



University of Tehran Press

Journal of Environmental Studies
Vol. 50, No. 1, Spring 2024

Journal Homepage: www.Jes.ut.ac.ir
Print ISSN: 1025-8620 Online ISSN 2345-6922

Estimation of Water Turbidity by Remote Sensing and Random Forest Algorithm, Case Study: Chitgar Persian Gulf Martyrs Lake, Tehran

Behnaz Karimi ¹, Seyed Hossein Hashemi ², Hossein Aghighi ³

1. Research Institute of Environmental Sciences, Shahid Beheshti University, Tehran, Iran, Email: b_karimi@sbu.ac.ir
2. Research Institute of Environmental Sciences, Shahid Beheshti University, Tehran, Iran, Email: h_hashemi@sbu.ac.ir
3. Faculty of Remote Sensing, Shahid Beheshti University, Tehran, Iran, Email: h_aghighi@sbu.ac.ir

Article Info

ABSTRACT

**Research Article:
Research Paper**

Article history:

Received 3 November 2022
Received in revised form
29 November 2023
Accepted 11 December 2023
Publish online 21 May 2024

Keywords:

Random Forest, Remote Sensing, Water Turbidity, Water Quality Monitoring

Water turbidity is one of the most important parameters of water quality, which represents the transparency of water and is effective in eutrophication. This research was done to estimate the amount of water turbidity using remote sensing data and the random forest technique. For this purpose, the water quality monitoring data of Chitgar Lake in Tehran were used, which is an artificial shallow lake with recreational and urban scenery usage. The Landsat 8 OLI/TIRS and Sentinel 2 MSI satellite images were extracted after matching the date of field observation data and satellite images from 2016 to 2021. Data were divided into calibration and validation datasets. After performing pre-processing processes on satellite images, important bands were recognized using the random forest method. Afterward, appropriate band composition and algorithms were selected and regression models were fitted and validated. The optimum model was able to estimate water turbidity with $Adj.R^2=0.6$, $RMSE=1.07$ NTU, and $NRMSE=12\%$ for Landsat-8 as well as with $Adj.R^2=0.73$, $RMSE=1.23$ NTU and $NRMSE=9\%$ for Sentinel-2 satellite and estimated with a power of 80% for Chitgar Lake. Consequently, the optimal predictive model in Sentinel-2 was chosen with the assistance of the random forest. Moreover, the predictive model was able to estimate the water turbidity in Chitgar Lake with acceptable accuracy.

Cite this article: Karimi, B., Hashemi, S. H., Aghighi, H. (2024). Estimation of Water Turbidity by Remote Sensing and Random Forest Algorithm, Case Study: Chitgar Persian Gulf Martyrs Lake, Tehran. *Journal of Environmental Studies*, 50 (1), 1- 15.

DOI: <http://doi.org/10.22059/JES.2023.350214.1008365>

© The Author(s).

Publisher: University of Tehran Press.



DOI: <http://doi.org/10.22059/JES.2023.350214.1008365>

Extended abstract**Introduction**

Water turbidity is one of the optical characteristics of water and is influenced by the concentration of organic and inorganic substances suspended and soluble in water, such as chlorophyll-a, suspended solids, and dissolved organic matter. Increasing water turbidity affects the light access for aquatic life and the health of aquatic ecosystems. On the other hand, suspended particles that cause water turbidity contain nutrients and are one of the causes of eutrophication in water sources. Therefore, monitoring the turbidity and the pattern of its changes is necessary for water quality management.

In recent years, the prediction of water turbidity by satellite images has been successfully carried out and validated in surface waters. Correlation between field measurement data, satellite bands, and construction regression models are methods to find retrieval turbidity algorithms. This study was conducted to improve the water quality monitoring process of Chitgar Lake as a shallow urban recreational lake. Effective water quality management of these water bodies often requires high-frequency monitoring and trend analysis. This research tries to provide a method to obtain the optimal water turbidity forecasting model with higher accuracy, by combining remote sensing and machine learning methods to overcome problems caused by geographical conditions and the inherent characteristics of the desired area on the prediction models.

Materials and methods

Chitgar Lake in the northwest of Tehran, with an average depth of 5 meters is the study area in this research.

Satellite data

The images of Landsat-8 and Sentinel-2 satellites were used with a maximum of one day difference from the field data sampling and less than 10% cloud cover in the studied area.

Field measurement

The field data of water turbidity was extracted from the lake quality monitoring data from 2016 to 2021 on 5 monitoring stations. After matching the date of field observation and satellite images, 105 and 104 data sets were acquired for Landsat-8 and Sentinel-2 satellites, respectively. After assessment of the statistical distribution, Data were divided into calibration and validation data sets.

Random forest

Random forest is a machine learning method based on classification and regression through several decision trees as classifiers. One of the benefits of the random forest algorithm is that determines the importance of a feature by considering the number of tree nodes that use that feature. This preponderance of the random forest algorithm has been used in this research for identifying important variables.

Accuracy

The accuracy evaluation was performed using RMSE, NRMSE, and Adj.R² was used to express the accuracy of regression models. In this research, the Power of model has also been calculated.

Discussion of results

Evaluating the relative importance coefficients of the bands by random forest in the Landsat-8 satellite illustrated the higher importance coefficients for b2 and b3 bands as important bands. Then single-band and band composition algorithms of these bands were evaluated in linear and non-linear regression models and the band ratio b2/b3 in the exponential equation estimated the best prediction.

In the Sentinel-2 satellite, bands b2, b3, and b4 were acquired as effective bands, and the best result in this satellite was achieved from the combination of three bands in an exponential equation. The models were verified with validation data sets for both satellites. The statistical results and spatial distribution maps show that the Sentinel-2 satellite has a better performance in estimating the turbidity values in the lake. On the other hand, according to the result, the Landsat-8 satellite can be used for achieving high temporal coverage to predict and trend analysis of water turbidity, with acceptable accuracy. The results showed agreement in the effective bands in water turbidity estimation algorithms in this study and similar research. It should be noted that the depth of the lake is one of the most effective factors in causing errors in water turbidity estimation due to the effects of reflection from the bottom.

Conclusions

In this research, the efficiency of Landsat-8 and Sentinel-2 satellites in estimating water turbidity value by remote sensing data and using the random forest method was investigated in a shallow urban artificial lake. The results showed that the Sentinel-2 satellite has appropriate performance in estimating the water turbidity compared to the Landsat-8 satellite in the lake. However, due to the relatively acceptable accuracy of the prediction model for the Landsat-8 satellite, this satellite can also be used for higher temporal coverage in monitoring the lake. The use of this method can provide the required information with higher temporal and spatial coverage, to improve the monitoring process simultaneously reducing costs in Chitgar Lake.



