

Research Paper

Application of Machine Learning Algorithms in Predicting the Water Area of Gorgan Bay Using Hydrological and Climatic Data

Alireza Yousefi Kebriya¹,Mehdi Nadi^{2*}

¹ Ph.D. of Agricultural Meteorology, Department of Water Engineering, Sari Agricultural Sciences and Natural Resources University, Sari, Iran.

² Associated Professor of Agricultural Meteorology, Department of Water Engineering, Sari Agricultural sciences and Natural Resources University, Sari, Iran.



10.22125/iwe.2025.495508.1844

Received:
December 23, 2024
Accepted:
April 21, 2025
Available online:
April 25, 2025

Keywords:
Caspian Sea Water
Level, Remote Sensing,
Machine Learning,
Statistical Modeling,
Water Surface Area
Changes.

Abstract

Gorgan Bay, located in northern Iran, is a sensitive and ecologically significant coastal ecosystem facing increasing threats due to climate change, reduced freshwater inflow, and human activities. This study aimed to model and predict changes in the bay's water area using hydrological, satellite, and climatic data from 2000 to 2023, including temperature, precipitation, discharge, water level, and the MNDWI-derived water area of the Caspian Sea. Several advanced machine learning models—XGBoost, Radial Basis Function Network (RBFN), Random Forest, and Linear Regression—were applied due to their ability to model complex, nonlinear relationships. Correlation analysis and linear regression were used to explore inter-variable relationships. Results showed that the Caspian Sea's water level had the highest correlation with Gorgan Bay's water area ($r = 0.90$, $p < 0.001$), followed by the Caspian water area ($r = 0.66$, $R^2 = 0.43$) and inflow discharge ($r = 0.63$, $R^2 = 0.40$). Precipitation and temperature had weaker effects ($r = 0.43$ and -0.34 , respectively). Among models, XGBoost showed the highest accuracy ($R^2 = 0.93$, $MAPE = 4.9\%$, $RMSE = 16.9$, $MAE = 15.3$), outperforming others. Five years (2000, 2008, 2011, 2016, 2018) were used for validation. While all models simulated fluctuations, XGBoost captured variations most precisely, making it the most suitable model for predicting Gorgan Bay's water area changes.

1. Introduction

Gorgan Bay, a shallow water body in northern Iran connected to the Caspian Sea, holds critical ecological and economic value. However, it is increasingly threatened by climate change, freshwater depletion, and human activities, leading to reduced water area and adverse impacts on biodiversity and local livelihoods. Fluctuations in temperature, precipitation, water levels, and discharge exacerbate these challenges. To support sustainable management, this study focuses on predicting water area changes using advanced machine learning models, evaluating their accuracy in capturing the complex interactions among key hydrological and climatic variables.

2. Materials and Methods

* Corresponding Author: Mehdi Nadi

Address: Department of Water Engineering, Sari Agricultural sciences and Natural Resources University, Sari, Iran. Email: mehdi.nadi@gmail.com Tel: 011-33687574

To achieve the study's objective, data from 2000 to 2023—including temperature, precipitation, discharge, water level, and MNDWI-derived water area—were collected from hydrological, climatic, and satellite sources. MNDWI was used to estimate the water area of Gorgan Bay due to its effectiveness in detecting water bodies. Advanced machine learning models such as XGBoost, RBFN, Random Forest, and Linear Regression were applied to predict water area changes, chosen for their ability to capture complex, nonlinear relationships. The relationships among variables were analyzed using Pearson correlation, statistical tests, and nonlinear regression. Model performance was assessed using RMSE, MAE, MAPE, MBE, and R^2 . Water level showed the strongest correlation with water area, while temperature and precipitation influenced it indirectly through discharge and evaporation. Model accuracy was validated by comparing predictions to actual data from five randomly selected years.

3. Results

The analysis showed that the Caspian Sea's water level had the strongest correlation with the water area of Gorgan Bay (Pearson = 0.90, $R^2 = 0.83$, $p = 0.001$), indicating that sea-level fluctuations are a key driver of changes in the Bay's extent. Similarly, the Caspian Sea's water area also showed a notable but weaker correlation ($r = 0.66$, $R^2 = 0.43$), suggesting its contribution to the Bay's drying. Inflow discharge from upstream sources had a moderate influence ($r = 0.63$, $R^2 = 0.40$), indicating its partial role in maintaining water levels. Precipitation and temperature showed the weakest correlations 0.43 and -0.34 respectively highlighting their limited direct impact. Precipitation may influence inflow, but its effect is moderated by evaporation and other dynamics. The negative correlation with temperature likely reflects increased evaporation, which reduces water coverage. Overall, while multiple variables affect the water area of Gorgan Bay, the Caspian Sea's water level remains the most significant factor. To assess the performance of the predictive models, five evaluation metrics were used: RMSE, MAE, MAPE, MBE, and R^2 . The XGBoost model emerged as the top-performing model, with an R^2 value of 0.93 and a MAPE of 4.9%. These results indicate that XGBoost was able to accurately capture the complex relationships between the input variables and the water area, providing precise predictions with minimal error. The Random Forest model, which is known for its robustness in handling high-dimensional data, came second with an R^2 of 0.76 and a MAPE of 8.08%. This suggests that Random Forest is a reliable model, but it does not perform as well as XGBoost in terms of accuracy. The Linear Regression model, a simpler algorithm, ranked third with an R^2 of 0.74 and a MAPE of 8.3%. While it is less complex and computationally expensive than other models, its performance was still satisfactory for predicting water area changes. The RBFN model, despite its ability to model nonlinear relationships, exhibited the weakest performance, with the highest systematic error. The RBFN model had an RMSE of 32.5, an MAE of 30.9, and a MAPE of 10.1%. This suggests that the model struggled to capture the underlying patterns in the data, possibly due to its inability to handle the complex interactions between the variables as effectively as other models. To further evaluate the models, five years (2000, 2008, 2011, 2016, and 2018) were randomly selected to predict the water area of Gorgan Bay. The results demonstrated that all models were capable of simulating the fluctuations in the water area, but the XGBoost model provided the most accurate predictions, particularly in predicting severe fluctuations. This is especially important for environmental planning, where predicting extreme events such as droughts or heavy rainfall is crucial for managing water resources. The XGBoost model's ability to accurately capture such fluctuations confirms its superiority in predicting the water area of Gorgan Bay.

4. Discussion and Conclusion

This study demonstrates the effectiveness of advanced machine learning models in predicting changes in the water area of Gorgan Bay. The results show that Caspian Sea water level changes are the most significant predictor, followed by inflow discharge. Precipitation and temperature have a weaker, more indirect influence. Among the tested models, XGBoost was the most accurate, highlighting the strength of gradient boosting in modeling complex

environmental data. These findings emphasize the value of machine learning in managing water resources and forecasting changes in sensitive ecosystems. Accurate predictions can support environmental planning and sustainable management, helping to reduce the impacts of climate change and human activities. This study also lays the groundwork for future research to incorporate more environmental variables and assess long-term climate effects on the region's hydrology.

5. Six important references

- 1) Kazempour, Z., Danesh-Yazdi, M., Asadifakhr, K., & Raie, M. (2023). Spatiotemporal dynamics of chlorophyll-a in the Gorgan Bay and Miankaleh Peninsula biosphere reserve: Call for action. *Remote Sensing Applications: Society and Environment*, 30, 100946. <https://doi.org/10.1016/j.rsase.2023.100946>
- 2) Li, L., Herrera, M., Mukherjee, A., Zheng, G., Chen, C., Dhada, M., Brice, H., Parekh, A., & Parlikad, A. K. (2025). Predictive alarm models for improving radio access network robustness. *Expert Systems with Applications*, 259, 125312. <https://doi.org/10.1016/j.eswa.2024.125312>
- 3) Li, S., Chen, F., Song, K., Liu, G., Tao, H., Xu, S., Wang, X., Wang, Q., & Mu, G. (2022). Mapping the trophic state index of eastern lakes in China using an empirical model and Sentinel-2 imagery data. *Journal of Hydrology*, 608, 127613. <https://doi.org/10.1016/j.jhydrol.2022.127613>
- 4) Nadi, M., & Yousefi Kebriya, A. (2024). A Method for Correction of Tropical Rainfall Measuring Mission Satellite Temperature Network in Mazandaran Province. *Iranica Journal of Energy and Environment (IJEE)*, 15(1), 100-110. <https://doi.org/10.5829/ijee.2024.15.01.10>
- 5) Riley, J. W., & Stillwell, C. C. (2023). Predicting inundation dynamics and hydroperiods of small, isolated wetlands using a machine learning approach. *Wetlands*, 43(63). <https://doi.org/10.1007/s13157-023-01706-2>
- 6) Yousefi kebriya, A., nadi, M., Ghanbariparmehr, E., & Zhongchang, S. (2024). Modeling and Forecasting Dust Pollution Using Satellite Data and Machine Learning Techniques: A Case Study of Ahvaz). *Journal of Agricultural Meteorology*, (). doi: 10.22125/agmj.2024.478677.1174

Conflict of Interest

Authors declared no conflict of interest.

Acknowledgments

This research is supported by research deputy of Sari Agricultural Sciences and Natural Resources University under grant number of: 06-1403-02. We would like to express our sincere gratitude for their support.

کاربرد الگوریتم‌های یادگیری ماشین در پیش‌بینی مساحت آبی خلیج گرگان با

بهره‌گیری از داده‌های هیدرولوژیکی و اقلیمی

علیرضا یوسفی کبری^۱، مهدی نادی^{۲*}

تاریخ ارسال: ۱۴۰۳/۱۰/۰۳

تاریخ پذیرش: ۱۴۰۴/۰۱/۳۰

مقاله پژوهشی

چکیده

خلیج گرگان یکی از زیست‌بوم‌های حساس و مهم در شمال ایران است که از لحاظ اکولوژیکی و اقتصادی از اهمیت بالایی برخوردار است. این مطالعه به پیش‌بینی تغییرات مساحت آبی خلیج گرگان با استفاده از داده‌های هیدرولوژیکی، ماهواره‌ای و اقلیمی شامل دما، بارش، دبی، تراز آب و مساحت آبی (MNDWI) دریای خزر در بازه زمانی ۱۳۷۸ تا ۱۴۰۱ پرداخت. به‌منظور مدل‌سازی این تغییرات، روش‌های پیشرفته یادگیری ماشین شامل XGBoost، شبکه عصبی RBFN، جنگل تصادفی و رگرسیون خطی استفاده شد. ارزیابی روابط میان متغیرها با استفاده از ضریب همبستگی و تحلیل رگرسیونی خطی انجام شد. شاخص‌های ارزیابی مدل شامل RMSE، MAE، MAPE، MBE و R^2 بوده و پنج سال به‌طور تصادفی داده‌های واقعی با داده‌های پیش‌بینی مقایسه شد. بر اساس نتایج، تغییرات تراز آبی دریای خزر بیشترین همبستگی را با مساحت آبی خلیج گرگان داشت (ضریب همبستگی ۰/۹۰ و معنی داری ۰/۰۰۱). ارتباط مساحت دریای خزر با خشک شدن خلیج گرگان ضریب همبستگی ۰/۶۶ و ضریب تبیین ۰/۴۳، و دبی ورودی با ضریب پیرسون ۰/۶۳ و ضریب تبیین ۰/۴۰ بود. بارش و دما با ضریب همبستگی ۰/۴۳ و ۰/۳۴- کمترین تأثیر را داشته است. نتایج ارزیابی مدل‌ها نشان داد که مدل XGBoost با ضریب تبیین ۰/۹۳ و خطای میانگین درصد مطلق ۰/۴۹، ریشه میانگین مربعات خطا ۱۶/۹ کیلومتر مربع و میانگین مطلق خطا ۱۵/۳ کیلومتر مربع بهترین عملکرد را داشت. مدل‌های جنگل تصادفی و رگرسیون خطی رتبه‌های دوم و سوم قرار گرفت. مدل RBFN با بیشترین خطای سیستماتیک و خطای میانگین درصدی مطلق ۱۰/۱٪ ضعیف‌ترین عملکرد را داشت. در نهایت، پنج سال به‌طور تصادفی از جمله سال‌های ۱۳۷۸، ۱۳۸۶، ۱۳۸۹، ۱۳۹۴ و ۱۳۹۶ برای پیش‌بینی مساحت آبی خلیج گرگان توسط مدل انتخاب شد. نتایج نشان داد که تمامی مدل‌ها توانستند نوسانات داده‌ها را به‌طور مناسب شبیه‌سازی کنند، اما مدل XGBoost بهترین عملکرد را داشت و نوسانات شدید مساحت آبی را دقیق‌تر پیش‌بینی کرد.

