند. T_i ها تعداد داده‌هایی هستند که از تقسیم کردن در گره مورد نظر براساس صفت انتخاب شده به دست آمده‌اند. Sd نیز انحراف معیار نمونه‌هایی است که به گره مورد نظر می‌رسند. فرایند شاخه زدن، هرس کردن و هموارسازی تا زمانی که تمام نمونه‌هایی که به یک گره می‌رسند، تفاوت اندکی داشته باشد، ادامه خواهند یافت. پس از تکمیل درخت، برای نمونه‌های موجود در هر زیرشاخه گره داخلی، یک مدل رگرسیون خطی چندمتغیری برازش داده می‌شود. این مدل، براساس داده‌های وابسته به آن گره و تمام صفاتی که در زیر درخت یا ریشه آن گره مورد استفاده قرار گرفته‌اند، ایجاد می‌شود.

۲. مدل حداقل مربعات ماشین بردار پشتیبان (LS-SVM)

ماشین بردار پشتیبان نوعی سیستم یادگیری کارآمد بر مبنای تئوری بهینه‌سازی مقید است که توسط کورتز و وپنیک^۱ (۱۹۹۵) ارائه شد. در این مدل از اصل استقراری کمینه‌سازی خطای ساختاری استفاده و به یک جواب بهینه کلی منجر می‌شود و برای حل معادلات از برنامه‌ریزی دوگانه استفاده می‌شود (خطی، قربانی و حسن‌پور کاشانی، ۲۰۱۱). بنابراین، حل مسائل در ابعاد بزرگ با استفاده از مدل یادشده، هزینه محاسباتی را به شدت افزایش می‌دهد. برای رفع این مشکل سویکنز^۲ و همکاران (۲۰۰۲) روش کمترین مربعات ماشین بردار پشتیبان را ارائه کردند. در این مدل از معادلات خطی ساده برای حل مسائل استفاده می‌شود که به‌طور مؤثری پیچیدگی الگوریتم را کاهش می‌دهد. تابع رگرسیونی استفاده شده در مدل SVM برای تخمین مسائل به‌صورت رابطه (۲) است.

$$y(x_i) = w^T \cdot \varphi(x_i) + b \quad \text{رابطه (۲)}$$

که در آن $\varphi(x_i)$ ترسیم غیرخطی ورودی‌ها در فضای ویژگی با ابعاد زیاد است. پارامترهای w و b به ترتیب مقادیر وزن‌ها و بایاس تابع رگرسیون هستند که از طریق حداقل‌سازی تابع هدف در رابطه (۳) تعیین می‌شوند.

$$\min_{w, e, b} j(w, e) = \frac{1}{2} w^T w + \frac{\gamma}{2} \sum_{i=1}^N e_i^2 \quad \text{رابطه (۳)}$$

$$Y_i = w^T \cdot \varphi(x_i) + b + e_i \quad i=1, 2, \dots, N \quad \text{رابطه (۴)}$$

که در آن γ پارامتر تنظیم‌کننده بخش خطا و e_i خطای داده‌های آموزش است. در نهایت، تابع تخمین مدل LS-SVM به‌صورت رابطه (۵) تعریف می‌شود.

$$y(x) = \sum_{i=1}^N a_i K(x_i, x_j) + b \quad \text{رابطه (۵)}$$

در رابطه فوق $K(x_i, x_j)$ تابع کرنل نامیده می‌شود که مطابق زیر به عنوان تابعی با ایجاد ضرب داخلی در فضای ویژگی معرفی می‌شود (سویکنز و همکاران، ۲۰۰۲).

$$K(x_i, x_j) = \varphi(x_i) \cdot \varphi(x_j) \quad i, j=1, 2, \dots, N \quad \text{رابطه (۶)}$$

برای تحلیل حساسیت مدل (LS-SVM) توابع کرنل خطی (Linear)، چندجمله‌ای (Polynomial) و پایه شعاعی (Radial Basis Function Kernel) آزموده شد. مقادیر پارامترهای تنظیم‌کننده و کرنل به ترتیب برابر ۱۰ و ۰/۴ در نظر گرفته شد.