بر آورد کدورت آب با استفاده از سنجش از دور و الگوریتم جنگل تصادفی، مطالعه موردی: دریاچه شهدای خلیج فارس چیتگر تهران

بهناز کریمی^۱، سیدحسین هاشمی^۲✉، حسین عقیقی^۳

۱. پژوهشکده علوم محیطی دانشگاه شهید بهشتی، تهران، ایران، رایانامه: b_karimi@sbu.ac.ir

۲. پژوهشکده علوم محیطی دانشگاه شهید بهشتی، تهران، ایران، رایانامه: h_hashemi@sbu.ac.ir

۳. دانشکده سنجش از دور دانشگاه شهید بهشتی، تهران، ایران، رایانامه: h_aghghi@sbu.ac.ir

اطلاعات مقاله	چکیده
نوع مقاله: مقاله پژوهشی	کدورت آب از مهم ترین پارامترهای کیفیت آب محسوب می شود که معرف شفافیت آب و موثر بر تغذیه گرایبی است. این پژوهش با هدف برآورد مقدار کدورت آب با استفاده از داده های سنجش از دور و تکنیک جنگل تصادفی انجام شده است. بدین منظور، از داده های پایش کیفیت آب دریاچه شهدای خلیج فارس چیتگر تهران که دریاچه ای شهری و کم عمق، با کاربری تفرج و منظر شهری است، استفاده شد. تصاویر ماهواره های لندست-۸ و سنتینل-۲ پس از انطباق تاریخ داده های میدانی و تصاویر ماهواره ای برای دوره زمانی سال ۱۳۹۵ تا ۱۴۰۰، انتخاب و داده ها به دو گروه جهت تولید و اعتبارسنجی مدل تقسیم شدند. نخست عملیات پیش پردازش روی تصاویر ماهواره ای انجام شد. سپس با استفاده از تکنیک جنگل تصادفی باندهای موثر شناسایی گردیدند، پس از آن، ترکیب های باندهای بهینه انتخاب و مدل های رگرسیون برازش و اعتبارسنجی شدند. مدل به دست آمده، میزان کدورت آب را با $Adj.R^2=0.6$ ، $RMSE=1.07$ NTU و $NRMSE=12\%$ در ماهواره لندست-۸ و $Adj.R^2=0.73$ ، $RMSE=1.23$ NTU و $NRMSE=9\%$ در ماهواره سنتینل-۲ و با توان آماری ۸۰ درصد برای دریاچه چیتگر پیش بینی کرد. بدین ترتیب، مدل برآوردی بهینه با کمک تکنیک جنگل تصادفی براساس داده های ماهواره سنتینل-۲ به دست آمد و مدل پیش بینی توانست مقادیر کدورت آب را در دریاچه چیتگر با دقت قابل قبولی برآورد کند.
تاریخ دریافت: ۱۴۰۱/۸/۱۲	
تاریخ بازنگری: ۱۴۰۲/۰۹/۰۸	
تاریخ پذیرش: ۱۴۰۲/۰۹/۲۰	
تاریخ انتشار: ۱۴۰۳/۰۳/۰۱	
کلیدواژه ها: پایش کیفیت آب، جنگل تصادفی، سنجش از دور، کدورت آب	

استناد: کریمی، بهناز؛ هاشمی، سیدحسین؛ عقیقی، حسین. (۱۴۰۳). برآورد کدورت آب با استفاده از سنجش از دور و الگوریتم جنگل تصادفی، مطالعه موردی: دریاچه شهدای خلیج فارس چیتگر تهران. نشریه محیط شناسی، ۵۰(۱)، ۱-۱۵.

DOI: <http://doi.org/10.22059/JES.2023.350214.1008365>

DOR: 20.1001.1.10258620.1403.50.1.2.9

© نویسندگان.

ناشر: انتشارات دانشگاه تهران.



DOI: <http://doi.org/10.22059/JES.2023.350214.1008365>

۱. مقدمه

کدورت آب یکی از ویژگی‌های نوری آب و معیاری برای شفافیت آب است که از میزان نور پراکنده شده توسط مواد در یک پهنه آبی به دست می‌آید و تحت تاثیر غلظت مواد آلی و معدنی معلق و محلول در آب، از جمله کلروفیل‌آ، جامدات معلق و مواد آلی محلول رنگی می‌باشد (Wass et al., 1997). کدورت آب از طریق کاهش منابع غذایی، تخریب مکان تخم‌ریزی و تاثیر بر عملکرد آبشش به ماهیان و سایر موجودات آبی آسیب می‌رساند (Quang et al., 2017). کدورت و ذرات معلق به دلیل ارتباط مستقیم‌شان با میزان عبور نور فتوسنتزی موثر بر رشد پلانکتون‌ها و سایر جلبک‌ها، دو متغیر مهم در مطالعه پهنه‌های آبی هستند (Hafeez et al., 2019). در میان پارامترهای کیفیت آب، رسوبات معلق رایج‌ترین مشکل در آب‌های داخلی هستند که با افزایش کدورت آب، میزان نور در دسترس برای آبزیان را کاهش می‌دهند (Ritchie et al., 2003). همچنین ذرات معلق حاوی مواد مغذی هستند و از دلایل ایجاد اوتروفیکاسیون در منابع آبی می‌باشند. از سوی دیگر، کدورت آب، از نظر زیبایی‌شناختی بر مطلوبیت منابع آب اثر منفی دارد (Moore, 1980). بنابراین، پایش و ثبت اطلاعات کدورت و غلظت رسوبات معلق و نیز الگوی تغییرات آنها در بدنه‌های آبی برای درک بهتر کیفیت منابع آبی ضروری است (Hafeez et al., 2019). اندازه‌گیری کدورت آب به طور سنتی، به صورت بصری با استفاده از یک دیسک Secchi (Baughman et al., 2015) یا به طور مستقیم از طریق روش‌های وزن‌سنجی یا تجزیه آزمایشگاهی یا به صورت نوری و مستقیماً با استفاده از کدورت‌سنج‌های نوری در محل انجام می‌شود (Pavelsky & Smith, 2009).

در ۵۰ سال گذشته پیش‌بینی کدورت آب و سایر پارامترهای شفافیت آب، با بهره‌گیری از ماهواره‌های فضایی در پهنه آب‌های سطحی مانند: دریاچه‌ها، مصب‌ها، مخازن و سواحل با موفقیت انجام و با استفاده از اندازه‌گیری‌های میدانی اعتبارسنجی شده است (Harrington et al., 1992). کدورت آب به دلیل ویژگی‌های فعال نوری آن معمولاً پتانسیل بالایی برای ارزیابی با تکنیک سنجش از دور دارد (Hossain et al., 2021). سنجش از دور کدورت، براساس میزان شار انرژی نوری از رنگ آب و عوامل ایجاد کدورت آب است. ذرات معلق، عمدتاً مسئول پراکندگی نور هستند، در حالی که مواد آلی محلول رنگی و کلروفیل‌آ خواص جذب نور آب را کنترل می‌کنند (Wass et al., 1997). رنگ، جذب نور در آب را افزایش و سیگنال سنجش از راه دور را کاهش می‌دهد زیرا انرژی بیشتری از خورشید در آب جذب می‌شود. افزایش کدورت آب ناشی از ذرات معلق باعث افزایش پراکندگی نور و سیگنال سنجش از راه دور می‌گردد، زیرا انرژی خورشیدی بیشتری توسط ذراتی که کدورت آب را ایجاد می‌کنند منعکس یا پراکنده می‌شود. برای غلظت‌های کم مواد معلق، بازتاب طیفی عمدتاً تحت تاثیر ویژگی‌های جذب آب است و برای غلظت‌های بالاتر، ویژگی‌های نوری ذرات معلق مهم‌ترین عامل است. یکی از عمده‌ترین موارد ایجاد خطا در آب شفاف و کم‌عمق، افزایش سیگنال ناشی از بازتاب مقداری از انرژی خورشیدی از کف می‌باشد (Moore, 1980).