واژه‌های کلیدی: تراز آب دریای خزر، سنجش از دور، یادگیری ماشین، مدل‌سازی آماری، تغییرات مساحت آب

^۱ دکتری هواشناسی کشاورزی، گروه مهندسی آب، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ساری، ایران، alirezau3fikrbriya@gmail.com

^۲ دانشیار گروه مهندسی آب، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ساری، ایران، mehdi.nadi@gmail.com (نویسنده مسئول)

کمک می‌کند (Rezaei et al., 2019). به دلیل اهمیت زیست‌محیطی، شبه‌جزیره میانکاله و خلیج گرگان در سال ۱۹۷۵ به‌عنوان سایت رامسر و در سال ۱۹۷۶ به‌عنوان ذخیره‌گاه زیست‌کره یونسکو معرفی شده است. با این حال، در سال‌های اخیر، خلیج گرگان با کاهش شدید سطح آب مواجه شده است (Ranjbar et al., 2019). با توجه به تهدیدات جدی که تالاب‌های ساحلی به‌ویژه خلیج گرگان با آن مواجه می‌باشد، استفاده از روش‌های نوین مانند یادگیری ماشین می‌تواند به‌عنوان یک ابزار قدرتمند برای پیش‌بینی و مدیریت وضعیت این اکوسیستم‌ها در نظر گرفته شود (Yousefi kebriya et al., 2024). یادگیری ماشین قادر است به‌طور دقیق و کارآمد الگوهای پیچیده‌ای که در داده‌های اقلیمی و هیدرولوژیکی نهفته است را شناسایی کرده و پیش‌بینی‌هایی مبنی بر تغییرات سطح آب، کیفیت آب و وضعیت اکولوژیکی تالاب‌ها ارائه دهد (Jayathilake et al., 2023). از جمله مهم‌ترین پارامترهایی که در این پیش‌بینی‌ها تأثیرگذار هستند می‌توان به دما، بارش، دبی رودخانه‌ها، تراز آب و سایر داده‌های محیطی اشاره کرد (Ouma et al., 2022). این مدل‌ها با پردازش حجم زیادی از داده‌ها و استفاده از الگوریتم‌های پیچیده، می‌توانند روندهای تغییرات را شبیه‌سازی کرده و راهکارهای مدیریتی مبتنی بر پیش‌بینی‌های دقیق‌تری را ارائه دهد (Rane et al., 2024). پیرعلی زفره‌ئی و همکاران (۱۳۹۸) در پژوهشی به بررسی قابلیت مدل ماشین بردار پشتیبان (SVM) در شبیه‌سازی پارامترهای کیفی آب در تالاب بین‌المللی چغاخور طی یک دوره ۲۳ ساله پرداختند. در این مطالعه، پارامترهای کلروفیل a و شفافیت آب (عمق سشی دیسک) به‌عنوان شاخص‌های کیفی آب مورد بررسی قرار گرفتند. داده‌ها به دو مجموعه آموزشی (۷۰ درصد) و آزمایشی (۳۰ درصد) تقسیم شدند. نتایج ارزیابی مدل با استفاده از ضریب تبیین (R^2) و خطای RMSE نشان داد که SVM در پیش‌بینی این پارامترها عملکرد مناسبی دارد ($R^2=0.73$) و RMSE برای کلروفیل a برابر ۷.۳۴ میلی‌گرم بر مترمکعب

مقدمه

تالاب‌های ساحلی حدود ۱۲ درصد از خط ساحلی جهان را تشکیل می‌دهند (Stein et al., 2021). این اکوسیستم‌ها خدمات اکولوژیکی مهمی ارائه می‌دهند، از جمله ایجاد زیستگاه برای گونه‌های مختلف، حمایت از ماهیگیری، حفاظت در برابر سیلاب‌ها و جذب گردشگران (Zheng et al., 2021). این آب‌ها معمولاً کم‌عمق هستند و از طریق یک یا چند کانال به منابع آبی بزرگ‌تری متصل می‌شوند و معمولاً با سازه‌های طبیعی مانند جزایر بارری، شبه‌جزیره‌ها یا هرم‌های شنی از دریا یا اقیانوس جدا می‌شوند (Li et al., 2022). با رشد سریع شهرنشینی، صنعتی شدن و گسترش فعالیت‌های کشاورزی، تالاب‌ها با آلودگی‌های ناشی از منابع نقطه‌ای و غیرنقطه‌ای مواجه شده‌اند که این امر چالش‌های فزاینده‌ای را در سطح جهانی به وجود آورده است. این تغییرات منجر به آلودگی شدید و فرآیند اوتروفیکاسیون^۳ (افزایش بیش از حد مواد مغذی مانند نیترژن و فسفر در منابع آبی و رشد غیرطبیعی جلبک‌ها) در بسیاری از اکوسیستم‌های آبی شده است. ساختار و ویژگی‌های اکولوژیکی این اکوسیستم‌ها همچنین به نوسانات سطح آب منابع آبی متصل به آن‌ها بسیار حساس است (Kazempour et al., 2023). با توجه به اهمیت بالای تالاب‌های ساحلی در حفظ تنوع زیستی و تأمین خدمات اکولوژیکی، مدیریت صحیح این اکوسیستم‌ها برای مقابله با تهدیدات ناشی از تغییرات اقلیمی، افزایش آلودگی و کاهش سطح آب ضروری است. خلیج گرگان، که یک تالاب ساحلی کم‌عمق در شمال ایران است، بین شبه‌جزیره میانکاله و دریای خزر قرار دارد. این خلیج خدمات اکولوژیکی مهمی را با فراهم آوردن زیستگاه‌های مناسب برای انواع گونه‌های آبی و پرندگان مهاجر ارائه می‌دهد. همچنین، خلیج گرگان از نظر اقتصادی به‌ویژه از طریق فعالیت‌های ماهیگیری و گردشگری به اقتصاد محلی

³. Eutrophication

داشتند. نتایج نشان داد که ۷۵.۷٪ از سطح منطقه در طبقه توان زیاد و ۶۶.۳۸٪ در طبقه توان خیلی زیاد قرار دارند. در مطالعه‌ای به نقشه‌برداری از تالاب‌های جنوب شرقی خلیج جورجیان^۴ پرداخته شد و از ترکیب تکنیک‌های یادگیری ماشین (مانند جنگل تصادفی و ماشین‌های بردار پشتیبان) با تحلیل تصاویر مبتنی بر شیء استفاده شد. الگوریتم‌های یادگیری ماشین با داده‌های موجود از منابع آزاد (تصاویر سنتینل ۲)^۵ آموزش داده شد و دقت آن‌ها بر روی یک مجموعه داده ذخیره‌شده، آزمایش شد. نتایج نشان داد که دقت‌های تولیدکننده ۸۵ درصد و ۷۹ درصد و دقت‌های کاربری ۷۸ درصد و ۸۴ درصد برای مدل‌های جنگل تصادفی و ماشین‌های بردار پشتیبان به ترتیب به دست آمده است (Luymes et al., 2021). در مطالعه‌ای از تکنیک‌های پیشرفته جغرافیایی و یادگیری ماشین برای ارزیابی آسیب‌پذیری‌های دریاچه کینجیر در پاکستان استفاده کرده است. داده‌های سنجش از دور چندزمانه از سال ۲۰۰۰ تا ۲۰۲۰ برای بررسی تغییرات پوشش زمین، تغییرات مساحت و تهدیدات محیطی تجزیه و تحلیل شد. الگوریتم جنگل تصادفی با دقت ۸۹.۵ درصد بهترین عملکرد را در طبقه‌بندی پوشش زمین نشان داد. نتایج نشان داد که از سال ۲۰۰۰، ۱۱ درصد از آب‌های آزاد دریاچه کاهش یافته و ۳۰ درصد از مناطق کنونی دریاچه در معرض آسیب‌پذیری متوسط تا بالا قرار دارند. مدل CA-Markov پیش‌بینی کرد که به دلیل فشارهای انسانی مانند رشد جمعیت ۲۹ میلیون نفری اطراف دریاچه، کاهش‌ها ادامه خواهد داشت (Aslam et al., 2024). در تحقیق انجام‌شده در رودخانه مدا، مالزی، سه الگوریتم یادگیری ماشین شامل شبکه عصبی چندلایه (MLP-NN)، شبکه عصبی حافظه بلندمدت کوتاه‌مدت (LSTM) و تقویت گرادیان شدید (XGBoost) برای پیش‌بینی سطح آب مورد بررسی قرار گرفتند. داده‌های روزانه سطح آب و هواشناسی برای دوره ۲۰۱۶ تا ۲۰۱۸ به کار گرفته شدند. نتایج نشان داد مدل MLP با ضریب تبیین ۰.۸۷۱ عملکرد بهتری نسبت به LSTM (۰.۸۶۵) و XGBoost (۰.۸۳۱) داشته است. مدل LSTM در پیش‌بینی بلندمدت (۷ روزه)

و برای شفافیت آب ۱.۳۷ متر گزارش شد). در مطالعه‌ای توسط تراهی و همکاران (۱۳۹۶) به ارزیابی عملکرد الگوریتم‌های ماشین بردار پشتیبان (SVM) و حداکثر احتمال در طبقه‌بندی کاربری جنگل‌های رودخانه‌ای مارون با استفاده از تصاویر ماهواره‌ای لندست ۸ پرداخت. داده‌ها پس از پیش‌پردازش شامل تصحیحات رادیومتریک و اتمسفری، با دو مجموعه کلاس‌بندی شامل هفت کلاس (جنگل، مرتع، کشاورزی، آیش، رودخانه، مسکونی و جاده) و سه کلاس (جنگل، رودخانه و سایر مناطق) مورد تحلیل قرار گرفتند. برای ارزیابی الگوریتم SVM، سه مجموعه نمونه تعلیمی با تعداد ۲۴۱، ۱۴۱ و ۴۱ نمونه و چهار نوع هسته (خطی، چندجمله‌ای، شعاعی و حلقوی) استفاده شد. نتایج نشان داد که بالاترین دقت با صحت کلی ۹۹.۲۴٪ و ضریب کاپای ۰.۹۷ مربوط به الگوریتم SVM با هسته چندجمله‌ای در باندهای اصلی سنجنده OLI و در طبقه‌بندی سه‌کلاسه است. همچنین، کاهش تعداد کلاس‌ها از هفت به سه موجب افزایش دقت شد، اما کاهش بیش از حد تعداد نمونه‌ها، دقت طبقه‌بندی را کاهش داد. سلامت و همکاران (۱۴۰۲) در پژوهشی به تحلیل توان‌یابی آب زیرزمینی در منطقه‌ی بجنورد پرداختند. عوامل مؤثر در این تحلیل شامل بلندی، شیب، جهت، شاخص رطوبت پستی‌بلندی، فاصله از آبراهه، تراکم زهکشی، فاصله از گسل، سنگ‌شناسی، شاخص موقعیت پستی‌بلندی، شاخص ناهمواری زمین، موقعیت شیب نسبی و شاخص هم‌گرایی جریان بودند. داده‌های موقعیت ۳۵۹ چشمه از شرکت آب منطقه‌ای دریافت شد. برای تقسیم‌بندی داده‌ها، ۷۰٪ به عنوان نقطه‌های آموزشی و ۳۰٪ به عنوان نقطه‌های اعتبارسنجی با استفاده از دستورالعمل تقسیم‌بندی تصادفی اختصاص داده شد. نتایج ارزیابی دقت مدل‌ها نشان داد که مدل ترکیبی SVM و دستورالعمل بهینه‌سازی فراکاوچی ازدحام ذرات (SVM-PSO) با دقت ACU برابر ۰.۹۴ از دیگر مدل‌ها مانند SVM (۰.۹۱۸) و SVM-BA (۰.۹۳۲) عملکرد بهتری دارد. در ارزیابی حساسیت، موقعیت شیب نسبی با ۱۴.۵٪، فاصله از گسل با ۱۳.۴٪ و سنگ‌شناسی با ۱۲.۳٪ بالاترین اهمیت را در پیش‌بینی توان آب زیرزمینی

^۴ . Georgian Bay

^۵ . Sentinel-2

سال‌های اخیر به‌طور قابل‌توجهی تحت تأثیر تغییرات اقلیمی و فعالیت‌های انسانی قرار گرفته است. کاهش شدید سطح آب این خلیج، که از سال ۱۹۹۵ تا ۲۰۲۰ حدود ۱۹۰ کیلومتر مربع از مساحت آن کاهش یافته، بیانگر عمق بحران زیست‌محیطی موجود است (Sharbaty., 2020). این تغییرات زیست‌محیطی و افت کیفیت آب، نه تنها به‌عنوان یک چالش اساسی زیست‌محیطی در کشور مطرح شده بلکه تأثیرات زیادی بر تنوع زیستی و معیشت جوامع محلی دارد. در این راستا، این مطالعه به‌منظور پیش‌بینی تغییرات مساحت آبی خلیج گرگان، از داده‌های اقلیمی و هیدرولوژیکی شامل دما، بارش، دبی، تراز آب دریای خزر و مساحت دریای خزر استفاده کرده است. برای مدل‌سازی این فرآیند، از روش‌های پیشرفته یادگیری ماشین شامل XGBoost، RBFN، جنگل تصادفی و رگرسیون خطی بهره گرفته شده و همچنین داده‌های ماهواره‌ای شامل تراز آب دریای خزر و مساحت آبی دریای خزر به‌عنوان ورودی‌های مدل در نظر گرفته شده است. برخلاف مطالعات پیشین که به بررسی تأثیر هر یک از عوامل اقلیمی و هیدرولوژیکی بر خشکی خلیج گرگان پرداخته است، این پژوهش بر مدل‌سازی جامع و پیش‌بینی دقیق‌تر مساحت آبی خلیج گرگان با استفاده از روش‌های نوین یادگیری ماشین تمرکز دارد. نتایج حاصل از این مدل‌سازی می‌تواند در برنامه‌ریزی‌های زیست‌محیطی، مدیریت منابع آبی و اتخاذ تصمیمات مؤثر در راستای حفظ و احیای خلیج گرگان بسیار مفید باشد.