۳. معیارهای ارزیابی مدل‌ها

در تحقیق حاضر از شاخص‌های آماری ریشه میانگین مربعات خطا (RMSE)، میانگین مطلق خطا (MBE) و ضریب تعیین (R^2) به منظور ارزیابی صحت و اعتبارسنجی نتایج مدل‌های حداقل مربعات ماشین بردار پشتیبان و درخت تصمیم استفاده شده است (روابط ۷ تا ۹).

$$RMSE = \sqrt{\frac{\sum (X_o - X_s)^2}{N}} \quad \text{رابطه ۷}$$

$$MAE = \frac{\sum |X_o - X_s|}{N} \quad \text{رابطه ۸}$$

$$NSE = [1 - \frac{\sum (X_o - X_s)^2}{\sum (X_o - \bar{X_o})^2}] \quad \text{رابطه ۹}$$

در روابط فوق X_o و X_s به ترتیب داده‌های مشاهداتی و شبیه‌سازی شده توسط مدل، X_o میانگین داده‌های مشاهداتی و N تعداد کل داده‌ها هستند. در هریک از مدل‌های مورد استفاده که مقادیر به دست آمده برای شاخص‌های میانگین مطلق خطا و مجذور میانگین مربعات خطا به صفر و معیار نش-ساتکلیف به یک نزدیک‌تر باشد، حاکی از دقت بالاتر روش مورد نظر در شبیه‌سازی است.

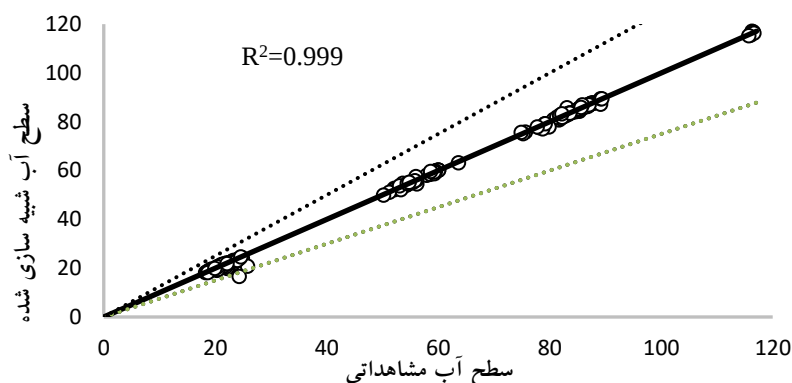
نتایج

جدول ۲، نتایج عملکرد مدل حداقل مربعات ماشین بردار پشتیبان (LS-SVM) به ازای ترکیبی از ورودی‌های مختلف نشان می‌دهد. مقدار بالای R^2 (۰/۹۹۹) و مقادیر پایین MBE و RMSE به ترتیب ۰/۱۶۸- و ۰/۹۴۹، در سناریوی ۴ (باران- باران ماه قبل-تبخیر- دما- دمای ماه قبل- تراز ماه قبل) در مرحله آزمون نشان می‌دهد سطح آب با دقت خوبی پیش‌بینی شده است. از طرف دیگر، مدل LS-SVM با ترکیب ورودی‌های مربوط به سناریوی ۹ (باران دو ماه قبل- باران- تبخیر- دمای دو ماه قبل- دمای ماه قبل- دما- تراز ماه قبل) نیز پیش‌بینی مناسبی از سطح ایستابی داشته، به طوری که R^2 ، MBE و RMSE در این سناریو به ترتیب ۰/۹۹۷، ۰/۰۱۸- و ۱/۷۵۵ به دست آمده است. مقایسه نتایج شاخص‌های ارزیابی مدل LS-SVM در سناریوهای مختلف نشان می‌دهد سناریوهای ۴ و ۹ بهترین ترکیب ورودی برای برآورد سطح آب زیرزمینی را دارند، که این برتری بیانگر تأثیر داده‌های اقلیمی در پیش‌بینی آب‌های زیرزمینی منطقه مورد مطالعه است. مقادیر خطای MBE در هر دو سناریو نشان می‌دهد مدل سطح آب زیرزمینی منطقه را نسبت به واقعیت کمتر برآورد نموده و به طور کلی هر دو سناریو توانسته‌اند سطح آب را با دقت قابل قبولی پیش‌بینی کنند. به طور کلی بررسی ورودی‌ها نشان می‌دهد که اضافه کردن پارامترهای بیشتر، همیشه به افزایش دقت پیش‌بینی پیش‌بینی منجر نمی‌گردد و گاهی ممکن است دقت را به مقدار زیادی کاهش دهد. در مجموع به ترتیب بیشترین و کمترین دقت در برآورد سطح آب زیرزمینی منطقه مورد مطالعه مربوط به سناریوهای ۴ و ۱۰ مدل حداقل مربعات ماشین بردار پشتیبان است.