از دیگر مشکلات عمده در محاسبه کدورت آب به‌وسیله انرژی بازتاب، تفاوت در جذب و پراکندگی نور در جو است که تصحیح جوی $DOS1^1$ می‌تواند تصحیح مناسبی برای این خطا باشد (Moore, 1980). در این روش، با این فرض که اشیاء تاریک نوری را منعکس نمی‌کنند، در هر باند تاریک‌ترین مقدار پیکسل جستجو می‌شود، در نتیجه هر مقدار بزرگتر از صفر باید از پراکندگی اتمسفر حاصل شود. اثر پراکندگی با کم کردن این مقدار از هر پیکسل در باند حذف می‌شود (El-Zeiny & El-Kafrawy, 2017).

نکته مهم دیگر این است که یک شار پراکنده شده ممکن است ترکیبی از رنگ آب، بازتاب کف، کدورت ایجاد شده توسط فیتوپلانکتون‌ها یا ناشی از ذرات معلق باشد. اما اغلب، تغییر در شار به دلیل تغییر در غلظت یک جزء ایجاد می‌شود که این مشکل ممکن است در روش اندازه‌گیری کدورت آب میدانی نیز وجود داشته باشد. ذکر این نکته ضروری است که حتی سطح انرژی اندازه‌گیری شده به‌وسیله سنجش از راه دور ممکن است دقیق‌تر از کدورت متوسط و نزدیک به سطح، نسبت به تجزیه آزمایشگاهی از یک نمونه نقطه‌ای باشد (Moore, 1980).

برقراری همبستگی و تحلیل‌های رگرسیونی با الگوریتم‌های تک یا چندباندی از روش‌های رایج برای ایجاد روابط بین بازتاب طیفی و اندازه‌گیری‌های کدورت آب در محل هستند (Harrington et al., 1992). تفاوت در رنگ و کدورت آب بر سیگنال‌ها در طول موج‌های

مرئی و بسیار نزدیک ماوراء بنفش و مادون قرمز تاثیر می‌گذارد. هم‌چنین در منابع ذکر شده است که بازتاب در ناحیه مرئی به ویژه ناحیه قرمز با افزایش رسوبات در آب یا کدورت آب افزایش می‌یابد (Ritchie et al., 2003; Pavelsky & Smith, 2009; Gholizadeh et al., 2016; Garg et al., 2020) و پیک بازتاب از ناحیه سبز به طیف قرمز تغییر می‌کند (M. Gholizadeh et al., 2016). مطالعات بیشتر نشان داده است که استفاده از ترکیب‌های باندی برای تخمین کدورت آب بهتر از تک باند است و مقادیر خطا را کاهش می‌دهد (Sebastiá-Frasquet et al., 2019).

تحقیقات متعددی برای یافتن الگوریتم‌های بازبانی کدورت آب با استفاده از ماهواره‌های مختلف صورت گرفته است که نشان می‌دهند، تصاویر ماهواره‌های LISS-III و IRS و باندهای ۱ و ۲ و ۳ در این ماهواره‌ها (Aghighi et al., 2008)، تصاویر ماهواره لندست-۵ و باندهای ۲، ۳ و ۴ این ماهواره (Nas et al., 2010; Atif et al., 2018) و باندهای مرئی آن (Akbar et al., 2010; Kalele A., 2019) و ماهواره سنتینل-۲ و باندهای مرئی و نزدیک مادون قرمز در این سنجنده‌ها با مقدار کدورت آب همبستگی خوبی دارند (Ellero et al., 2019) و در این مطالعات همبستگی مقادیر کدورت با بازتاب امواج باندهای ماهواره‌ای با مقادیر اندازه‌گیری‌های میدانی اعتبارسنجی شده است. (M., 2018; Katlane et al., 2020; Chen et al., 2022).

البته در برخی پژوهش‌ها از پیش‌بینی کدورت آب برای ارزیابی کلی شرایط کیفیت آب در شرایطی که داده‌های میدانی وجود ندارد، به‌عنوان مثال در طول قرنطینه کووید-۱۹، هم استفاده شده است. در این روش از شاخص‌هایی همچون شاخص کدورت تفاوت نرمال شده NDTI^۱ جهت بررسی کدورت آب در غیاب داده‌های میدانی استفاده می‌شود. در محاسبه این شاخص از مقادیر بازتاب نوار سبز و قرمز استفاده شده و با موفقیت در مطالعات استفاده شده است (Garg et al., 2020; Chen et al., 2022).

در سال‌های اخیر، تکنیک‌های یادگیری ماشین همچون جنگل تصادفی، رگرسیون بردار پشتیبان و شبکه‌های عصبی مصنوعی به کاربردهای زیست محیطی سنجش از دور برای تخمین پارامترهای کیفی آب مانند کلروفیل آ (Hafeez et al., 2019) و پارامترهای غیرفعال نوری مانند نیتروژن کل و فسفر کل (Huang et al., 2020) کمک می‌کنند. البته استفاده از این روش‌ها در برآورد شفافیت و کدورت آب محدودتر می‌باشد (Rubin et al., 2021).