مواد و روش

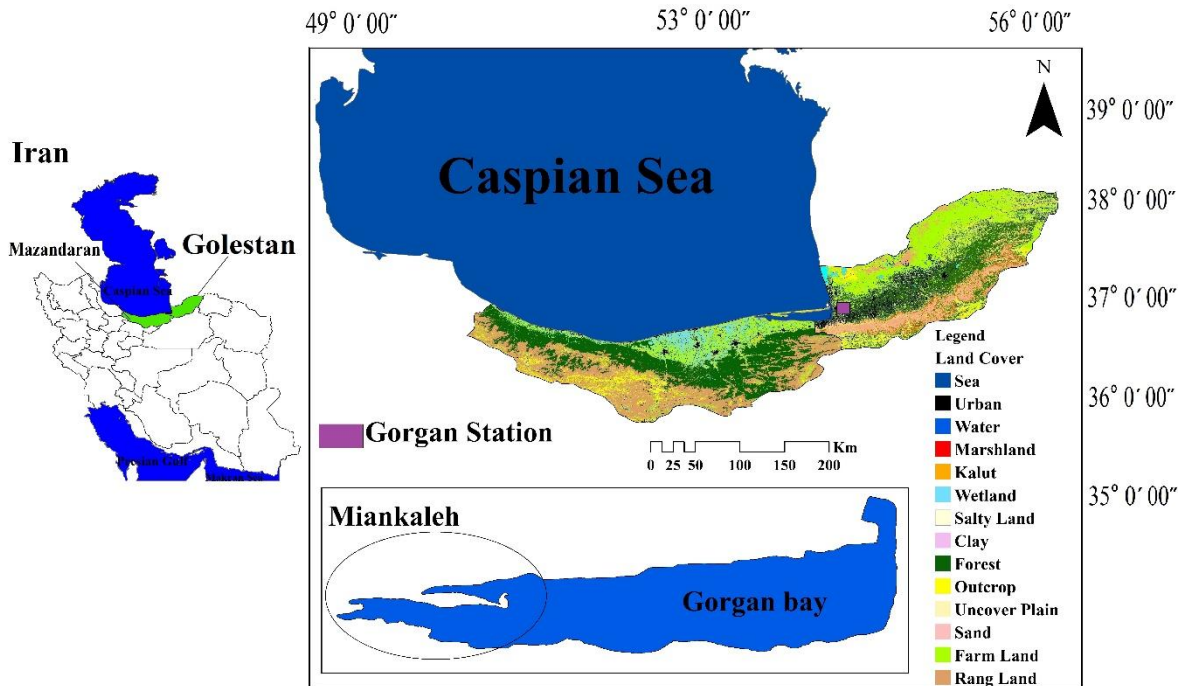
منطقه مورد مطالعه

خلیج گرگان عرض جغرافیایی $36^{\circ} 48'$ تا $36^{\circ} 57'$ شمالی و طول جغرافیایی $53^{\circ} 25'$ تا $54^{\circ} 2'$ شرقی در بخش جنوب‌شرقی دریای خزر جای گرفته است. این خلیج از شمال به شبه‌جزیره میانکاله و از جنوب به استان‌های گلستان و مازندران محدود می‌شود، که در شکل ۱ قابل مشاهده است. اقلیم معتدل دریای خزر بر این منطقه حاکم است و میانگین دمای سالانه آن حدود ۱۸ درجه سانتی‌گراد

نسبت به سایر مدل‌ها برتری نشان داد، در حالی که XGBoost به‌عنوان سریع‌ترین الگوریتم معرفی شد. این مطالعه بر تنوع عملکرد مدل‌ها و وابستگی آن‌ها به کیفیت و کمیت داده‌های آموزشی تأکید دارد (Zakaria et al., 2023). در مطالعه‌ای در منطقه‌ی پناهگاه ملی حیات وحش سنت مارکس پان هندل فلوریدا در امتداد ساحل خلیج مکزیک، روشی برای پیش‌بینی دینامیک روزانه غرقاب با استفاده از یک مجموعه داده بزرگ از سطح آب تالاب‌ها و الگوریتم جنگل تصادفی ارائه شد و از متغیرهایی استفاده شد که ویژگی‌های مورفولوژی حوضه و برآوردهای بودجه آبی جوی را در دوره‌های مختلف گذشته توصیف می‌کرد. عملکرد مدل با استفاده از دو معیار، یعنی هیدروپریود میانه و درصد روزهای پیش‌بینی‌شده به‌درستی، ارزیابی شد. مدل‌ها عملکرد کلی خوبی داشت و دقت متوسط ۸۳ درصد بر روی داده‌های اعتبارسنجی گزارش شد (Riley & Stillwell, 2023). در پژوهشی، برای پیش‌بینی تغییرات حجم ذخیره آب این تالاب‌ها در کانادا، از یک رویکرد ترکیبی یادگیری ماشین و مدل‌سازی هیدرولوژیکی استفاده شد. این مطالعه شامل استفاده از مدل‌های یادگیری ماشین برای تخمین WSV پایه با استفاده از داده‌های عمق آب و اطلاعات جغرافیایی، و شبیه‌سازی تعادل آبی با ابزار SWAT برای دوره ۲۰۱۸-۲۰۳۴ بود. مدل‌های مختلف یادگیری ماشین شامل رگرسیون خطی، درخت‌های رگرسیون، ماشین‌های بردار پشتیبانی (SVM)، درخت‌های تجمیعی و رگرسیون فرایند گاوسی (GPR) بکارگرفته شد. این رویکرد ترکیبی به‌ویژه در شبیه‌سازی تغییرات WSV تالاب‌ها تحت دو سناریو اقلیمی مختلف (RCP 2.6 و RCP 8.5) موثر واقع شد (Cui et al., 2021). در مجموع، این تحقیقات نشان‌دهنده نقش مؤثر روش‌های یادگیری ماشین در پیش‌بینی منابع آبی، تالاب‌ها و اکوسیستم‌های آبی همچون خلیج‌ها و دریاچه‌ها می‌باشد. این رویکردهای پیشرفته نه تنها قادر به ارزیابی دقیق‌تر تغییرات منابع آبی در شرایط مختلف اقلیمی و محیطی هستند، بلکه ابزارهایی برای پیش‌بینی دقیق‌تر وضعیت این منابع در آینده ارائه می‌دهد. خلیج گرگان، به‌عنوان یکی از زیست‌بوم‌های حساس و استراتژیک در شمال ایران، در

کیلومتر در راستای شرقی-غربی امتداد دارد و بیشترین پهنای آن به ۱۲ کیلومتر می‌رسد و عمق آن به‌طور تقریبی بین ۱ تا ۶ متر متغیر است (Kazempour et al., 2023).

است. میزان بارندگی سالانه خلیج برابر با ۱۹۵ میلیون مترمکعب و میزان تبخیر سالانه آن ۵۵۲ میلیون مترمکعب تخمین زده شده است. این خلیج به طول تقریبی ۶۰



شکل (۱): کاربری اراضی استان گلستان و استان مازندران و موقعیت جغرافیایی خلیج گرگان

می‌باشد، که به دلیل دقت و پوشش زمانی و مکانی مناسب، امکان تحلیل دقیق‌تری از تغییرات زمانی و مکانی این پهنه‌های آبی را فراهم می‌کند.

سنجش از دور

ماهواره‌های مودیس و سنتینل-۲ به عنوان ابزارهای پیشرفته و حیاتی در پایش محیط زیست و تغییرات اقلیمی شناخته می‌شوند. مودیس، که توسط NASA و NOAA طراحی شده است، با ارائه داده‌هایی با وضوح متوسط، نقش مهمی در پایش دما، تغییرات پوشش گیاهی، برف، گردوغبار، و دیگر شاخص‌های زیست‌محیطی ایفا می‌کند (NOAA, 2020). سنتینل-۲، توسعه‌یافته توسط آژانس فضایی اروپا (ESA)، با ارائه تصاویر با وضوح بالا (تا ۱۰ متر) امکان پایش دقیق تغییرات سطح زمین، مدیریت منابع طبیعی، و نظارت بر تغییرات کاربری اراضی را فراهم می‌آورد (Lorente et al., 2021). علاوه بر این، ماهواره

داده‌های مورد استفاده

در این پژوهش، به منظور بررسی ایجاد مدل‌های پیش‌بینی مساحت آبی خلیج گرگان از پارامترهای دما، بارش، دبی ورودی به خلیج گرگان، مساحت دریای خزر و تراز آب دریای خزر در بازه زمانی ۲۳ ساله و از دوره‌ی زمانی ۱۳۷۸ تا ۱۴۰۱ استفاده شد. پارامترهای هواشناسی شامل دما و بارش ایستگاه هواشناسی سینوپتیک گرگان استفاده شد که در شکل ۱ موقعیت جغرافیایی ایستگاه مشخص است. اطلاعات مربوط به دبی رودخانه‌های منتهی به خلیج گرگان، از سازمان‌های آب منطقه‌ای استان‌های گلستان و مازندران به دست آمد. برای بررسی مساحت پهنه آبی خلیج گرگان و دریای مازندران، از شاخص MNDWI ماهواره‌ی مودیس و سنتینل-۲ استفاده شد و تراز آب دریای مازندران از داده‌های جیسون-۳، استفاده شد. داده‌ها شامل تصاویر و اطلاعات ماهواره‌های مودیس^۶، سنتینل-۲^۷، جیسون-۳^۸

^۸. Jason 3

^۶. MODIS (Moderate Resolution Imaging Spectroradiometer)

^۷. Sentinel 2

مسائل پیش‌بینی عملکرد برجسته‌ای دارد (Li et al., 2025). مدل XGB برای پیش‌بینی مساحت آبی خلیج گرگان با تنظیمات خاص پیاده‌سازی شد. این مدل شامل ۱۰۰۰ درخت تصمیم‌گیری است و با نرخ یادگیری ۰.۰۱ و عمق درخت ۶ آموزش داده می‌شود. برای بهبود عملکرد و جلوگیری از بیش‌برازش (overfitting)، از قابلیت Early Stopping استفاده شد. به این صورت که آموزش مدل پس از ۱۰ دوره بدون بهبود متوقف شد. این ویژگی با استفاده از معیار خطای مجموعه آزمایشی (test error) پیاده‌سازی شد تا مطمئن شویم که مدل بیش از حد به داده‌های آموزشی تطبیق پیدا نمی‌کند.

^{۱۳} RBFN نوعی شبکه عصبی است که با استفاده از توابع شعاعی برای نگاشت داده‌ها به فضای ویژگی‌های جدید، الگوهای پیچیده و غیرخطی را مدل‌سازی می‌کند. این روش در شناسایی الگوهای چندبعدی کاربرد دارد. شبکه عصبی با توابع پایه شعاعی، این مدل دارای یک لایه مخفی با ۱۰ نرون است که از تابع فعال‌سازی tanh (تانژانت هیپربولیک) برای لایه‌های مخفی استفاده می‌کند. برای آموزش مدل از الگوریتم L-BFGS (یک روش بهینه‌سازی مبتنی بر گرادین) استفاده شد که بهینه‌سازی پارامترهای مدل را انجام می‌دهد. این مدل به مدت ۱۰۰۰ تکرار آموزش داده شد. پس از آموزش، مدل قادر به پیش‌بینی مقادیر بر اساس ویژگی‌های ورودی بود. همچنین، به منظور تکرارپذیری نتایج، از مقدار ثابت random_state=42 استفاده شد.

جنگل تصادفی^{۱۴} از مجموعه‌ای از درخت‌های تصمیم برای پیش‌بینی استفاده می‌کند (Al-Selwi et al., 2024). این روش به دلیل قابلیت کاهش واریانس و مقاومت در برابر داده‌های نویزی، یکی از ابزارهای محبوب در مسائل طبقه‌بندی و رگرسیون است. جنگل تصادفی، این مدل با استفاده از ۱۰۰ درخت تصمیم‌گیری پیاده‌سازی شد. جنگل تصادفی به‌طور خودکار ویژگی‌های مختلف ورودی را تجزیه و تحلیل کرده و مدل‌های تصمیم‌گیری را بر اساس داده‌های آموزشی ایجاد می‌کند. همچنین، از پارامتر

جیسون-۳، به عنوان یکی از ابزارهای پیشرفته در سنجش از دور، برای اندازه‌گیری تراز آب دریاها و ارزیابی تغییرات سطح آب اقیانوس‌ها طراحی شده و اطلاعات ارزشمندی برای مطالعات اقلیمی و مدیریت منابع آب ارائه می‌دهد. این ماهواره‌ها با داده‌های دقیق و جامع، پشتیبان کلیدی تحقیقات علمی و برنامه‌ریزی‌های محیط زیستی به شمار می‌روند (Esri, 2018).