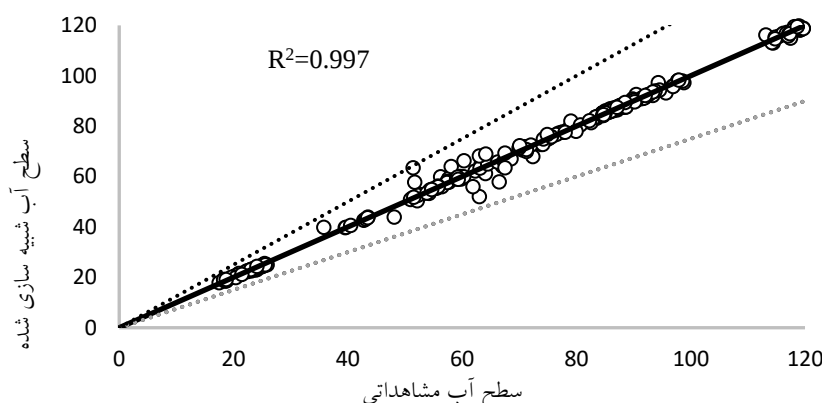
جدول ۲. نتایج ارزیابی سناریوهای مختلف در مدل LS-SVM

سناریو	آموزش			آزمون		
	R^2	RMSE	MBE	R^2	RMSE	MBE
۱	۰/۹۹۴	۲۴/۲۲	۴/۸۲۰	۰/۰۰۰۷۵	۲۸/۱۷	۲/۲۳۰
۲	۰/۶۵۸	۱۷/۶۴۹	۰/۶۶۲	۰/۰۰۵	۳۰/۲۴۹	- ۲/۹۳۱
۳	۰/۶۷۱	۱۷/۵۹۵	۰/۸۰۵	۰/۰۰۹	۳۱/۰۴۵	- ۰/۳۵۵
۴	۰/۹۹۹	۰/۹۵۱	- ۰/۰۳۳	۰/۹۹۹	۰/۹۹۲	- ۰/۱۵۱
۵	۰/۳۰۴	۲۶/۳۵	۰/۰۰۰	۰/۰۰۸	۳۰/۸۷	۳/۶۰۵
۶	۱	۶/۳۲۱	۲/۴۶۸	۰/۰۰۳	۲۸/۷۱	۲/۶۹۰
۷	۱	۱/۴۰۷	۳/۴۸۴	۰/۰۰۱	۲۸/۸۵	- ۱/۵۰۶
۸	۰/۷۳۵	۱۴/۲۶	- ۰/۲۵۱	۰/۰۰۱	۲۹/۲۳	- ۲/۱۷۹
۹	۰/۹۹۵	۲/۱۱۳	- ۰/۰۳۳	۰/۹۹۷	۱/۷۵۵	- ۰/۰۱۸
۱۰	۰/۹۷۸	۰/۹۸۴	- ۰/۱۹۷	۰/۰۰۰	۲۹/۳۹۰	- ۲/۴۲۰

ارزیابی مدل (LS-SVM) براساس نمودارهای پراکنش حول محور یک‌به‌یک در مرحله صحت‌سنجی در سناریوهای مورد استفاده نیز حاکی از پراکنش نامناسب حول محور یک‌به‌یک و همبستگی پایین بین مقادیر مشاهداتی و شبیه‌سازی شده در تمام سناریوها به جز سناریوهای ۴ و ۹ است؛ به‌طوری‌که در سناریو ۴ دقت پیش‌بینی مدل بسیار خوب بوده و توانسته سطح آب زیرزمینی را در تمام محدوده ۲۰ تا ۱۲۰ متر با خطای کمی پیش‌بینی کند (شکل ۲). همچنین در سناریوی ۹ تقریباً تمامی مقادیر سطح آب پیش‌بینی شده در محدوده ± 25 درصد واقع شده‌اند که نشان‌دهنده دقت قابل قبول در شبیه‌سازی است (شکل ۳).