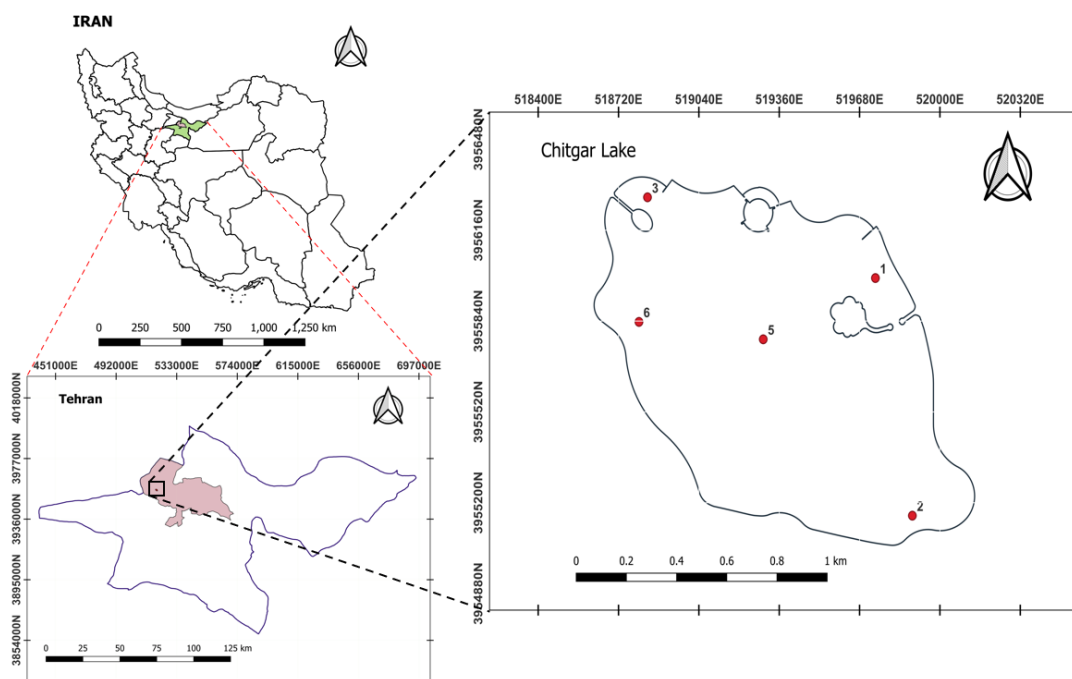
پایش کیفیت آب دریاچه‌های با کاربری تفرج مانند دریاچه شهدای خلیج فارس چیتگر که اغلب کم‌عمق بوده و در محدوده‌های شهری قرار دارند، به دلیل موقعیت جغرافیایی و اهمیت بالای کیفیت آب برای گردشگران، از اهمیت ویژه‌ای برخوردار است. چرا که افزایش کدورت و شکوفایی جلبکی در این پهنه‌ها می‌تواند باعث شکل‌گیری مناظر نامناسب و اثرات نامطلوب بر فعالیت‌های تفریحی و گردشگری، انتشار بوی نامطبوع و سموم جلبکی، ایجاد مناطق عاری از اکسیژن و تهدید حیات آبیان (Bayat et al., 2019) گردد. عمق کم این گونه دریاچه‌ها، پهنه آبی را در برابر نوسانات آب و هوایی حساس‌تر نموده و بر اختلاط آب، غلظت کلروفیل آ و کدورت آب تاثیر می‌گذارد (Bohn et al., 2018). به این دلایل، مدیریت اثربخش این پهنه‌های آبی نیازمند پایش با فرکانس بالای کیفیت آب در گستره‌ای وسیع می‌باشد که اغلب نیازمند منابع انسانی، مالی و تجهیزاتی فراوان می‌باشد. به دلیل اهمیت مدیریت کیفیت آب و لزوم پایش با وضوح بالای زمانی و مکانی آن در دریاچه چیتگر تهران، این پژوهش جهت بررسی کارایی فناوری سنجش از دور در تخمین میزان کدورت آب، با به‌دست آوردن مدل برآوردی کدورت و با کمک الگوریتم جنگل تصادفی RF^۲، انجام شده است.

برآورد کدورت آب با تکنیک سنجش از دور، در آب‌های داخلی و به‌ویژه دریاچه‌های شهری به صورت محدود صورت گرفته است. علاوه بر آن دریاچه چیتگر به دلیل مصنوعی و کم‌عمق بودن منحصر به فرد است که این ویژگی‌ها از عوامل موثر بر کاهش دقت و عملکرد مدل‌های پیش‌بینی به سبب اثر بازتاب از کف و جنس بستر می‌باشد. این تحقیق سعی در ارائه روشی برای به دست آوردن مدل پهنه پیش‌بینی کدورت آب با دقت بالاتر، به وسیله تلفیق فناوری سنجش از دور و روش یادگیری ماشین با قابلیت کاهش اثر مشکلات ناشی از شرایط جغرافیایی و ویژگی‌های ذاتی پهنه موردنظر بر مدل‌های پیش‌بینی را دارد.

۲. روش‌شناسی پژوهش

۲-۱. محدوده مطالعه (دریاچه شهدای خلیج فارس چیتگر)

دریاچه شهدای خلیج فارس چیتگر، دریاچه‌ای مصنوعی و کم‌عمق در شمال غرب شهر تهران و در شمال پارک جنگلی چیتگر در منطقه ۲۲ شهرداری، با مساحت ۱۳۲ هکتار و حجم ۶/۸ میلیون مترمکعب است. عمق دریاچه از صفر در ساحل شمالی تا نه متر در جنوب متغیر و میانگین آن پنج متر می‌باشد. آب دریاچه با آبیگری از رودخانه کن در ماه‌های خاصی از سال تامین می‌شود و تبخیر آب تنها خروجی آن است (Bayat et al., 2019). این دریاچه از مهمترین اماکن گردشگری در تهران می‌باشد که اهدافی همچون، برگزاری برنامه‌های تفریحی مانند ماهیگیری و مسابقات ورزشی، تلطیف هوای منطقه، توسعه اقتصادی منطقه و همچنین توسعه امکانات و زیرساخت‌های تجاری را نیز برآورده می‌سازد. برای حفظ کیفیت آب دریاچه و جلوگیری از پدیده شکوفایی جلبکی، تصفیه‌خانه دریاچه در پاییز ۱۳۹۵ راه‌اندازی شده است (Bayat et al., 2019) که در فصل آبیگری دریاچه، کدورت ناشی از فرآیند آبیگری و ذرات معلق را تحت کنترل قرار می‌دهد. شکل (۱) موقعیت دریاچه و ایستگاه‌های پایش آن را نشان می‌دهد.



شکل ۱. دریاچه شهدای خلیج فارس چیتگر و ایستگاه‌های پایش آن، تهران، ایران (Bayat et al., 2019)

۲-۲. داده‌های ماهواره‌ای و تصحیحات مقدماتی

در این پژوهش، از تصاویر ماهواره‌های لندست-۸ (C1L1) و سنتینل-۲ (LIC) به‌عنوان داده‌های سنجش از دور استفاده شده است. از دلایل انتخاب این تصاویر می‌توان به قدرت تفکیک مکانی، رادیومتریک، طیفی و زمانی و نسبت سیگنال به نویز (SNR) بالای تصاویر این ماهواره‌ها و دسترسی رایگان به این داده‌ها اشاره نمود (Buma & Lee, 2020). تصاویر ماهواره‌های لندست-۸ و سنتینل-۲ منطبق و با حداکثر یک روز اختلاف با تاریخ نمونه‌برداری زمینی، با پوشش ابر کمتر از ۱۰ درصد در محدوده مورد مطالعه، از سایت سازمان زمین‌شناسی آمریکا دریافت شدند و سپس تصحیحات اتمسفری و هندسی بر روی تصاویر انجام شد. تصحیح اتمسفری تصاویر ماهواره‌ای به منظور حذف اثرات جوی یا آتروسول پیش از پردازش ضروری است. این تصحیحات در گستره‌های مکانی و زمانی بزرگ اهمیت دارند. زیرا، شرایط جوی می‌تواند به‌طور قابل توجهی متفاوت و متغیر باشد (Topp et al., 2020). همچنین برای داشتن خواص نوری واقعی آب در آب‌های کم‌عمق، اعمال تصحیح اتمسفری و حذف اثرات بازتاب کف بسیار مهم است (Katlane et al., 2020) این تصحیحات در مرحله

پیش پردازش تصاویر در نرم افزار QGIS (3.20.1) و SCP-Plugin انجام شد. روش تصحیح اتمسفری در این نرم افزار روش الگوریتم تصحیح DOS1 می باشد.