شاخص MNDWI

شاخص MNDWI^۹ برای تشخیص مناطق آبی در تصاویر ماهواره‌ای و تصاویر سنجش از دور استفاده می‌شود. این شاخص از اختلاف بین باندهای سبز^{۱۰} و مادون قرمز میانه^{۱۱} استفاده می‌کند و به عنوان یک شاخص برای تفکیک آب از سایر اشیاء مانند خاک و گیاهان استفاده می‌شود و با استفاده از رابطه (۱) محاسبه می‌شود. در این رابطه، باند سبز بیانگر شدت نور در باند سبز تصویر است و مادون قرمز میانه بیانگر شدت نور در باند مادون قرمز میانه تصویر است. این شاخص معمولاً مقادیر مثبت برای آب و مقادیر منفی برای سایر اشیاء نظیر خاک و گیاهان تولید می‌کند. به عبارت دیگر، مناطق آبی در تصویر با مقادیر مثبت MNDWI نشان داده می‌شوند، در حالی که سایر اشیاء با مقادیر منفی نشان داده می‌شوند (Gautama et al., 2015).

$$MNDWI = \frac{Green - SWIR}{Green + SWIR} \quad (1)$$

روش‌های یادگیری ماشین

در این پژوهش از روش‌های یادگیری ماشین شامل XGBoost، RBFN، جنگل تصادفی و رگرسیون خطی برای مدل‌سازی پیش‌بینی مساحت آبی خلیج گرگان استفاده شد، که در شکل ۲ مشاهده می‌شود. XGB^{۱۲} یک الگوریتم تقویتی گرادین کارآمد و سریع است که برای داده‌های بزرگ و پیچیده مناسب بوده و به دلیل قابلیت کنترل بیش‌برازش و تنظیم وزن ویژگی‌ها، در بسیاری از

¹³ . Radial Basis Function Network

¹⁴ . Random Forest

⁹ . Modified Normalized Difference Water Index

¹⁰ . Green

¹¹ . Mid-Infrared

¹² . Extreme Gradient Boosting

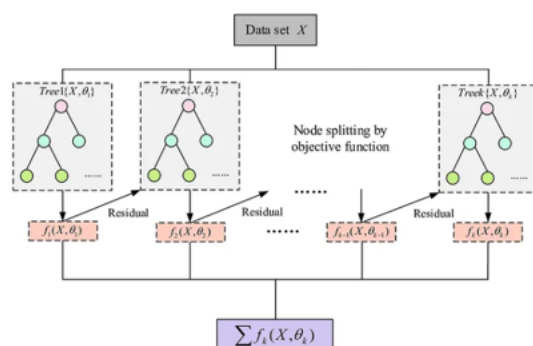
داده random_state=42 برای تضمین تکرارپذیری نتایج استفاده شد. پس از آموزش، پیش‌بینی‌هایی برای مجموعه داده آزمایشی انجام شد.

رگرسیون خطی^{۱۵} یک روش پایه‌ای و پراستفاده برای مدل‌سازی روابط خطی میان متغیرها است. این روش به دلیل سادگی و تفسیرپذیری بالا همچنان در بسیاری از تحلیل‌ها کاربرد دارد (Ganaie et al., 2022). در مدل رگرسیون خطی، این مدل با استفاده از کلاس LinearRegression از کتابخانه

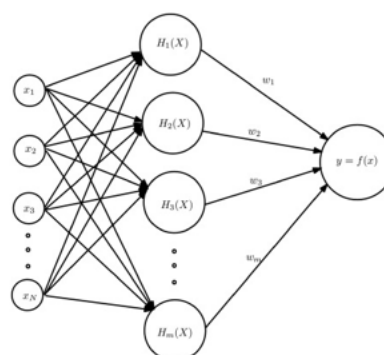
این روش‌ها با بهره‌گیری از مزایای منحصر به فرد خود، امکان پیش‌بینی دقیق‌تر و تحلیل جامع‌تر از داده‌ها را فراهم می‌کند. مدل‌سازی‌ها با استفاده از زبان برنامه‌نویسی Python پیاده‌سازی شد.

از کتابخانه LinearRegression از

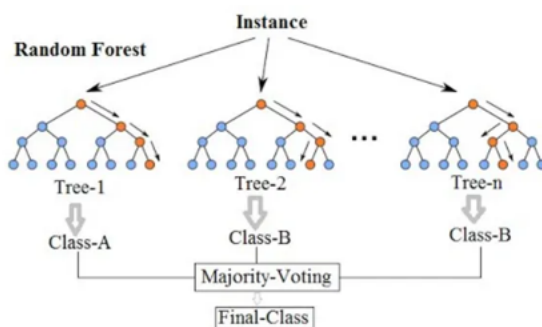
A) XGBoost



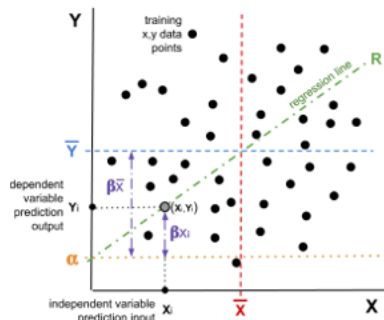
B) RBFN



C) Random Forest



D) Linear Regression



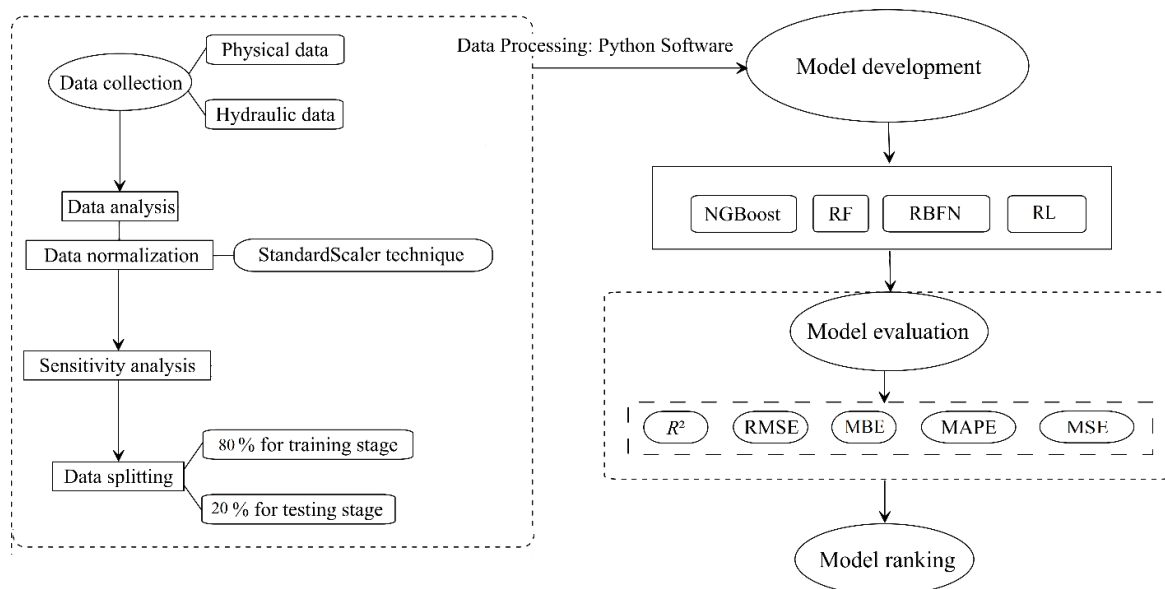
شکل (۲): مدل‌های پیش‌بینی XGBoost, RBFN, جنگل تصادفی و رگرسیون خطی

متغیر افزایش (کاهش) یابد، دیگری نیز همزمان افزایش (کاهش) می‌یابد (Nadi and Yousefi Kebriya, 2024). شاخص p-value نشان‌دهنده احتمال رد نکردن فرضیه‌ی صفر است. اگر این شاخص کوچک باشد (معمولاً کمتر از ۰.۰۵)، این به این معنی است که داده‌ها به اندازه کافی قوی می‌باشد تا فرضیه‌ی صفر رد شود (Nadi & Yousefi Kebriya., 2024). شاخص‌های خطا شامل

شاخص ارزیابی

برای بررسی ارتباط بین داده‌ها از ضریب همبستگی پیرسون (r) و آزمون فرضیات آماری (p-value)، استفاده شد. ضریب همبستگی پیرسون بین بازه ۱- و ۱ قرار می‌گیرد. مقدار r برابر با ۱ نشان‌دهنده یک رابطه خطی مستقیم کامل بین دو متغیر است، به این معنی که اگر یک

رگرسیون خطی برای مدلسازی پیش‌بینی مساحت آبی خلیج گرگان استفاده می‌شود که در شکل ۳ مراحل مختلف این مطالعه قابل مشاهده است. برای ارزیابی کارایی مدل‌ها، نتایج پنج سال مختلف به صورت تصادفی با داده‌های واقعی مقایسه می‌شود. این مطالعه با استفاده از داده‌های ماهواره‌ای و تکنیک‌های مدلسازی پیشرفته، دقت پیش‌بینی مساحت خلیج گرگان را به طور قابل توجهی افزایش می‌دهد. نتایج به دست آمده می‌تواند نقشی کلیدی در ارتقای مدیریت منابع طبیعی، پیش‌بینی تغییرات زیست‌محیطی و اتخاذ تدابیر پیشگیرانه مؤثر در حفاظت از اکوسیستم‌های آبی ایفا کند.



شکل (۳): فلوچارت مراحل مختلف روش‌شناسی پژوهش

شده است. نتایج تحلیل‌ها که در شکل ۴ و ۵ ارائه شده‌اند، نشان می‌دهند که تغییرات تراز و مساحت آبی دریای مازندران به همراه دبی ورودی به خلیج گرگان، تأثیر معنادار و قابل توجهی بر تغییرات مساحت آبی خلیج داشته‌اند. این تأثیرات در سطح اطمینان ۰.۰۱ تأیید شده‌اند، در حالی که متغیرهای دما و بارش در سطح ۰.۰۵ معنادار بوده و اثر کمتری بر تغییرات مساحت خلیج نشان داده‌است.

نتایج و بحث

تجزیه و تحلیل آماری

در این پژوهش، به منظور ارزیابی و تحلیل عوامل مؤثر بر پیش‌بینی تغییرات مساحت آبی خلیج گرگان^{۲۰}، از مجموعه‌ای جامع از داده‌های هواشناسی (دما (T) و بارش (R))، اطلاعات ماهواره‌ای مرتبط با تراز (Level) و مساحت آبی دریای مازندران^{۲۱} (ACS)، و آمار مربوط به دبی ورودی (Q) به خلیج گرگان طی بازه زمانی ۱۳۸۷ تا ۱۴۰۲ استفاده

19 . Mean Absolute Percentage Error - MAPE

20 . Area Gorgan Bay

21 . Area Caspian Sea

16 . Mean Absolute Error - MAE

17 . Mean Bias Error - MBE

18 . Root Mean Square Error - RMSE

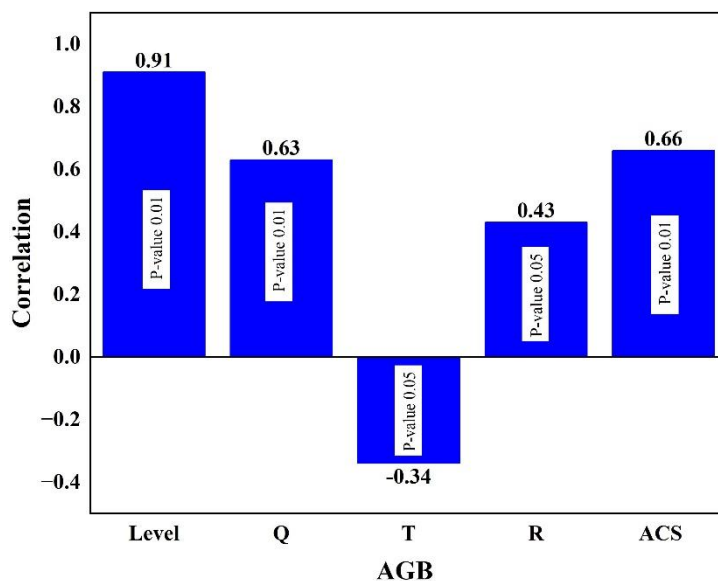
برای بررسی بیشتر پارامترهای پیش‌بینی و باتوجه به اینکه پارامترهای تراز آبی دریای خزر و مساحت دریای خزر برای پیش‌بینی مساحت آبی خلیج گرگان نسبت به سایر پارامترهای بکار گرفته شده، قوی تر بوده است، نمودار برازش رگرسیونی خطی آن‌ها رسم شد. بر اساس نتایج تجزیه و تحلیل داده‌ها و بررسی رابطه بین مساحت آبی خلیج گرگان و تراز آب و مساحت آبی دریای خزر، همان‌طور که در شکل ۵ نشان داده شده است، ارتباطی غیرخطی و درجه دوم بین متغیرها مشاهده می‌شود و مدل رگرسیونی نیز ارائه شده است. ضریب تبیین تعدیل‌شده (Adjusted R²) این مدل به ترتیب برابر با ۰.۸۳ و ۰.۴۳ است، که نشان‌دهنده دقت بالای مدل در تبیین بیش از ۸۳٪ و ۴۳٪ از تغییرات مساحت آبی خلیج گرگان بر اساس تغییرات تراز آبی و مساحت آبی دریای خزر می‌باشد، که معادله آن نیز در شکل ۴ آورده شده است. همچنین، باند اطمینان ۹۵ درصد (Confidence Band ۹۵٪) که به رنگ آبی مشخص شده است، نشان می‌دهد که مدل به‌خوبی داده‌های پراکنده را پوشش می‌دهد و داده‌های واقعی در بیشتر موارد در این محدوده قرار گرفته‌اند. این تطابق بیانگر دقت و قابلیت اعتماد مدل رگرسیونی ارائه شده می‌باشد. به طور کلی، نتایج این تحلیل، رابطه قوی بین تراز آب و مساحت آبی دریای خزر با مساحت آبی خلیج گرگان را تأیید می‌کند.