شکل ۲. سناریوی چهارم مدل کمترین مربعات ماشین بردار پشتیبان در مرحله صحت‌سنجی



شکل ۳. سناریوی نهم مدل کمترین مربعات ماشین بردار پشتیبان در مرحله صحت‌سنجی

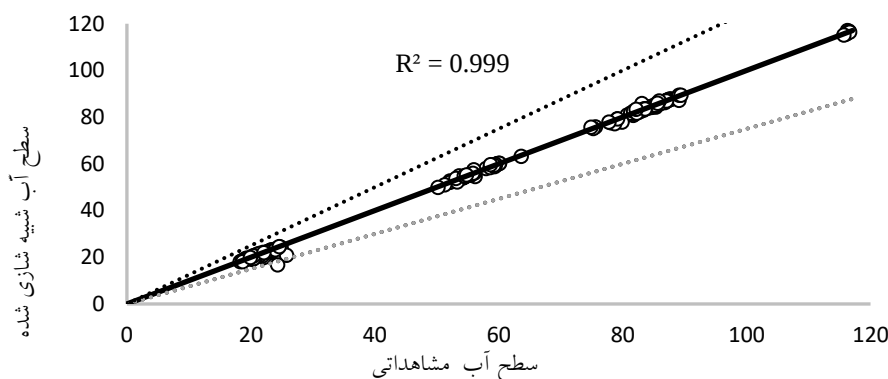
با توجه به مقدار بالای R^2 (۰/۹۹۹) در ترکیب سناریوی ۴ (باران یک ماه قبل - باران - تبخیر - دمای یک ماه قبل - دما - تراز ماه قبل) و در کنار آن مقادیر پایین MBE و RMSE به ترتیب حدود $-0/168$ و $0/949$ نسبت به شاخص‌های ارزیابی تخمین مدل M5 با سناریوهای دیگر، می‌توان ورودی‌های سناریوی ۴ را به عنوان ترکیب ورودی برتر معرفی کرد. در سناریو ۲ که با حذف بارندگی دو ماه قبل و دمای ماه قبل از سناریو ۱ ایجاد شده، مقدار خطای مجذور میانگین مربعات حدود ۳۰ درصد کاهش می‌یابد. در سناریوی ۴ نیز که نسبت به سناریو ۱ فقط بارندگی دو ماه قبل را به عنوان ورودی ندارد، خطای میانگین مربعات به مقدار قابل توجهی کاهش پیدا کرده اما مدل کم‌برآوردی نشان می‌دهد. بنابراین می‌توان نتیجه گرفت که در مدل درخت تصمیم بارندگی دو ماه قبل باعث افزایش خطای شبیه‌سازی سطح آب شده است. اضافه شدن دمای سه ماه قبل به سناریوی ۳ مقدار خطای سناریو ۵ نسبت به سناریو ۳ حدود ۳۰ درصد افزایش می‌یابد که بیانگر تأثیر منفی دمای سه ماه قبل در ترکیب ورودی‌های سناریو ۳ بر پیش‌بینی سطح آب است. در سناریو ۶ بارندگی ماه قبل با بارندگی همان ماه جایگزین شده است و مقدار خطا نسبت به سناریو ۵ حدود ۶۵ درصد کاهش یافته است. با اضافه شدن دمای دو ماه قبل به سناریو ۱ و ایجاد سناریو ۷ خطای پیش‌بینی به مقدار قابل توجهی کاهش یافته و کم‌برآوردی نیز بهبود پیدا می‌کند. با حذف بارندگی ماه قبل از سناریو ۷ و تولید سناریو ۸ مقدار خطای مدل به مقدار زیادی افزایش نشان می‌دهد. با حذف بارندگی ماه قبل از سناریو ۴، سناریو ۹ ایجاد شده که خطای آن سناریو بیش از دو برابر سناریو ۴ است، که بیانگر تأثیر بارندگی ماه قبل در برآورد سطح آب زیرزمینی بوده و تأثیر زیادی بر بهبود نتایج سناریوهای ۴ و ۷ نشان داده و باعث افزایش دقت شبیه‌سازی سطح آب شده

است. با جایگزینی تبخیر ماه قبل با دمای ماه قبل در سناریو ۹ و ایجاد سناریو ۱۰ بیش از ۳۰ درصد افزایش خطا در شبیه‌سازی سطح آب زیرزمینی مشاهده می‌شود. به‌طور کلی سناریوهای ۴ و ۵ مدل درخت تصمیم به‌ترتیب بیشترین و کمترین دقت را در برآورد سطح آب نسبت به سایر سناریوها به خود اختصاص می‌دهند (جدول ۳).

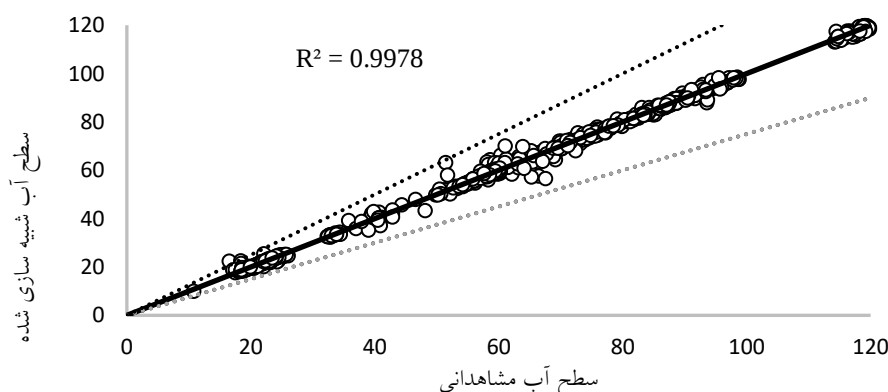
جدول ۳. نتایج ارزیابی سناریوهای مختلف در مدل M5