۳-۲. داده های زمینی

داده های زمینی کدورت آب، از داده های پایش کیفیت دریاچه از سال ۱۳۹۵ تا ۱۴۰۰ در ایستگاه های پایش استخراج شده است. پس از انطباق داده های زمینی و تاریخ تصاویر ماهواره ای منطبق و با حداکثر اختلاف یک روز ۱۰۵ و ۱۰۴ سری داده به ترتیب برای ماهواره ی لندست-۸ و سنتینل-۲ جهت ارزیابی پارامتر کدورت آب استخراج شد. بازه مقادیر در داده های لندست-۸ در محدوده ۰/۸۴ الی ۹/۷۵ NTU و در ماهواره سنتینل-۲ بین ۰/۸۳ الی ۱۴/۴ NTU قرار داشتند. سپس توزیع آماری داده ها با کمک آزمون های شپرو-ویلک^۱ و Q-Q Plot در نرم افزار R ویرایش ۴,۱,۲ بررسی شد. توزیع داده ها در هر دو ماهواره نرمال نبوده و پس از تبدیل لگاریتم طبیعی با $(p>0.05)$ نرمال شدند. در مرحله بعد، داده ها به دو دسته جهت تولید و اعتبارسنجی مدل به ترتیب با نسبت های $\frac{2}{3}$ و $\frac{1}{3}$ تقسیم شدند (Wang et al., 2020; He et al., 2021; Kapalanga et al., 2021).

۴-۲. روش جنگل تصادفی (RF)

جنگل تصادفی یک الگوریتم یادگیری ماشین از نوع «یادگیری ماشین نظارت شده» است که مبتنی بر طبقه بندی و رگرسیون از طریق تعداد زیادی درخت تصمیم به عنوان طبقه بندی کننده است که از روش بسته بندی و به طور تصادفی درخت های تصمیم گیری مرتبط را انتخاب می کند (Li et al., 2021) و در حالی که از داده ها برای ساخت مدل های متعدد استفاده می کند. نتایج مدل سازی همه مدل ها را با طبقه بندی یا میانگین گرفتن (رگرسیون) یکپارچه می کند، به طوری که نتایج مدل نهایی دارای دقت عملکرد و تعمیم بالایی باشد. در اجرای جنگل تصادفی، تعداد مساوی نمونه آموزشی به طور تصادفی از همه نمونه ها استخراج می شود، سپس m مجموعه ویژگی های فرعی از تمامی ویژگی های ورودی انتخاب می شود و در نهایت یک ویژگی بهینه از مجموعه ویژگی های فرعی برای تقسیم گره انتخاب می شود. هنگام حل مسایل رگرسیون، هر درخت تصمیم یک درخت رگرسیون است (Huang et al., 2020).

یکی از مزایای الگوریتم جنگل تصادفی این است که اندازه گیری اهمیت نسبی هر ویژگی روی مدل پیش بینی در آن آسان است. این ابزار، اهمیت یک ویژگی را با در نظر گرفتن تعداد گره های درخت که از آن ویژگی استفاده می کنند، محاسبه می کند. از طریق بررسی اهمیت ویژگی ها، می توان ویژگی هایی که به اندازه کمتر در فرآیند تصمیم گیری نقش دارند یا نقشی ندارند را حذف کرد (IBM, 2020; ESRI, 2022) که از این مزیت الگوریتم جنگل تصادفی در این پژوهش برای تعیین متغیرهای مهم استفاده شده است. روش جنگل تصادفی با تعیین تابع مهم متغیر برای شناسایی باندهای مهم برای مدل های برآورد کدورت آب در دریاچه چیتگر اجرا شد. بعد از نصب بسته جنگل تصادفی در نرم افزار R داده های زمینی و ماهواره ای از دسته داده های آموزشی (تولید مدل) مورد تجزیه و تحلیل قرار گرفت. ابتدا تعداد درختان تصمیم ۵۰۰ در نظر گرفته شد و میزان خطای مدل رگرسیون پیشنهادی نرم افزار ثبت گردید. با تغییر تعداد درختان تصمیم، الگوریتم RF مجدداً اعمال شده و با رسیدن به کمترین مقدار خطای مدل، باندهای مهم پس از تثبیت مقادیر خطا مورد انتخاب نهایی قرار گرفت. بدین ترتیب، مهمترین باندها با استفاده از ضرایب معنی دار پارامترهای مدل که همان مقادیر داده های باندهای ماهواره ای هستند، شناسایی شدند. باندهای مشخص شده، در تعیین الگوریتم و سپس رابطه مدل پیش بینی مورد استفاده قرار گرفته است.

۵-۲. دقت و توان آماری

ارزیابی دقت مدل های برآوردی با استفاده از ریشه میانگین مربعات خطا (RMSE) (رابطه ۱) و ریشه میانگین مربعات خطای نرمال شده (NRMSE) (رابطه ۲) انجام شد. دلیل استفاده از ریشه میانگین مربعات خطا، یکسان بودن بُعد و مقیاس آن با ویژگی هدف است (Lim &

1. Shapiro-Wilk

می‌باشد. در این تحقیق، معیار توان آماری (Power) مدل که نشان‌دهنده اطمینان از داده‌های ورودی درست از نظر حجم نمونه کافی برای کمترین خطا است. از رابطه (۳) محاسبه شد. توان مدل با تعداد نمونه‌ها و انحراف معیار داده‌ها مرتبط است (Verzani, 2004).

$$RMSE = \sqrt{\frac{\sum_i^n (Y_{obs} - Y_{est})^2}{n}} \quad \text{رابطه (۱)}$$

$$NRMSE = \frac{RMSE}{MAX(Y_{obs}) - MIN(Y_{obs})} \quad \text{رابطه (۲)}$$

در این روابط Y_{obs} غلظت اندازه‌گیری شده در نقطه i تخمین مدل از غلظت پارامتر در نقطه i و n تعداد کل نقاط اعتبارسنجی می‌باشد. در این تحقیق، معیار توان آماری (Power) مدل که نشان‌دهنده اطمینان از داده‌های ورودی درست از نظر حجم نمونه کافی برای کمترین خطا است. از رابطه (۳) محاسبه شد. توان مدل با تعداد نمونه‌ها و انحراف معیار داده‌ها مرتبط است (Verzani, 2004).

$$Power = 1 - \beta \quad \text{رابطه (۳)}$$

که در آن β خطای نوع دوم است و به‌طور معمول برای بازه اطمینان ۹۵ درصد، ۲۰ درصد در نظر گرفته می‌شود و توان آماری مطلوب ۸۰ درصد به‌دست می‌آید (Verzani, 2004). همچنین معمولاً جهت بیان دقت مدل‌های خطی مقدار R^2 گزارش می‌شود که نسبت واریانس متغیر وابسته را که با مدل رگرسیون در نمونه محاسبه می‌شود، نشان می‌دهد. مقدار R^2 به‌طور سیستماتیک مقدار واریانس توضیح داده شده در جامعه هدف را خوش‌بینانه تخمین می‌زند، برای تخمین مقدار واریانس توضیح داده شده، معمولاً از R^2 تعدیل شده^۱ استفاده می‌شود که واقعی‌تر بوده و معمولاً مقدار آن کمتر و در بهترین حالت مساوی R^2 است (Karch, 2020). در این تحقیق از $Adj.R^2$ برای بیان دقت مدل‌های رگرسیون استفاده شده است. اعتبار مدل‌های برآوردی با کمک این معیارهای آماری و با استفاده از داده‌های اعتبارسنجی، ارزیابی شده و به انتخاب مدل پیش‌بینی بهینه منتج شده است.