بر اساس نتایج شکل ۴ و جدول ۱، مشخص شد که بیشترین همبستگی مربوط به تغییرات تراز آبی دریای مازندران وجود داشته است؛ به‌گونه‌ای که ضریب پیرسون بیش از ۰/۹۰ و ضریب تبیین R² برابر با ۰/۸۳ بوده است، که در شکل ۵ نیز آورده شده است. این نشان‌دهنده رابطه مستقیم و قوی بین افزایش یا کاهش تراز آب دریای مازندران و تغییرات مساحت خلیج گرگان است، به‌طوری که با افزایش تراز آب دریای مازندران، مساحت خلیج نیز افزایش یافته است. در بررسی سایر پارامترها، ارتباط آن‌ها با مساحت آبی خلیج گرگان نسبت به تراز آب دریای مازندران ضعیف‌تر ارزیابی شد. برای نمونه، تغییرات مساحت آبی دریای مازندران با خشک شدن خلیج گرگان ضریب پیرسون ۰/۶۶ را نشان داده است، که در شکل ۴ آورده شده است. از سوی دیگر، دبی ورودی به خلیج گرگان با ضریب پیرسون ۰/۶۳ تأثیر نسبی قابل توجهی داشته است. اما میزان بارش و دما، اثرگذاری کمتری بر مساحت آبی خلیج داشته‌است، به‌طوری که ضریب پیرسون برای میزان بارش برابر ۰/۴۳ محاسبه شد. نهایتاً، کمترین میزان ارتباط در این پژوهش مربوط به دما بود که با ضریب پیرسون ۰/۳۴- نشان داد هر چه دما افزایش یابد، مساحت آب خلیج کاهش می‌یابد. این یافته‌ها تأیید می‌کنند که از میان عوامل بررسی شده، تغییرات تراز آبی دریای مازندران بیشترین نقش را در تغییرات مساحت خلیج گرگان ایفا می‌کند، در حالی که دبی ورودی و سایر عوامل نظیر بارش و دما اثرگذاری کمتری داشته‌اند. بنابراین، این پارامترها ابزارهای کلیدی برای پیش‌بینی تغییرات آینده مساحت آبی خلیج گرگان به شمار می‌آید.

جدول (۱): همبستگی پارامترهای مورد بررسی با مساحت خلیج گرگان

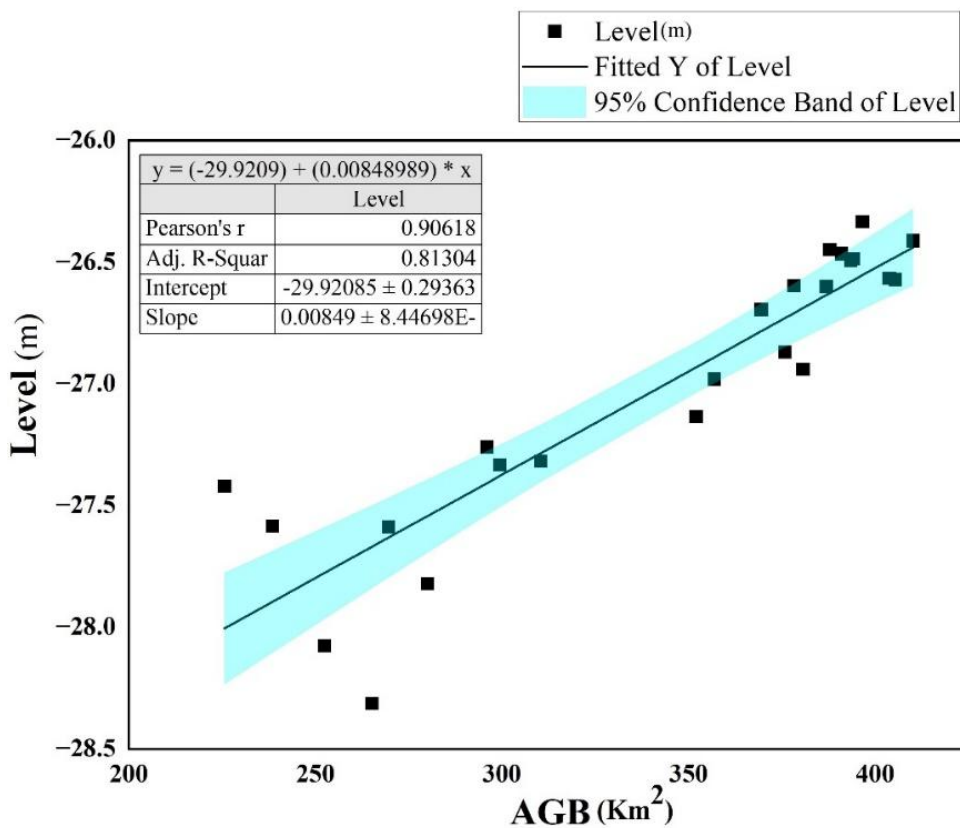
پارامتر	پیرسون	P-value	ضریب تبیین
تراز آب دریای خزر	۰/۹۱	۰/۰۰۰۱	۰/۸۲
دبی ورودی	۰/۶۳	۰/۰۰۱	۰/۴
دما	-۰/۳۴	۰/۰۶۴	۰/۱۱
بارش	۰/۴۳	۰/۰۳	۰/۱۹
مساحت دریا خزر ^{۲۲}	۰/۶۶	۰/۰۰۰۱	۰/۴۳

^{۲۲}. Area Caspian Sea (ACS)

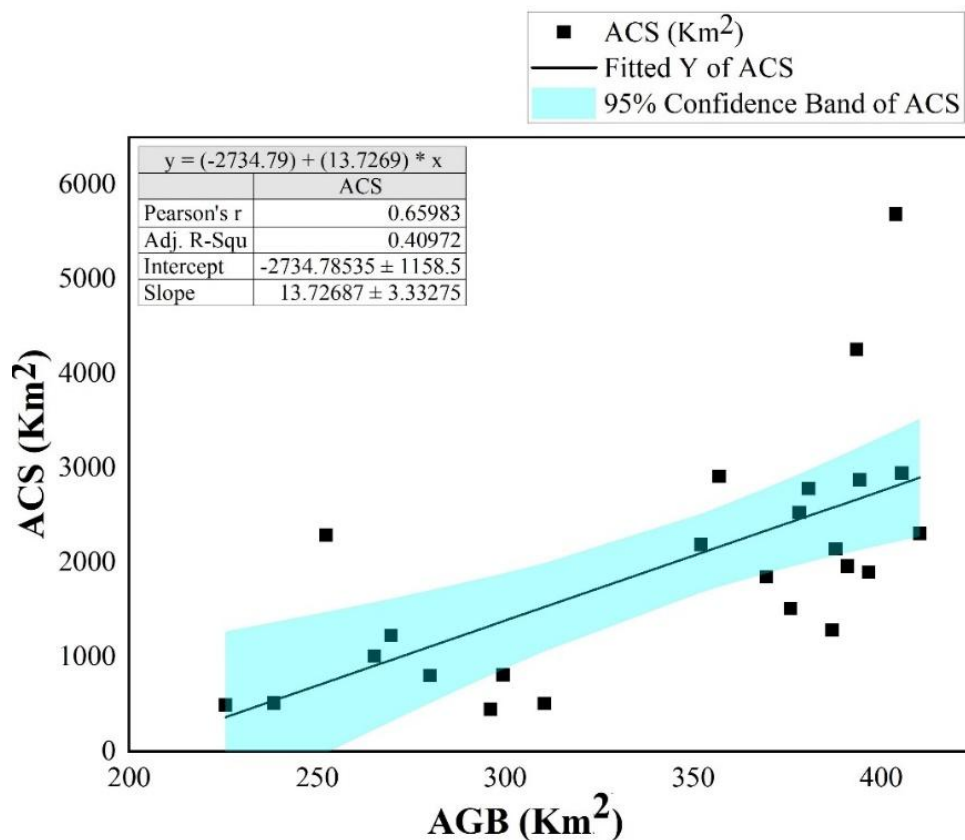


شکل(۴): تجزیه و تحلیل روابط آماری پارامترهای پیش‌بینی با مساحت خلیج گرگان

A)



B)



شکل (۵): نمودار برازش رگرسیونی غیر خطی تراز آب (A) و مساحت آبی دریای خزر (B) با مساحت آبی خلیج گرگان

میانگین)، R^2 (ضریب تبیین)، MBE (خطای میانگین بایاس) و MAPE (خطای درصدی مطلق میانگین) است. این معیارها به طور گسترده برای سنجش دقت و کارایی مدل‌های پیش‌بینی در مطالعات مختلف استفاده می‌شود. برای ارزیابی مدل‌ها، پیش‌بینی‌هایی برای پنج سال به طور تصادفی انجام شد تا مقایسه‌ای میان مقادیر واقعی و پیش‌بینی‌شده صورت گیرد.

بر اساس نتایج ارائه شده در شکل ۶ و جدول ۲، مدل XGBoost عملکرد بسیار دقیق و قابل اطمینانی از خود نشان داد. معیارهای ارزیابی شامل میانگین مطلق خطا برابر با ۱۵.۳، ریشه میانگین مربعات خطا معادل ۱۶.۹، و خطای میانگین درصد مطلق برابر با ۴.۹ درصد بودند که همگی بیانگر دقت بالا در پیش‌بینی مساحت آبی می‌باشد. علاوه بر این، ضریب تبیین (R^2) مدل برابر با ۰.۹۳ محاسبه شد که نشان‌دهنده همبستگی قوی و تطابق مناسب بین مقادیر پیش‌بینی‌شده و واقعی است. از منظر تحلیل خطای سیستماتیک، میانگین خطای بایاس مدل ۷.۶ به دست آمد،

تحلیل و ارزیابی مدل‌های پیش‌بینی

در این تحقیق، برای پیش‌بینی مساحت آبی خلیج گرگان، از چهار مدل یادگیری ماشین شامل XGBoost، شبکه عصبی با توابع پایه شعاعی (RBFN)، جنگل تصادفی (Random Forest) و رگرسیون خطی استفاده شد. برای ارزیابی عملکرد این مدل‌ها، داده‌ها به دو بخش آموزشی (۸۰ درصد) و آزمایشی (۲۰ درصد) تقسیم شد. تمامی مدل‌ها با استفاده از داده‌های مشابه آموزش داده شد. ابتدا به دو مجموعه آموزشی و آزمایشی تقسیم شد. برای این منظور، از تابع `train_test_split` از کتابخانه `sklearn.model_selection` استفاده شد که به طور تصادفی ۸۰ درصد از داده‌ها را به مجموعه آموزشی و ۲۰ درصد را به مجموعه آزمایشی اختصاص داد. به منظور تکرارپذیری نتایج، از مقدار ثابت `random_state=42` برای تضمین اجرای یکسان در هر بار استفاده شد. برای ارزیابی عملکرد مدل‌ها، از پنج معیار رایج استفاده شد که شامل RMSE (خطای مربع ریشه‌ای)، MAE (خطای مطلق