سناریو	آموزش			آزمون		
	MBE	RMSE	R ²	MBE	RMSE	R ²
۱	-۰/۹۹۵	۱/۸۸۷	-۰/۰۳۷	-۰/۲۲۲	۲/۳۷۶	۰/۹۹۳
۲	-۰/۹۹۵	۲/۰۵۴	-۰/۰۲۶	-۰/۱۱۳	۱/۷۶۸	۰/۹۹۷
۳	-۰/۹۹۸	۱/۳۰۰	-۰/۰۱۴	-۰/۲۸۱	۲/۱۰۲	۰/۹۹۵
۴	-۰/۹۹۹	۰/۹۲۱	-۰/۰۳۴	-۰/۱۶۸	۰/۹۴۹	۰/۹۹۹
۵	-۰/۹۹۸	۱/۴۲۲	-۰/۰۲۲	-۰/۲۸۳	۳/۱۵۳	۰/۹۸۶
۶	-۰/۹۹۵	۲/۱۷۶	-۰/۰۴۴	-۰/۰۶۸	۱/۸۸	۰/۹۹۶
۷	-۰/۹۹۴	۲/۰۸۳	-۰/۰۵۰	-۰/۰۸۴	۱/۴۲۳	۰/۹۹۸
۸	-۰/۹۹۵	۱/۹۷۷	-۰/۰۴۵	-۰/۳۱۵	۲/۲۴۵	۰/۹۹۴
۹	-۰/۹۹۷	۱/۷۵۵	-۰/۰۱۸	-۰/۰۳۳	۲/۱۱۳	۰/۹۹۵
۱۰	-۰/۹۹۷	۱/۶۴۱	-۰/۰۰۸	-۰/۰۱۴	۲/۸۳۶	۰/۹۹۱

ارزیابی مدل درخت تصمیم M5 براساس نمودارهای پراکنش حول محور یک‌به‌یک در مرحله صحت‌سنجی در سناریوهای مورد استفاده حاکی از پراکنش مناسب سطح آب زیرزمینی مشاهداتی شبیه‌سازی شده حول محور یک‌به‌یک و همبستگی مناسب بین مقادیر مشاهداتی و شبیه‌سازی شده در تمام سناریوهای مورد استفاده است؛ به‌طوری‌که در سناریو ۴ دقت پیش‌بینی مدل بسیار خوب بوده و توانسته سطح آب زیرزمینی را در تمام محدوده ۲۰ تا ۱۲۰ متر با خطای کم و با ضریب همبستگی نزدیک به یک (۰/۹۹۹) پیش‌بینی کند (شکل ۴). همچنین در سناریوی ۵ که دقت کمتری نسبت به سایر سناریوها داشته، تمام نقاط در محدوده ± 25 درصد قرار داشته است. پراکندگی نقاط در این محدوده زیاد بوده ولی در مجموع پیش‌بینی قابل قبولی انجام شده است (شکل ۵).

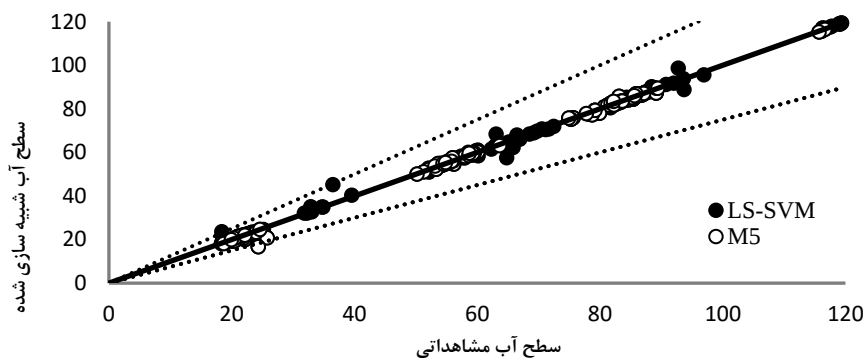


شکل ۴. سناریوی چهارم مدل درخت تصمیم (M5) در مرحله صحت‌سنجی



شکل ۵. سناریوی پنجم مدل درخت تصمیم (M5) در مرحله صحت‌سنجی

مقایسه نتایج سناریو ۴ مدل حداقل مربعات ماشین بردار پشتیبان و مدل درخت تصمیم نشان می‌دهد که هر دو مدل سطح آب را با دقت قابل قبولی پیش‌بینی کرده‌اند و تمام نقاط در محدوده $\pm 25\%$ درصد قرار دارد (شکل ۶). اگرچه پراکندگی نقاط حول خط ۱:۱ در مدل حداقل مربعات ماشین بردار پشتیبان بیشتر از مدل درخت تصمیم است، هر دو مدل توانسته‌اند در تمام محدوده مورد مطالعه سطح آب را با دقت تقریباً یکسانی تخمین بزنند. به‌طور کلی مدل LS-SVM در مقایسه با درخت تصمیم نمی‌تواند نتایج منطقی درخصوص پیش‌بینی سطح آب ارائه دهد و حساسیت این مدل نسبت به ورودی‌های مختلف، زیاد است.