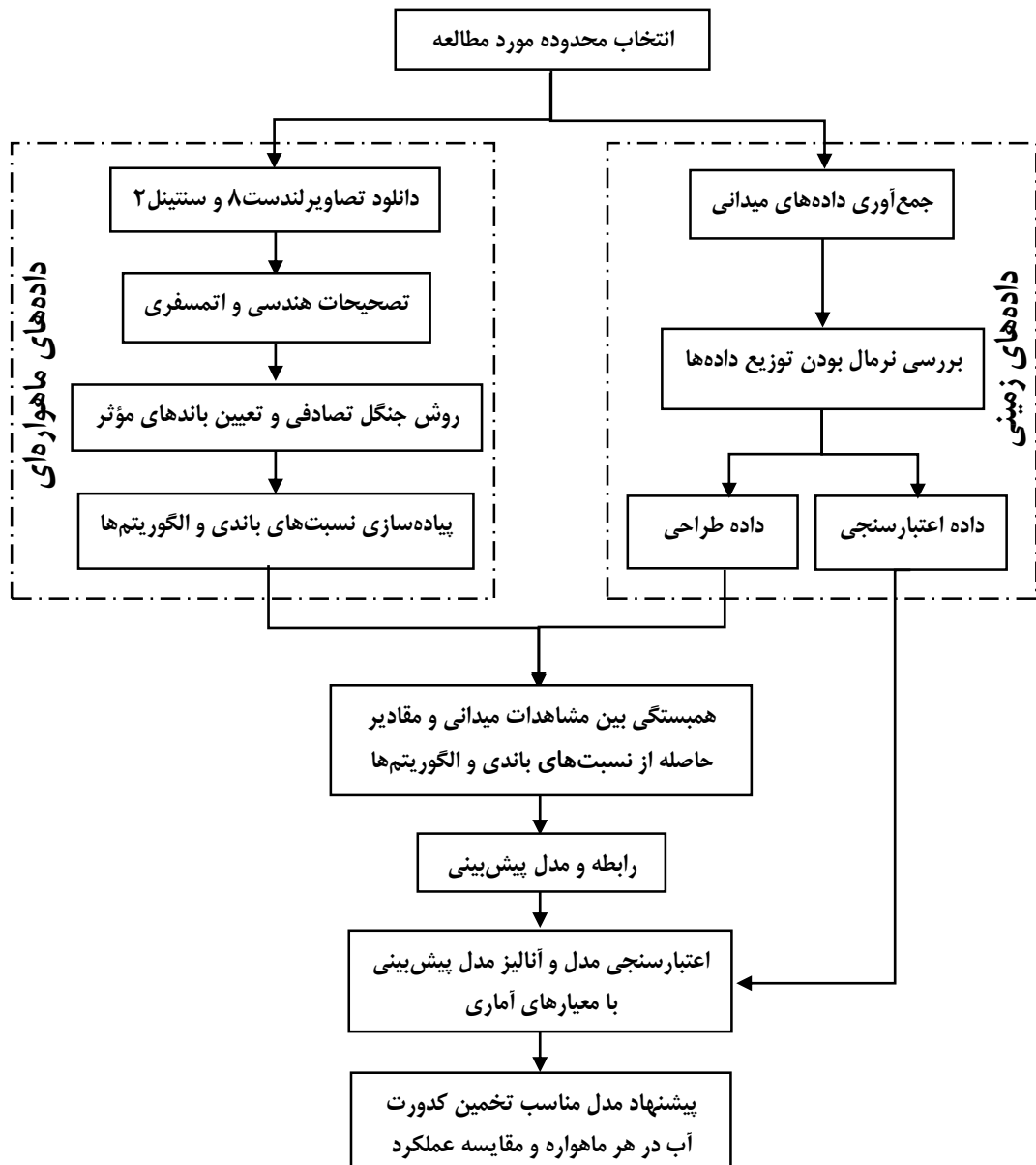
۶-۲. مراحل انجام پژوهش

شماتیک نمودار جریان کار این پژوهش در شکل (۲) نشان داده شده است. توضیح مراحل در زیربندهای همین بخش تشریح شده است (رجوع شود به زیر بندهای ۱-۲ تا ۵-۲).

۳. یافته‌های پژوهش

پس از انجام پیش‌پردازش بر روی تصاویر ماهواره‌ای، داده‌های ماهواره‌ای و داده‌های زمینی در محیط نرم‌افزار R وارد شدند و با استفاده از بسته و برنامه جنگل تصادفی، برای تعیین باندهای با درجه اهمیت بالاتر در تخمین کدورت آب بوسیله محاسبه ضرایب متغیرهای مهم، مورد تحلیل و ارزیابی قرار گرفتند. جدول (۱) ضرایب اهمیت نسبی باندهای ماهواره لندست-۸ را که توسط روش جنگل تصادفی به‌دست آمده نشان می‌دهد. براساس این ضرایب در ماهواره لندست-۸ باندهای $b_2(482 \text{ nm})$ و $b_3(562 \text{ nm})$ به‌عنوان باندهای موثر شناسایی شدند. که با توجه به مرور منابع صورت گرفته، با نتایج تحقیقات مشابه انطباق داشت. سپس الگوریتم‌های تک باندی و نسبت‌های باندی بین این باندها در روابط و مدل‌های رگرسیون خطی و غیرخطی ارزیابی شد و نسبت باندی $\frac{b_2}{b_3}$ در معادله‌نمایی بهترین پیش‌بینی را برآورد کرد.

^۱ $Adj.R^2$ (Adjusted R^2)



شکل ۲. نمودار جریان و روش انجام پژوهش

در جدول (۲) ضرایب اهمیت نسبی به‌دست آمده برای باندهای ماهواره سنتینل-۲ از روش جنگل تصادفی درج شده است. با انطباق این باندها با نتایج تحقیقات مشابه، باندهای با ضرایب اهمیت نسبی بالاتر که در انتخاب الگوریتم‌ها از آنها استفاده شده، تفکیک شده است. که باندهای $b_2(490 \text{ nm})$ ، $b_3(560 \text{ nm})$ و $b_4(665 \text{ nm})$ به‌عنوان باندهای مؤثر استخراج گردید و بهترین نتیجه در این ماهواره از ترکیب سه باند در یک معادله نمایی به‌دست آمد. سپس اعتبار روابط به‌دست آمده با داده‌های صحت‌سنجی بررسی شد. نتایج آماری معادلات اعتبارسنجی و الگوریتم‌های وابسته در جدول (۳) بیان شده است. همچنین پس از به‌دست آمدن مدل‌ها، نرمال بودن توزیع مقادیر باقیمانده‌ها^۱ برای تعیین بهترین مدل برازش شده، با آزمون شیپرو-ویلک بررسی شد (Verzani, 2004).