مختلف مانند تنظیم وزن نمونه‌ها یا تحلیل دقیق‌تر توزیع ویژگی‌های ورودی استفاده کرد. در مدل رگرسیون خطی میانگین مطلق خطا برابر با ۲۲.۷ و ریشه میانگین مربعات خطا معادل ۳۲.۸ نشان‌دهنده وجود خطاهای متوسط در پیش‌بینی‌ها است. خطای میانگین درصد مطلق نیز برابر با ۸.۳۰ درصد بود، که تطابق نسبی بین پیش‌بینی‌ها و مقادیر واقعی را تأیید می‌کند. ضریب تبیین (R^2) برابر با ۰.۷۴ نشان می‌دهد که مدل قادر به توضیح ۷۴ درصد از تغییرات متغیر هدف بر اساس ویژگی‌های ورودی است. هرچند این مقدار نشان‌دهنده دقت نسبی مدل است، اما در مقایسه با مدل‌های پیچیده‌تر مانند جنگل تصادفی یا XGBoost، ممکن است نشان‌دهنده محدودیت‌های رگرسیون خطی در مدل‌سازی روابط غیرخطی یا پیچیده در داده‌ها باشد. تحلیل خطای سیستماتیک مدل با استفاده از میانگین خطای بایاس مقدار ۴.۲ را نشان داد که حاکی از تمایل مدل به بیش‌برآوردی در پیش‌بینی‌ها است. این مقدار مثبت، به احتمال زیاد ناشی از عدم توانایی مدل در مدیریت داده‌هایی با تغییرات غیرخطی یا داده‌های پرت است. با این حال، مقدار نسبتاً کوچک MBE در مقیاس داده‌ها نشان می‌دهد که این خطا اثر قابل توجهی بر پیش‌بینی‌های کلی مدل ندارد. در این تحقیق، مدل‌های مختلف پیش‌بینی بر اساس معیارهای ارزیابی استاندارد مورد مقایسه قرار گرفته شد. نتایج نشان داد که مدل XGBoost با ضریب تبیین (R^2) ۰.۹۳ بهترین عملکرد را در بین مدل‌های مورد بررسی داشت. این مدل با کمترین خطای میانگین درصدی و کمترین میزان خطای سیستماتیک توانست بالاترین دقت را در پیش‌بینی‌ها ارائه دهد و به‌عنوان مدل بهینه برای این مطالعه انتخاب شود. در مقابل، مدل‌های جنگل تصادفی و رگرسیون خطی با دقت کمتر و مقادیر ضریب تبیین به‌ترتیب برابر با ۰.۷۶ و ۰.۷۴، نتایج متوسطی داشته و در رتبه‌های دوم و سوم قرار می‌گیرند. مدل توابع شعاعی، با بالاترین مقادیر خطای میانگین مطلق و ریشه میانگین مربعات خطا، همچنین با بیشترین میزان خطای سیستماتیک، ضعیف‌ترین عملکرد را به نمایش گذاشت. در مجموع، مدل XGBoost با ترکیب دقت بالا و کمترین میزان خطای سیستماتیک، به‌عنوان بهترین مدل پیش‌بینی

که به وضوح تمایل مدل به بیش‌برآوردی (Overestimation) را نشان می‌دهد. این مقدار بایاس قابل مدیریت است، اما به‌عنوان یک نقطه بهبود، می‌توان از تکنیک‌های تنظیم بیشتر مدل برای کاهش این خطا استفاده کرد. بر اساس نتایج حاصل از ارزیابی مدل RBFN، معیارهای ارزیابی عملکرد شامل میانگین مطلق خطا برابر با ۳۰.۹، ریشه میانگین مربعات خطا معادل ۳۲.۵ و خطای میانگین درصد مطلق برابر با ۱۰.۱ درصد بود. این مقادیر نشان‌دهنده دقت قابل قبول مدل در پیش‌بینی داده‌های هدف است. علاوه بر این، ضریب تبیین (R^2) مدل برابر با ۰.۷۵ به دست آمد که بیانگر تطابق خوب بین پیش‌بینی‌ها و مقادیر واقعی است. محاسبه میانگین خطای بایاس، که مقدار آن برابر با ۱۲.۰۱- بود، این مقدار منفی نشان‌دهنده تمایل مدل به کم‌برآوردی در پیش‌بینی‌ها است. چنین خطای بایاسی می‌تواند نشان دهد که مدل در تشخیص برخی از ویژگی‌های خاص داده‌ها یا تغییرات ناگهانی در مقادیر هدف دقت کافی را ندارد. با این حال، مقدار مطلق خطای بایاس نسبت به مقیاس داده‌ها نسبتاً کم بوده و نشان می‌دهد که مدل از نظر بایاس عملکرد معقولی دارد.

عملکرد مدل RF بر اساس معیارهای ارزیابی نشان داد که میانگین مطلق خطا برابر با ۲۱.۱ و ریشه میانگین مربعات خطا معادل ۳۱.۶ بود، که این مقادیر بیانگر دقت نسبی مدل در پیش‌بینی‌ها می‌باشد. علاوه بر این، خطای میانگین درصد مطلق معادل ۸.۰۸ درصد است که نشان‌دهنده تطابق نسبی مدل با مقادیر واقعی داده‌ها است. ضریب تبیین (R^2) برابر با ۰.۷۶ نشان‌دهنده مناسب بودن توضیح‌دهندگی مدل در پیش‌بینی تغییرات داده‌های هدف است. با این حال، تحلیل میانگین خطای بایاس که مقدار آن برابر با ۱۳.۳ محاسبه شد، حاکی از تمایل مدل به بیش‌برآوردی در پیش‌بینی‌ها است. این مقدار مثبت می‌تواند نشان‌دهنده این باشد که مدل در برخی موارد، تأثیر ویژگی‌های خاص یا روندهای موجود در داده‌ها را بیش از حد تخمین زده است. در عین حال، مقیاس خطای بایاس نسبت به دامنه داده‌ها نشان می‌دهد که مدل از نظر سیستماتیک بودن خطا عملکردی نسبتاً مناسب دارد، اما برای بهبود دقت و کاهش این خطا، می‌توان از روش‌های

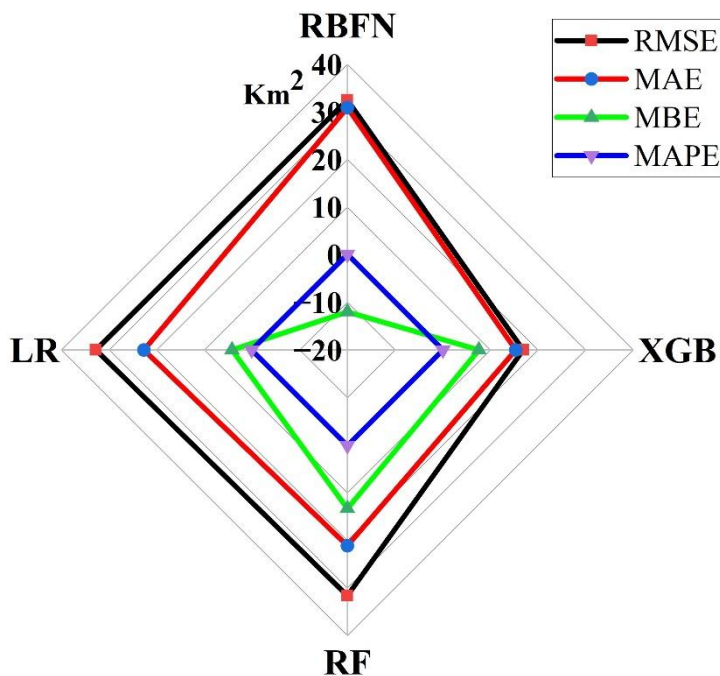
و قابلیت پردازش داده‌های پیچیده، می‌تواند در زمینه‌هایی همچون پیش‌بینی تغییرات سطح آب تالاب‌ها، تحلیل تأثیرات تغییرات اقلیمی بر منابع آبی و برنامه‌ریزی مدیریت منابع طبیعی موثر باشد.

شناخته شد و از نظر علمی، انتخاب مناسب‌تری برای مسائل پیچیده پیش‌بینی محسوب می‌شود. این مدل‌ها به‌ویژه برای مدیریت منابع آب، مدیریت تالاب‌ها، پیش‌بینی تغییرات اقلیمی، مدل‌سازی اثرات خشکسالی و پیش‌بینی کیفیت منابع آبی کاربرد دارند. مدل XGBoost، به دلیل دقت بالا

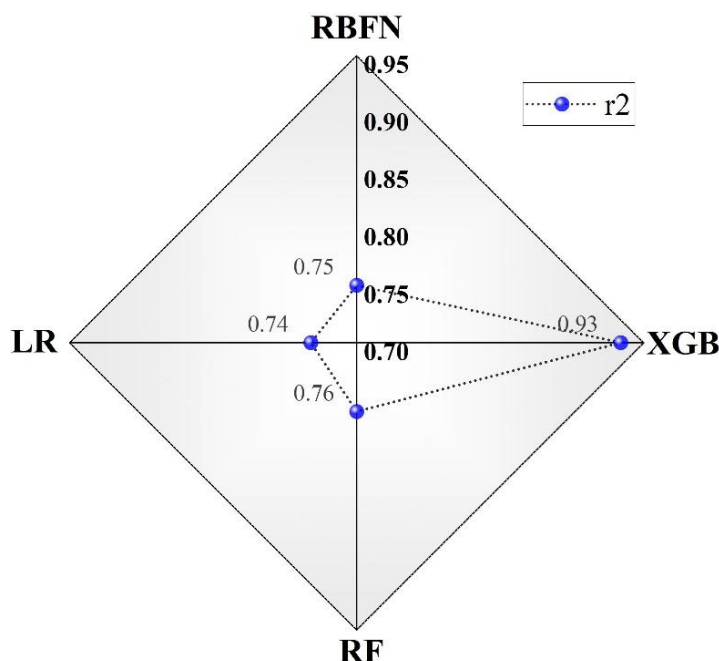
جدول (۲): ارزیابی مدل‌های پیش‌بینی

مدل	RMSE	MAE	MBE	MAPE	R ²
RBFN	۳۲/۵	۳۰/۹	-۱۲/۰۱	٪۱۰/۱۰	۰/۷۵
XGB	۱۶/۹	۱۵/۳۴	۷/۶	٪۴/۹۰	۰/۹۳
RF	۳۱/۶	۲۱/۱	۱۳/۳	٪۸/۰۸	۰/۷۶
LR	۳۲/۸	۲۲/۷	۴/۲	٪۸/۳۰	۰/۷۴

A)



B)



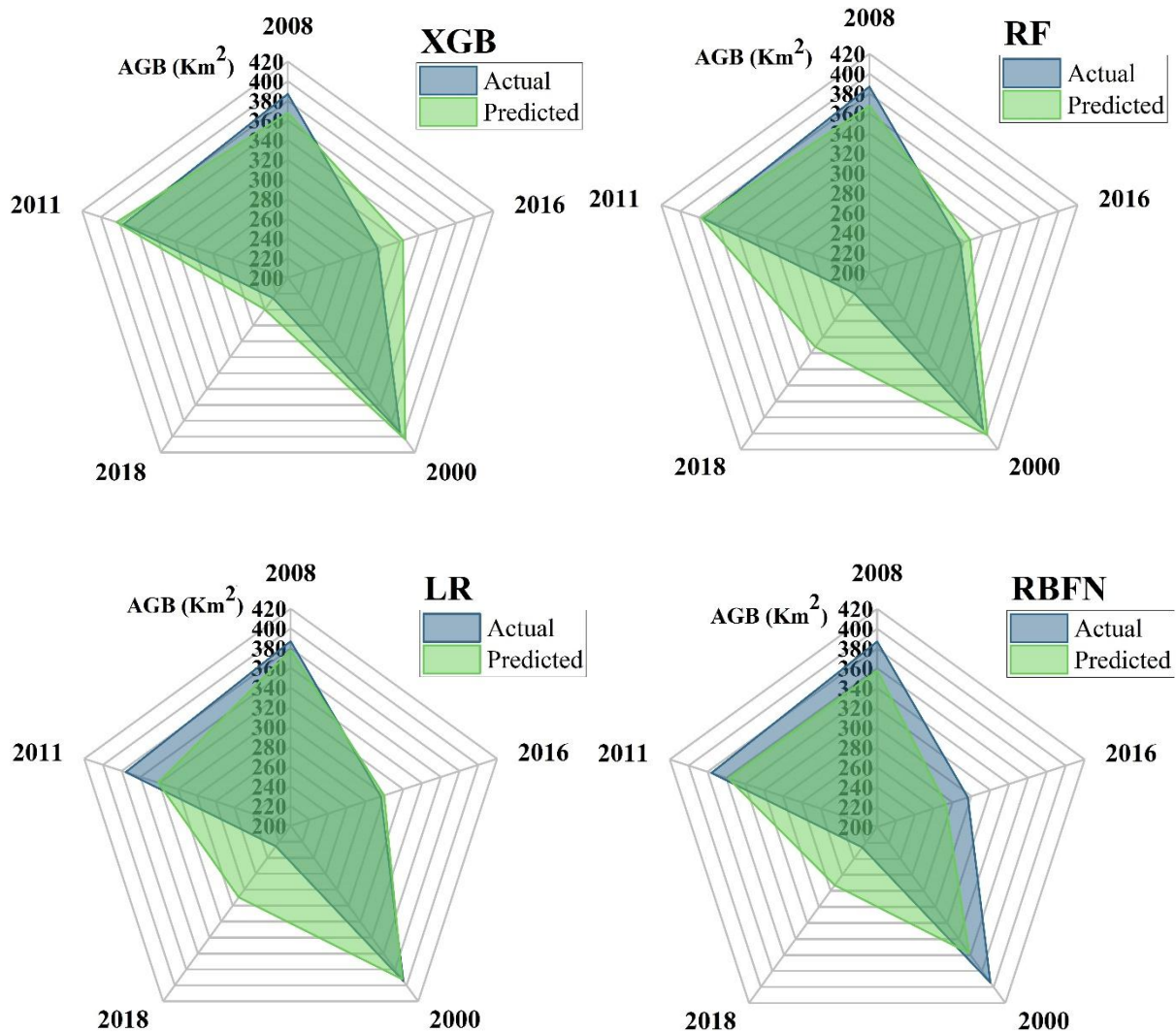
شکل (۶): ارزیابی مدل‌های پیش‌بینی با استفاده از شاخص‌های میانگین مربعات خطا، خطای بایاس، میانگین قدر مطلق خطا و درصد قدر مطلق خطا (A) و ضریب تبیین (B)