شکل ۶. مقایسه مدل LS-SVM و M5 به‌ازای سناریوی چهارم در مرحله صحت‌سنجی

بحث و نتیجه‌گیری

بررسی سناریوهای مختلف در مدل حداقل مربعات ماشین بردار پشتیبان نشان می‌دهد این مدل فقط در سناریو ۴ و ۹ نتایج قابل قبولی ارائه داده و در سناریوهای دیگر مقدار خطا بسیار زیاد است. تأثیر ورودی‌های مختلف بر مدل حداقل مربعات ماشین بردار پشتیبان بسیار زیاد بوده، به‌طوری‌که با اضافه و یا کم شدن یک ورودی مقدار خطا تغییرات زیادی دارد. مدل حداقل مربعات ماشین بردار پشتیبان با بسیاری از ترکیب ورودی‌ها نتوانسته شبیه‌سازی مناسبی برای سطح آب منطقه مورد مطالعه داشته باشد. در بعضی از سناریوهای مدل LS-SVM مانند ۷ و ۱۰ مقدار زیاد ضریب تعیین و خطای پایین در مرحله آموزش نشان می‌دهد مدل در مرحله آموزش نتایج خوبی ارائه داده اما در مرحله آزمون مقادیر خطا به مقدار زیادی افزایش می‌یابد. اختلاف زیاد نتایج در مرحله آموزش و آزمون به این معناست که مدل نتوانسته با ترکیب ورودی‌ها به‌خوبی آموزش ببیند و فقط توانسته داده‌هایی را که با آن آموزش دیده است شبیه‌سازی کند. به‌طور کلی بررسی سناریوهای مختلف در مدل درخت تصمیم (M5) حاکی از دقت قابل قبول همه سناریوهای ورودی در شبیه‌سازی سطح آب است، به‌طوری‌که پراکنش داده‌های سطح آب مشاهداتی و شبیه‌سازی‌شده در تمام سناریوها در محدوده $\pm 25\%$ درصد قرار دارد، اما دقت سناریوها با یکدیگر متفاوت است. بررسی تمام سناریوهای مدل درخت تصمیم نشان می‌دهد در بعضی از سناریوها دقت مرحله آموزش بیشتر از آزمون است و در برخی سناریوها دقت مرحله آزمون بیشتر از آموزش است که می‌تواند به‌دلیل تصادفی بودن داده‌ها باشد. بهترین نتایج مدل درخت تصمیم مربوط به سناریو ۴ و سپس سناریو ۷ و ضعیف‌ترین شبیه‌سازی مربوط به سناریو ۵ است. مقایسه سناریوها ۴ و ۷ نشان می‌دهد اگرچه تعداد ورودی‌ها در سناریو ۷ بیشتر است، دقت پیش‌بینی آن کمتر است؛ بنابراین می‌توان نتیجه گرفت بارندگی و دمای دو ماه قبل (در سناریو ۷) باعث کاهش دقت شبیه‌سازی و افزایش مقدار خطا شده‌اند. مقایسه سناریوهای مختلف در مدل‌های درخت تصمیم و حداقل مربعات ماشین بردار پشتیبان نشان می‌دهد هر دو مدل با سناریو ۴ بهترین شبیه‌سازی از سطح آب منطقه مورد مطالعه را ارائه داده‌اند. از طرفی، بررسی سناریوهای دیگر نشان می‌دهد که کم و یا اضافه شدن پارامتر ورودی به ترکیب پارامترهای موجود در سناریو ۴ منجر به افزایش مقدار خطای مدل خواهد شد. در مجموع با وجود ترکیب پارامترهای مجموع بارندگی ماه قبل، دمای ماه قبل و سطح آب ماه قبل در سناریو ۴، مدل‌های مورد استفاده نتایج مطلوب‌تری ارائه کرده‌اند که با نتایج خاشعی سیوکی و همکاران (۱۳۹۲) مطابقت دارد. همچنین خلیلی نفت چالی و همکاران (۱۳۹۶) در ارزیابی مدل‌های الگوریتم تنبیل و درخت تصمیم در برآورد سطح آب زیرزمینی دشت نیشابور نشان دادند اضافه شدن پارامتر مجموع بارندگی ماه قبل به سناریوها باعث بهبود دقت مدل‌های فوق در برآورد سطح آب زیرزمینی شده است که با نتایج تحقیق حاضر همخوانی دارد. به‌طور کلی حساسیت مدل M5 نسبت به تغییر پارامترهای ورودی بسیار کمتر از