جدول ۱. ضرایب اهمیت نسبی باندهای ماهواره لندست-۸ در تخمین کدورت آب، مشخص شده با روش RF

شماره باند	۱	۲	۳	۴	۵	۶	۷
ضریب اهمیت	۲/۵۸	۲/۹۱	۳/۸۷	۲/۸۸	۲/۵۰	۱/۵۳	۱/۵۸

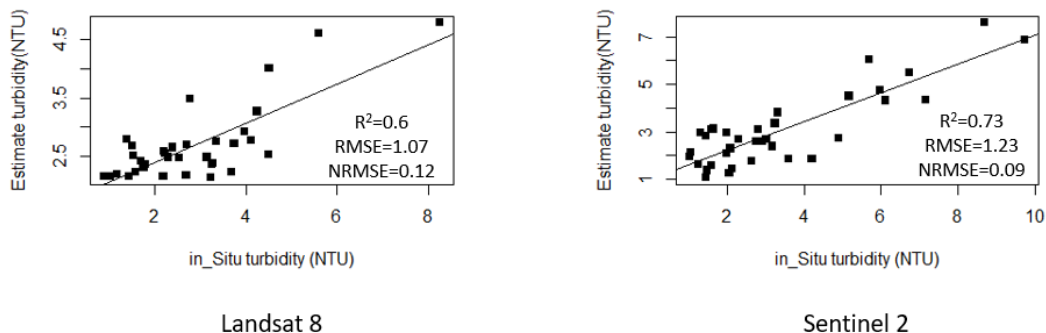
جدول ۲. ضرایب اهمیت نسبی باندهای ماهواره سنتینل-۲ در تخمین کدورت آب، مشخص شده با روش RF

شماره باند	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۱	۱۲	
ضریب اهمیت	۱/۴۵	۲/۶۸	۶/۹۰	۲/۱۲	۱/۴۷	۲/۰۲	۲/۱۰	۱/۴۵	۱/۵۹	۰/۹۰	۱/۳۹	۱/۲۲

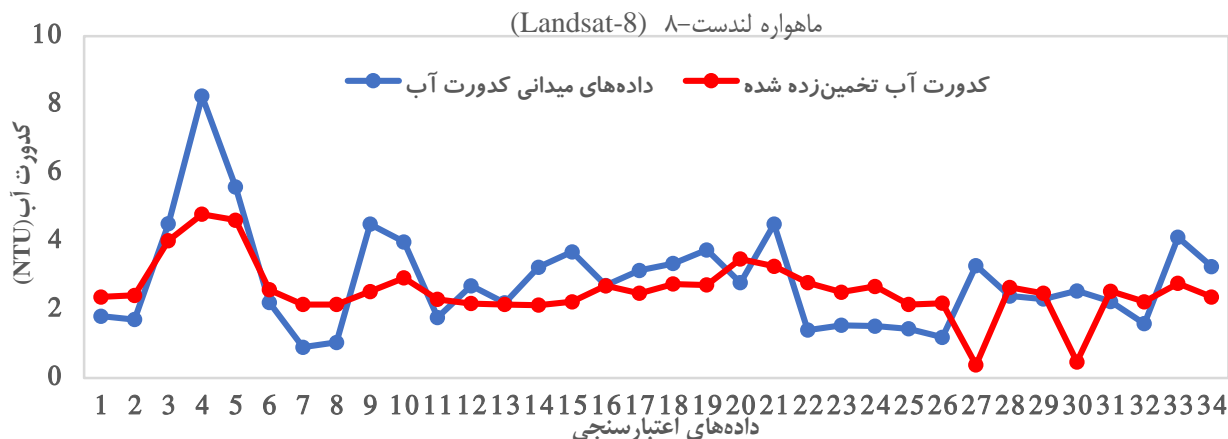
جدول ۳. نتایج آماری الگوریتم‌های تخمین کدورت آب مورد استفاده در دریاچه شهدای خلیج فارس چیتگر

ماهواره	الگوریتم	معادله	R ²	Adj. R ²	RMSE NTU	NRMSE	Power
Landsat 8 OLI/TIRS	$\frac{b2}{b3}$	$Ln(y) = 8 \times x^2 - 17.2 \times x + 10$	۰/۶	۰/۶	۱/۰۷	۰/۱۲	۰/۸
Sentinel 2 MSI	b2, b3, b4	$Ln(y) = -21.1 \times b2 + 65.1 \times b3 - 35.8 \times b4$	۰/۷۳	۰/۷۳	۱/۲۳	۰/۰۹	۰/۸

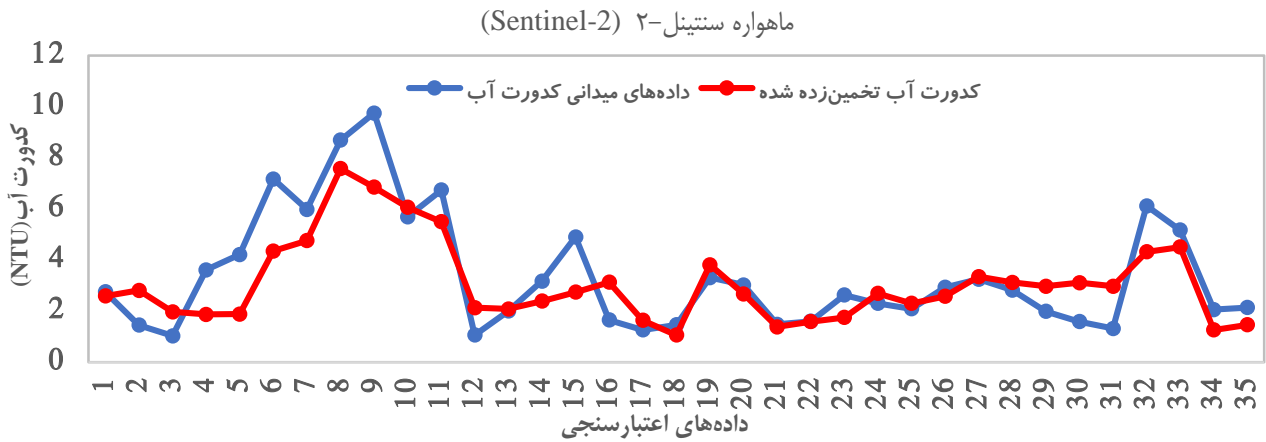
شکل شماره (۳) نمودار پراکندگی بین مقادیر برآوردی مدل و نمونه‌های میدانی را در محدوده داده‌های اعتبارسنجی برای هر دو ماهواره نمایش می‌دهد. و شکل‌های (۴ و ۵) نمودار خطی مقایسه‌ای بین مقادیر اندازه‌گیری میدانی و برآورد ماهواره‌ای را برای داده‌های اعتبارسنجی نشان می‌دهد.