رگرسیون خطی و شبکه عصبی RBFN تمایل به بیش‌برآوردی داشتند. در مقابل، مدل XGBoost به‌طور قابل توجهی نزدیک‌ترین پیش‌بینی‌ها را به داده‌های واقعی ارائه داد و توانست نوسانات شدید کاهش و افزایش مساحت آبی خلیج را به‌خوبی شبیه‌سازی کند. این امر نشان‌دهنده عملکرد برتر مدل XGBoost در شبیه‌سازی رفتارهای پیچیده و نوسانات زیاد مساحت آبی خلیج گرگان است. در مجموع، مدل XGBoost به‌عنوان مدل برتر برای پیش‌بینی مساحت آبی خلیج گرگان شناخته شد، زیرا توانسته است نوسانات شدید مساحت آبی را به‌دقت شبیه‌سازی کند و رفتار مناسب‌تری در شرایط مختلف از خود نشان دهد. این ویژگی‌ها نشان می‌دهند که این مدل در پیش‌بینی‌های دقیق و تحلیل‌های پیچیده منابع آبی و اکوسیستم‌های حساس می‌تواند کاربرد قابل توجهی داشته باشد. این مدل همچنین می‌تواند به‌عنوان پایه‌ای برای شبیه‌سازی تغییرات در سایر اکوسیستم‌های حساس نظیر تالاب‌ها و رودخانه‌ها به کار گرفته شود و در برنامه‌ریزی‌های کلان محیط‌زیستی،

بررسی عملکرد پیش‌بینی مدل‌ها

در این تحقیق، از چهار مدل یادگیری ماشین شامل XGBoost، شبکه عصبی با توابع پایه شعاعی (RBFN)، جنگل تصادفی و رگرسیون خطی برای پیش‌بینی مساحت آبی خلیج گرگان در بازه زمانی ۱۳۷۸ تا ۱۴۰۱ استفاده شد. داده‌های این تحقیق به‌صورت سالانه و برای پنج سال به‌طور تصادفی انتخاب شدند، که شامل سال‌های ۱۳۷۸، ۱۳۸۶، ۱۳۸۹، ۱۳۹۴ و ۱۳۹۶ بودند. این سال‌ها به‌دلیل ویژگی‌های خاص خود از نظر مساحت (بالا، متوسط و پایین‌ترین میزان مساحت آبی خلیج) انتخاب شدند. در شکل ۷، مقایسه‌ای بین داده‌های واقعی و مقادیر پیش‌بینی شده توسط مدل‌ها ارائه شده است. نتایج نشان می‌دهد که تمامی مدل‌ها توانسته‌اند روند کلی و نوسانات داده‌ها را شبیه‌سازی کنند. به‌ویژه زمانی که مساحت خلیج گرگان به بیشترین میزان خود در سال‌های مختلف رسید، مدل‌ها به‌خوبی توانستند نوسانات را پیش‌بینی کنند. با این حال، در شرایطی که مساحت آبی خلیج گرگان به شدت کاهش یافته است (مانند سال ۱۳۹۶)، مدل‌های جنگل تصادفی،

مانند مقابله با کاهش سطح آب و مدیریت پایدار منابع آبی،
موثر باشد.



شکل (۷): پیش‌بینی مدل‌های XGBoost، شبکه عصبی با توابع پایه شعاعی، جنگل تصادفی و رگرسیون خطی و داده‌های واقعی به صورت پادساعتگرد است. نرخ انتقال رسوب به دهانه خلیج گرگان ۰.۰۰۱ مترمکعب بر ثانیه بر هر متر و در زبانه میانکاله کمتر از ۰.۰۰۰۰۲ مترمکعب بر ثانیه بود. همچنین، عملکرد دهانه آشوراده در طوفان‌های بزرگ منطقه تحت شرایط کنونی در وضعیت پایدار قرار دارد (Roodbari et al., 2017). در تحقیقی، به بررسی تغییرات سطح آب و مناطق آبی خلیج گرگان با استفاده از شاخص‌های سنجش از دور و مدل‌سازی عددی پرداخته است. داده‌های ماهواره‌ای و شاخص‌هایی چون NDWI، در سال‌های اخیر، مطالعات متعددی بر روی خلیج گرگان انجام شده است که به بررسی جنبه‌های مختلف این اکوسیستم منحصربه‌فرد پرداخته‌اند. به عنوان مثال، در پژوهش به بررسی سازوکار و عملکرد دهانه خلیج گرگان و ساحل میانکاله با مدل‌سازی عددی در تراز فعلی (۲۶.۷- متر) پرداخته است. نتایج نشان داد که انرژی امواج و سرعت جریان‌های ساحلی از غرب به شرق کاهش می‌یابد و دهانه آشوراده تحت تأثیر انرژی کمتری از امواج قرار دارد. جهت جریان‌ها در بیرون خلیج از غرب به شرق و درون خلیج



سطح آب و اثرات آن بر حیات وحش و فعالیت‌های انسانی در مناطق ساحلی و تالاب‌ها مفید باشد. این تحقیق با ارائه یک رویکرد نوین، راهکارهای موثری برای پایش و مدیریت منابع آبی ارائه می‌دهد که نه تنها در مدیریت بحران‌ها، بلکه در برنامه‌ریزی‌های توسعه پایدار نیز قابل استفاده است.

نتیجه‌گیری

با توجه به بحران‌های ناشی از خشکی و کاهش مساحت آبی خلیج گرگان، که ناشی از تغییرات اقلیمی و فعالیت‌های انسانی است، پیش‌بینی دقیق روندهای این تغییرات برای مدیریت منابع آبی و حفاظت از اکوسیستم‌های آبی این منطقه ضروری به نظر می‌رسد. در این پژوهش به منظور پیش‌بینی تغییرات مساحت آبی خلیج گرگان (MNDWI) با استفاده از داده‌های، بارش، دما، تراز آبی و مساحت آبی دریای خزر و دبی ورودی به خلیج گرگان در بازه زمانی ۱۳۷۸ تا ۱۴۰۱ پرداخته شد. برای دریافت داده‌های تراز آب و مساحت آبی از ماهواره‌های مودیس، سنتینل ۲ و جیسون ۳، بکار گرفته شد. به منظور مدل‌سازی این تغییرات، از روش‌های پیشرفته یادگیری ماشین شامل مدل‌های گرادینت تقویت کننده (XGBoost)، شبکه عصبی با توابع پایه شعاعی (RBFN)، جنگل تصادفی و رگرسیون خطی استفاده شد. ارزیابی روابط میان متغیرها با استفاده از ضریب همبستگی پیرسون، آزمون فرضیات آماری و تحلیل رگرسیونی غیرخطی انجام شد و نمودار نیز رسم شد. همچنین، شاخص‌های ارزیابی مدل شامل RMSE، MAPE، MAE، MBE و R^2 بوده است. برای بررسی بهتر مدل، داده‌های واقعی و داده‌های پیش‌بینی در پنج سال مختلف مورد مقایسه قرار گرفت. نتایج نشان داد که، تغییرات تراز آبی و مساحت دریای خزر و دبی ورودی به خلیج بیشترین همبستگی را با مساحت آبی خلیج گرگان داشته که به ترتیب دارای ضریب پیرسون ۰/۹۰، ۰/۶۶، ۰/۶۳ و ضریب تبیین ۰/۸۳، ۰/۴۳، ۰/۴۰ بوده است و در سطح ۰/۰۱ معنی دار می‌باشند. همچنین بارش و دما کمترین تأثیر را داشتند، به طوری که به ترتیب ضریب پیرسون ۰/۴۳، ۰/۳۴- و ضریب تبیین ۰/۱۹، ۰/۱۱ بوده است و در سطح ۰/۰۵ معنی دار بوده است. ارزیابی مدل-

AWEI و WRI برای تحلیل تغییرات آبی استفاده شده‌اند. برای طبقه‌بندی تصاویر به دو کلاس آب و خشکی، از روش‌های پردازش تصویر و ماشین‌های بردار پشتیبان بهره گرفته شد. نتایج نشان می‌دهد که استفاده از شاخص AWEI و کرنل تابع شعاعی دقت طبقه‌بندی را به ۹۷.۵۱ درصد افزایش داده و این روش پتانسیل بالایی برای پایش مناطق آبی وسیع دارد (Gholizadeh et al., 2024). در پژوهشی به مدل‌سازی هیدرودینامیکی خلیج گرگان پرداخته است. نتایج نشان می‌دهند که سرعت جریان متوسط داخل خلیج ۰.۱ متر بر ثانیه است و دما بین ۰.۵۳- و ۳۶.۵۳+ درجه سانتیگراد تغییر می‌کند. جریان عمومی آب معمولاً خلاف عقربه‌های ساعت است و تغییرات دما و ارتفاع آب الگوی سینوسی فصلی دارند. تأثیر رودخانه‌ها به دلیل تخلیه کم نادیده گرفته شده و فشار هوا تأثیر زیادی بر سطح آب دارد. در دهه گذشته، سرعت جریان کاهش و دما تقریباً ۷ درجه سانتی‌گراد افزایش یافته است (Siadatmousavi & Eftekhari, 2024). مطالعات یادشده نشان می‌دهند که اکثر پژوهش‌ها به موضوعاتی نظیر کیفیت آب، افزایش شوری، پایش تغییرات مناطق آبی، افزایش خشکی کانال‌ها و تحلیل عوامل موثر بر خشک شدن خلیج گرگان پرداخته‌اند. با این حال، تاکنون پژوهش جامعی با استفاده از روش‌های یادگیری ماشین برای پیش‌بینی مساحت آبی خلیج گرگان بر اساس داده‌های هیدرولوژیکی و اقلیمی انجام نشده است. در این پژوهش، از روش‌های پیشرفته یادگیری ماشین شامل XGBoost، شبکه‌های عصبی با توابع پایه شعاعی (RBFN)، جنگل تصادفی و رگرسیون خطی استفاده شد. نتایج تحقیق نشان داد که مدل‌های یادگیری ماشین، به‌ویژه XGBoost، ابزار قدرتمندی برای پیش‌بینی تغییرات مساحت منابع آبی هستند و می‌توانند برای مدیریت بهینه این منابع در برابر تغییرات اقلیمی و فعالیت‌های انسانی به کار روند. پیش‌بینی دقیق تغییرات سطح آب، نقشی حیاتی در برنامه‌ریزی برای مقابله با خشکسالی‌ها و سیلاب‌ها ایفا می‌کند و می‌تواند به شبیه‌سازی اثرات تغییرات اقلیمی بر اکوسیستم‌های آبی، نظیر تالاب‌ها و خلیج‌ها، کمک کند. همچنین، این مدل‌ها می‌توانند برای ارزیابی خطرات زیست‌محیطی نظیر کاهش

خوبی از روند داده‌ها داشته است. با این حال، در شرایطی که مساحت آبی خلیج گرگان به شدت کاهش یافته است (مانند سال ۲۰۱۸)، مدل‌های جنگل تصادفی، رگرسیون خطی و شبکه عصبی RBFN در چار خطای بیش برآوردی شده‌اند. اما مدل XGBoost به‌طور قابل توجهی نزدیک‌ترین روند را به داده‌های واقعی داشته است و توانست نوسانات شدید کاهش و افزایش مساحت آبی خلیج را به‌خوبی تشخیص دهد. نتایج نشان داد که مدل XGBoost نسبت به سایر مدل‌ها دقت بالاتری داشته و پیش‌بینی‌های دقیق‌تری صورت می‌دهد، لذا به‌منظور مدیریت بهتر منابع آبی و کاهش اثرات بحران خشکی، توصیه می‌شود که از مدل‌های پیشرفته یادگیری ماشین برای پیش‌بینی دقیق‌تر تغییرات در این مناطق استفاده شده و تدابیری برای بهبود شرایط اکولوژیکی و آبی خلیج گرگان اتخاذ گردد.