مدل LS-SVM است، به طوری که مدل درخت تصمیم در تمام ترکیب‌ها نتایج قابل قبولی ارائه کرده است؛ اما خطای بعضی از سناریوها در مدل حداقل مربعات ماشین بردار پشتیبان بسیار زیاد است. پیری و همکاران (۱۴۰۱) نیز با مقایسه سه مدل شبکه عصبی مصنوعی، درخت تصمیم و ماشین بردار پشتیبان نتایج مشابهی به دست آورده و نشان دادند مدل درخت تصمیم دقت خوبی در برآورد سطح آب زیرزمینی دشتستان بوشهر دارد. مقایسه روش‌های مختلف با مدل درخت تصمیم در مطالعه یانگ^۱ و همکاران (۲۰۰۹) نیز نشان داد اگرچه مدل‌های درخت تصمیمی و ماشین بردار پشتیبان دقت مناسبی در برآورد سطح آب زیرزمینی داشته‌اند، مدل شبکه عصبی مصنوعی نسبت به دو مدل دیگر خطای کمتری در پیش‌بینی داشته است. همچنین مطالعه رواتی و لمشکی (۲۰۲۳) نیز نشان داد دو مدل ماشین بردار پشتیبان و درخت تصمیم دارای دقت قابل قبولی در شبیه‌سازی سطح آب زیرزمینی است. به طور کلی نتایج تحقیق حاضر بر این موضوع تأکید دارد که انتخاب مناسب ورودی‌ها و پارامترهای اقلیمی می‌تواند تأثیر زیادی بر دقت پیش‌بینی داشته باشد به‌ویژه در سناریوهایی که شامل داده‌های بارندگی و دما در زمان‌های مختلف بودند، بارندگی مؤثرترین پارامتر در پیش‌بینی سطح ایستابی است. تأثیر پارامترهای مختلف در هر دو مدل یکسان نیست و از طرفی حساسیت مدل‌ها نیز نسبت به پارامترهای یکسان، متفاوت است؛ بنابراین آنالیز حساسیت هر مدل می‌تواند پارامترهای مؤثر را برای آن مدل تعیین نماید.

References

- Aalem, H., Fallahi, M., & Farmanieh, S. (2019). Estimating Runoff Using SCS - CN Based On GIS: A Case Study (Shirvan, Bojnord, Faruj, Safiabad and Meshkan Cities). *New findings in applied geology*, 13(26), 156-16. (In Persian)
- Abareshi, F., Meftah Halghi, M., Sanikhani, H., & Dehghani, A.A. (2014). Comparison of three intelligence techniques for predicting water table depth fluctuations (Case study: Zarringol plain). *Journal of Water and Soil Conservation Research*, 21(1), 163-180. (In Persian)
- Cortes, C., & Vapnik, V. (1995). Support-vector network, *Mach. Learn*, 20: 273–297.
- Dogani, A., Dourandish, A., & Ghorbani, M. (2020). Ranking of resilience indicators of Mashhad plain to groundwater resources reduction by Bayesian best-worst method. *Iranian Journal of Water and Irrigation Management*, 10(2), 301–316.
- Fallahi, M.R., Varvani, H., & Golian, S. (2012). Precipitation prediction using tree regression model to flood control. *Fifth National Conference on Watershed Management and Soil and Water Resources Management*. Kerman, Iran. (In Persian)
- Fallahzade, M., Rezaei, P., Eslamian, S., & Abbasi, A. (2019). Relationship of Drought and Teleconnection Patterns; Case Study of Qara-Qom Basin. *Geographical Researches Quarterly Journal*, 34(2), 153-164. (In Persian)
- Jabaalbarez, B., & Malekian, A. (2019). Comparison of the performance of artificial neural networks and gene expression to predict the groundwater level in arid and semi-arid areas (Case study: Jiroft plain). *Iranian Journal of Range and Desert Research*, 26(2), 292-301. (In Persian)
- Khalili Naft Chali, A., Shahidi, A., & Khashei Siuki, A. (2017). Comparison of Lazy Algorithms and M5 Model to Estimate Groundwater Level (Case Study: Plain Neyshabur). *J. Water and Soil Sci*, 21(3), 15-26. (In Persian)
- Khashei Siuki, A., Ghahraman, B., & Kouchakzadeh, M. (2013). Comparison of ANN, ANFIS and Regression Models to Estimate Groundwater level of Neyshaboer Aquifer. *Iranian Journal of Irrigation and Drainage*, 1(7), 10-22. (In Persian)
- Khatibi, R., Ghorbani, M.A., Hasanpour Kashani, M., & Kisi, O. (2011). Comparison of three artificial intelligence techniques for discharge routing. *Journal of Hydrology*, 403(3 -4), 201 -212.
- Milan, S.G., Roozbahani, A., & Banihabib, M.E. (2018). Fuzzy optimization model and fuzzy inference system for conjunctive use of surface and groundwater resources. *Journal of Hydrology*, 566, 421-434.
- Misaghi, F., & Mohammadi, K. (2006). Zoning of rainfall data using classical statistical and geostatistical methods and comparison with artificial neural networks. *Scientific Journal of Agriculture*, 29(4), 1-13. (In Persian)
- Nahrin, F., Sattari, M. T., & Bigzali, N. (2013). Comparison of suspended load estimation using two methods: sediment gauge curve and M5 tree model (Case study: Liqvan Chay River). *12th Iranian Hydraulic Conference*. (In Persian)
- Pham, Q. B., Kumar, M., Di Nunno, F., Elbeltagi, A., Granata, F., Islam, A.R.M.T., Talukdar, S., Nguyen, X.C., Ahmed, A.N., & Anh, D.T. (2022). Groundwater level prediction using machine learning algorithms in a drought-prone area. *Neural Computing and Applications*, 34(13): 10751-10773.
- Piri, H., Mobaraki, M. & Siasar, Saleh. (2023). Temporal and spatial modeling of groundwater level in Bushehr plain using artificial intelligence and geostatistic. *Journal Watershed Management Research*, 13(26), 58-68. (In Persian)
- Poursalehi, F., KhasheiSiuki, A., & Hashemi, S. R. (2022). Investigating the performance of the random forest algorithm in predicting water table fluctuations in comparison with two decision tree models and artificial neural network in the Birjand plain aquifer. *Ecohydrology*, 8(4), 961-974. (In Persian)
- Rajaei, T., & Ebrahimi, H. (2016). Application of wavelet neural network model for forecasting groundwater level time series with non-stationary and nonlinear characteristics. *Journal of Water and Soil Conservation*, 22(5), 99- 115. (In Persian)