تقدیر و تشکر

این مقاله مستخرج از طرح پژوهشی با شماره ۱۴۰۳-۰۶-۰۲ دانشگاه علوم کشاورزی و منابع طبیعی ساری می‌باشد و مراتب تشکر و قدردانی از حمایت مادی و معنوی دانشگاه علوم کشاورزی و منابع طبیعی ساری به جا آورده می‌شود.

های پیش‌بینی مساحت آبی خلیج گرگان نشان داد که، مدل‌های XGBoost، RBFN، جنگل تصادفی و رگرسیون خطی به ترتیب دارای ضریب تبیین، ۰/۹۳، ۰/۷۵، ۰/۷۶ و ۰/۷۴ می‌باشد، همچنین شاخص خطا مدل‌ها نشان داد که RMSE، MAPE، MAE، MBE مدل XGBoost به ترتیب برابر، ۱۶/۹، ۱۵/۳۴، ۴/۹٪ و ۷/۶٪ بهترین عملکرد را نسبت به سایر مدل‌ها داشته است. مدل‌های جنگل تصادفی و رگرسیون خطی به ترتیب با خطای میانگین درصدی مطلق ۰/۸/۰۸، ۰/۸/۳٪ و میانگین مربعات خطای، ۳۱/۶ و ۳۲/۸ در رتبه‌های دوم و سوم قرار گرفته است. مدل شبکه عصبی توابع شعاعی (RBFN) با بیشترین خطای سیستماتیک ۱۲/۹- و خطای میانگین درصدی مطلق ۱۰/۱٪، ریشه میانگین مربعات خطا ۳۲.۵ و میانگین مطلق خطا ۳۰.۹ بیشترین خطا و ضعیف‌ترین عملکرد را داشته است. برای مقایسه عملکرد پیش‌بینی مدل‌ها پنج سال مختلف شامل سال‌های ۱۳۷۸، ۱۳۸۶، ۱۳۸۹، ۱۳۹۴ و ۱۳۹۶ به‌طور تصادفی انتخاب شد. این سال‌ها به دلیل ویژگی‌های خاص خود از نظر مساحت آبی (بالا، متوسط و پایین‌ترین میزان مساحت آبی خلیج) توسط مدل انتخاب گردید. هر چهار مدل مورد استفاده در این پژوهش نوسانات و روند داده‌ها را به‌طور مناسب پیش‌بینی کرده و درک

منابع

- Al-Selwi, S.M., M.F. Hassan, S.J. Abdulkadir, A. Muneer, E.H. Sumiea, A. Alqushaibi and M.G. Ragab. 2024. RNN-LSTM: From applications to modeling techniques and beyond—Systematic review. *Journal of King Saud University - Computer and Information Sciences*, 36(5): 102068. <https://doi.org/10.1016/j.jksuci.2024.102068>
- Aslam, R.W., H. Shu, I. Naz, A. Quddoos, A. Yaseen, K. Gulshad and S.S. Alarifi. 2024. Machine learning-based wetland vulnerability assessment in the Sindh Province Ramsar site using remote sensing data. *Remote Sensing*, 16(5): 928. <https://doi.org/10.3390/rs16050928>
- Cui, Q., M.E. Ammar, M. Irvani, J. Kariyeva and M. Faramarzi. 2021. Regional wetland water storage changes: The influence of future climate on geographically isolated wetlands. *Ecological Indicators*, 120: 106941. <https://doi.org/10.1016/j.ecolind.2020.106941>
- Esri. 2018. Sentinel-2, 10m Multispectral, Multitemporal, 13-band images with visual renderings and indices. ArcGIS. <https://www.arcgis.com/home/item.html?id=fd61b9e0c69c4e14bebd50a9a968348c>
- Ganaie, M.A., M. Hu, A.K. Malik, M. Tanveer and P.N. Suganthan. 2022. Ensemble deep learning: A review. *Engineering Applications of Artificial Intelligence*, 115: 105151. <https://doi.org/10.1016/j.engappai.2022.105151>
- Gautama, V., G. Gaurava, P. Murugana and M. Annadurai. 2015. Assessment of surface water dynamics in Bangalore using WRI, NDWI, MNDWI, supervised classification and K-T transformation. *International Conference on Water Resources, Coastal and Ocean Engineering*, 4: 739–746. DOI: 10.1016/j.aqpro.2015.02.095

Gholizadeh, S., R. Dezvareh and A. Kiani. 2024. Long-term monitoring of changes in the waterbody of Gorgan Bay using the combined method of remote sensing indicators and numerical modeling. *Amphibious Science and Technology*, 5(1): 59–75. <https://doi.org/10.22034/JAMST.2024.544346.1133>

Jayathilake, T., M.B. Gunathilake, E.M. Wimalasiri and U. Rathnayake. 2023. Wetland water-level prediction in the context of machine-learning techniques: Where do we stand? *Environments*, 10(5): 75. <https://doi.org/10.3390/environments10050075>

Kazempour, Z., M. Danesh-Yazdi, K. Asadifakhr and M. Raie. 2023. Spatiotemporal dynamics of chlorophyll-a in the Gorgan Bay and Miankaleh Peninsula biosphere reserve: Call for action. *Remote Sensing Applications: Society and Environment*, 30: 100946. <https://doi.org/10.1016/j.rsase.2023.100946>

Li, L., M. Herrera, A. Mukherjee, G. Zheng, C. Chen, M. Dhada, H. Brice, A. Parekh and A.K. Parlikad. 2025. Predictive alarm models for improving radio access network robustness. *Expert Systems with Applications*, 259: 125312. <https://doi.org/10.1016/j.eswa.2024.125312>

Li, S., F. Chen, K. Song, G. Liu, H. Tao, S. Xu, X. Wang, Q. Wang and G. Mu. 2022. Mapping the trophic state index of eastern lakes in China using an empirical model and Sentinel-2 imagery data. *Journal of Hydrology*, 608: 127613. <https://doi.org/10.1016/j.jhydrol.2022.127613>

Lorente, A., T. Borsdorff, A. Butz, O. Hasekamp, J. aan de Brugh, A. Schneider, L. Wu, F. Hase, R. Kivi, D. Wunch, D.F. Pollard, K. Shiomi, N.M. Deutscher, V.A. Velazco, C.M. Roehl, P.O. Wennberg, T. Warneke and J. Landgraf. 2021. Methane retrieved from TROPOMI: Improvement of the data product and validation of the first 2 years of measurements. *Atmospheric Measurement Techniques*, 14: 665–684. <https://doi.org/10.5194/amt-14-665-2021>

Luymes, N. and P. Chow-Fraser. 2021. Detection of potential vernal pools on the Canadian Shield (Ontario) using object-based image analysis in combination with machine learning. *Canadian Journal of Remote Sensing*, 47(4): 519-534. <https://doi.org/10.1080/07038992.2021.1900717>

Nadi, M. and A. Yousefi Kebriya. 2024. A Method for Correction of Tropical Rainfall Measuring Mission Satellite Temperature Network in Mazandaran Province. *Iranica Journal of Energy and Environment (IJEE)*, 15(1): 100-110. <https://doi.org/10.5829/ijee.2024.15.01.10>

Nadi, M. and A. Yousefi Kebriya. 2024. Evaluation and modification of TRMM temperature and precipitation products in Mazandaran. *Journal of Climate Research*, 1402(55): 147-160. DOI:10.21203/rs.3.rs-1521038/v1

NOAA (National Oceanic and Atmospheric Administration). 2020. NOAA ESRL global monitoring laboratory: SURFRAD aerosol optical depth. <https://gml.noaa.gov/grad/surfrad/aod/>, accessed in 2020 .

Ouma, Y.O., D.B. Moalafhi, G. Anderson, B. Nkwae, P. Odirile, B.P. Parida and J. Qi. 2022. Dam water level prediction using vector autoregression, random forest regression, and MLP-ANN models based on land-use and climate factors. *Sustainability*, 14(22): 14934. <https://doi.org/10.3390/su142214934>

Pirali Zafraei, A.R., S.A. Hedayati, S. Pourmanafi, O. Bireghdar Keshkuli and R. Ghorbani. 2020. Investigation of the efficiency of support vector machine in predicting changes in water quality parameters (Case study: Choghakhor International Wetland). *Journal of Aquatic Ecology*, 10(1): 23-34. <https://doi.org/10.1001.1.23222751.1399.10.1.3.9>

Rane, N.L., M. Paramesha, S.P. Choudhary and J. Rane. 2024. Artificial intelligence, machine learning, and deep learning for advanced business strategies: A review. *Partners Universal International Innovation Journal (PUIIJ)*, 02(03): 147-171. <https://doi.org/10.5281/zenodo.12208298>

Ranjbar, M.H., M. Jandaghi Alaei and M. Nazarali. 2019. A modeling study of the impact of increasing water exchange rate on water quality of a semi-enclosed bay. *Ecological Engineering*, 136: 177-184. <https://doi.org/10.1016/j.ecoleng.2019.05.024>

Rezaei, R., A. Jafarian, F. Mattern, U.K. Shukla, V. Senapathi and M. Bernecker. 2019. The sedimentology and development of a modern sandspit (Miankaleh Peninsula) and a lacustrine lagoon



(Gorgan Bay), Caspian Sea, Iran. *Marine Geology*, 415: 105974. <https://doi.org/10.1016/j.margeo.2019.105974>

Riley, J.W. and C.C. Stillwell. 2023. Predicting inundation dynamics and hydroperiods of small, isolated wetlands using a machine learning approach. *Wetlands*, 43: 63. <https://doi.org/10.1007/s13157-023-01706-2>

Roodbari Shahmiri, S., M. Adjami and H. Khoshrahan. 2017. Prediction of Gorgan Bay Inlet Performance and Morphological Stability. *Uqyānūs/Shināsī*, 8(31): 53–65. <https://doi.org/10.29252/joc.8.31.53>

Salamat, A., M. Ardestani and B. Malekmohammadi. 2023. Modeling groundwater potential using machine learning models. *Watershed Management Research*, 36(4): 114-132. <https://doi.org/10.22092/wmrj.2023.360675.1506>

Sharbaty, S. 2020. Investigating the rivers' role in the Gorgan Bay possible future after separation from the Caspian Sea. *Environmental Researches*, 12(23): 117-125. <https://doi.org/20.1001.1.20089597.1400.12.23.10.0>

Siadatmousavi, S.M. and A. Eftekhari. 2024. 3D Hydrodynamic Numerical Modeling of Gorgan Bay. *International Journal of Coastal, Offshore and Environmental Engineering (IJCOE)*. doi: 10.22034/ijcoe.2024.357431.1008

Stein, E.D., E.M. Gee, J.B. Adams, K. Irving and L. Van Niekerk. 2021. Advancing the science of environmental flow management for protection of temporarily closed estuaries and coastal lagoons. *Water*, 13(5): 595. <https://doi.org/10.3390/w13050595>

Torahi, A., M. FiroziNejad and A. Abdolkhani. 2017. Assessment of SVM and MLC algorithms on land use/land cover mapping of riparian forest using OLI sensor (Case study: Riparian Forest of Maroon, Behbahan). *Iranian Journal of Remote Sensing and GIS*, 9(1): 49-62. https://gisj.sbu.ac.ir/article_96239.html?lang=fa

Yousefi Kebria, A., M. Nadi and M. Jamei. 2021. Investigation of Statistical and Geostatistical Methods in Preparing the Rainfall Map of Mazandaran Province. *Journal of Watershed Management Research*, 12(23): 212-223. doi: 10.52547/jwmr.12.23.212

Yousefi Kebriya, A., M. Nadi and M. Jamei. 2022. Combining interpolation methods and precipitation products of TRMM satellite to increase the accuracy of rainfall maps in Mazandaran province. *Water and Soil Conservation*, 28(3): 49-70. doi: 10.22069/JWSC.2022.19286.3477

Yousefi Kebriya, A., M. Nadi, E. Ghanbariparmehr and S. Zhongchang. 2024. Modeling and Forecasting Dust Pollution Using Satellite Data and Machine Learning Techniques: A Case Study of Ahvaz. *Journal of Agricultural Meteorology*. doi: 10.22125/agmj.2024.478677.1174

Zakaria, M.N.A., A.N. Ahmed, M.A. Malek, A.H. Birima, M.M.H. Khan, M. Sherif and A. Elshafie. 2023. Exploring machine learning algorithms for accurate water level forecasting in Muda River, Malaysia. *Heliyon*, 9(7): e17689. <https://doi.org/10.1016/j.heliyon.2023.e17689>

Zheng, L., H. Wang, C. Liu, S. Zhang, A. Ding, E. Xie, J. Li and S. Wang. 2021. Prediction of harmful algal blooms in large water bodies using the combined EFDC and LSTM models. *Journal of Environmental Management*, 295: 113060. <https://doi.org/10.1016/j.jenvman.2021.113060>