- Revathy, R., & Lekshmy, D.C.A. (2023). Groundwater Level Prediction Using Support Vector Machine and M5 Model Tree-A Case Study. *Proceedings of the 7th Biennial International Conference on Emerging Trends in Engineering, Science & Technology (ICETEST 2023)*
- Rostaminezhad Dolatabad, H., Shahabi, S., & Madadi, M.R. (2023). Evaluation of the efficiency of decision tree in combination with wavelet transform for predicting groundwater level fluctuations in Kerman Baghin Plain. *Iranian Journal of Irrigation and Drainage*. 17:3 (99), 413-427. (In Persian)
- Samani, J.M., Tahmasbi, A., & Tahmasbi Sarvestani, Z. (2021). Water Resources Management and Sustainable Development. *Infrastructure Studies Office*, Volume 22, Serial Number 7374. (In Persian)
- Suykens, J.A.K., Gestel, T.V., Barbanter, J.D., Moor, B.D., & Vandewalle, J. (2002). Least squares support vector machines. *World Scientific Pub. Co. Inc.* ISBN: 978-981-277-665-5. 308 Pages.
- Vadiati, M., Rajabi Yami, Z., Eskandari, E., Nakhaei, M., & Kisi, O. (2022). Application of artificial intelligence models for prediction of groundwater level fluctuations: Case study (Tehran-Karaj alluvial aquifer). *Environmental Monitoring and Assessment*, 194(9), 619.
- Wang, X., Liu, T., Zheng, X., Peng, H., Xin, J., & Zhang, B. (2018). Short-term prediction of groundwater level using improved random forest regression with a combination of random features. *Applied Water Science*, 8(5): 1-12.
- Wei, A., Chen, Y., Li, D., Zhang, X., Wu, T., & Li, H. (2022). Prediction of groundwater level using the hybrid model combining wavelet transform and machine learning algorithms. *Earth Science Informatics*, 15(3), 1951-1962.
- Yang, Z.P., Lu, W.X., Long, Y.Q., & Li, P. 2009. Application and comparison of two prediction models for groundwater levels: A case study in Western Jilin Province. China. *Journal Arid Environ*, 73, 487-492.
- Zarei, M., Ghazavi, R., Abdollahi, KH., Ranzi, R., Ramesh, S.V.T., & Barontini, S. (2024). Spatiotemporal variation of water balance components in Mashhad catchment, Iran: Investigating the impact of changes in climatic data and land use. *Water Supply*, 24(2), 397-415. doi: 10.2166/ws.018