شکل ۳. همبستگی میان داده‌های میدانی و مقادیر برآوردی کدورت آب در ماهواره‌های لندست-۸ و سنتینل-۲ در دریاچه چیتگر

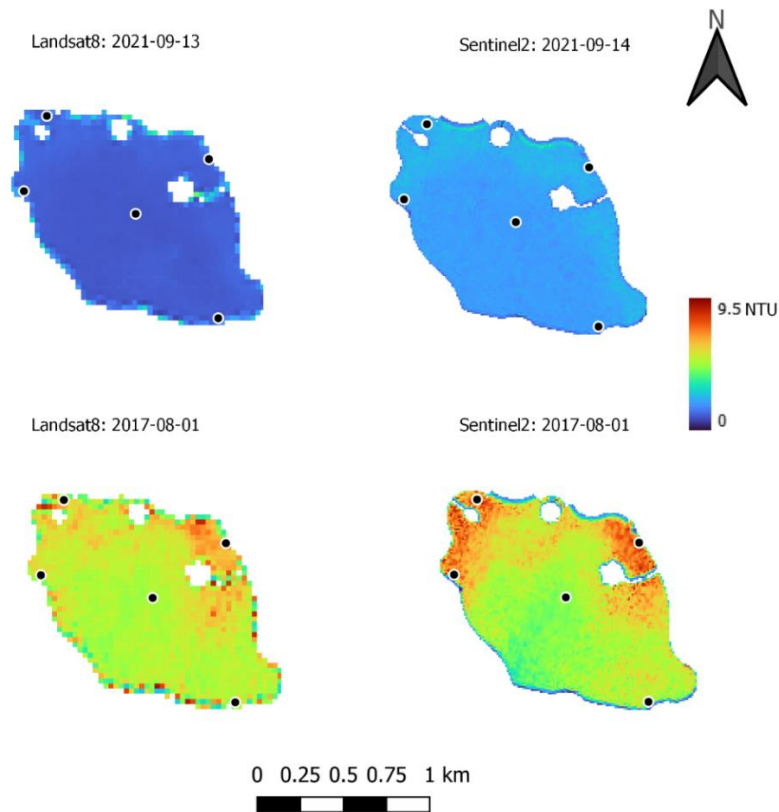


شکل ۴. نمودار خطی بین داده‌های میدانی و برآورد کدورت آب با ماهواره لندست-۸ در دریاچه چیتگر تهران، ایران



شکل ۵. نمودار خطی بین داده‌های میدانی و برآورد کدورت آب با ماهواره سنتینل-۲ در دریاچه چیتگر تهران، ایران

همچنین شکل شماره (۶) توزیع مکانی کدورت آب دریاچه را براساس مدل‌های پیش‌بینی به‌دست آمده در هر دو ماهواره، در دو محدوده مقداری پایین و بالا نمایش می‌دهد.



شکل ۶. توزیع مکانی کدورت دریاچه شهدای خلیج فارس چیتگر تهران با ماهواره‌های لندست-۸ و سنتینل-۲

۴. بحث

بررسی نتایج آماری و نقشه‌های توزیع مکانی نشان می‌دهد که ماهواره سنتینل-۲ دارای عملکرد بهتری در تخمین مقادیر کدورت آب در دریاچه می‌باشد. از سویی دیگر، با توجه به نتایج معادلات، می‌توان از ماهواره لندست-۸ با دقت قابل قبولی، در تاریخ‌هایی که تصویر

ماهواره‌ای سنتینل-۲ در دسترس نیست و به منظور پوشش زمانی بهتر جهت پیش‌بینی و مشاهده روند مقدار کدورت آب استفاده کرد. در جدول شماره (۴) الگوریتم، روابط و نتایج مورد استفاده در برخی از مطالعات برای تعیین کدورت در پهنه‌های آبی مختلف با استفاده از سنجنده‌ها عنوان شده است. بررسی مطالعات پیشین نشان می‌دهد که در اکثر مطالعات باندهای مورد استفاده در برآورد کدورت آب در ماهواره لندست شامل باندهای ناحیه مرئی و نزدیک مادون قرمز یعنی باندهای ۲ الی ۵ هستند و در ماهواره سنتینل-۲ نیز استفاده از باندهای مرئی نتایج خوبی در تخمین مقدار کدورت داشته‌اند. همچنین استفاده از باندهای مرئی و نزدیک مادون قرمز در ماهواره سنتینل-۲ و در شاخص تفاوت نرمال شده کدورت (NDTI) نتایج قابل قبولی در برآورد کدورت آب در غیاب داده‌های میدانی دارند (Garg et al., 2020; Chen et al., 2022). این توافق نتایج در باندهای موثر در الگوریتم‌های تخمین کدورت آب در این مطالعه نیز تایید شده است. تفاوت در الگوریتم‌ها و روابط به دلیل تفاوت در نوع پهنه آبی مورد بررسی، و علاوه بر آن متفاوت بودن عوامل ایجاد کدورت، تاثیر عمق، جنس بستر و فصل نمونه‌برداری در منابع آب‌های سطحی می‌باشد. در پژوهش گوانگ و همکاران (Quang et al., 2017) بیان شده است که در آب‌های کم‌عمق و فصول خشک تعلیق رسوبات بستر عامل ایجاد کدورت هستند.

جدول ۴. تحقیقات انجام شده در پهنه‌های آبی مختلف برای پیش‌بینی کدورت آب (WT) و نتایج الگوریتم‌های مورد استفاده

منبع	محدوده مورد مطالعه	الگوریتم	معادله	نتایج
(Aghighi et al., 2008)	Gorgan Bay-Iran	b_1, b_2, b_3	$WT = e^{a*x_1+b*x_2+c*x_3+d}$	$R^2(IRS) = 0.75$
(Nas et al., 2010)	Lake Beysehir-Turkey	b_2, b_3, b_4	$WT = a * x_1 + b * x_2 + c * x_3 + d$	$R^2(Landsat5) = 0.56$
(Quang et al., 2017)	Cam Ranh Bay and Thuy Trieu Lagoon-Vietnam	b_4	$WT = a * x + b$	$R^2(Landsat8) = 0.84$
(M. H. Gholizadeh & Melesse, 2017)	Florida Bay-USA	b_2, b_3, b_4, b_5	$WT = a_1 * x_1 + a_2 * x_2 + a_3 * x_3 + a_4 * x_4 + b$	$R^2(Landsat8)_{Dry} = 0.84$ $R^2(Landsat8)_{Wet} = 0.63$
(Ellero, 2018)	Umdloti Estuary, KwaZulu-Natal, Africa	b_2, b_3, b_4	$WT = a_1 * x_1 + a_2 * x_2 + a_3 * x_3 + b$	$R^2(Sentinel2) = 0.99$
(Atif et al., 2018)	Haleji Lake-Pakistan	b_2, b_3	$WT = a * x_1 + b * x_2 + c$	$R^2(Landsat5) = 0.83$
(Kalele, 2019)	Charles River-UK	b_2, b_3	$WT = a * x_1 + b * x_2 + c$	$R^2(Landsat8) = 0.84$
(Katlane et al., 2020)	Kneiss Archipelago Gulf of Gabes-Tunisia	b_4	$WT = a * x_1 + b$	$R^2(Sentinel2) = 0.7$
(Kapalanga et al., 2021)	Olushandja Dam-Namibia	b_2, b_3, b_4, b_5	$WT = a_1 * x_1 + a_2 * x_2 + a_3 * x_3 + a_4 * x_4 + b$	$R^2(Landsat8) = 0.98$
(Hossain et al., 2021)	Tennessee River-USA	b_4	$WT = a * x^b$	$R^2(Landsat8) = 0.95$
این مطالعه	Chitgar Lake	$\frac{b_2}{b_3}$	$WT(L8) = e^{a*x^2+b*x+c}$	$R^2(Landsat8) = 0.6$
		b_2, b_3, b_4	$WT(S2) = e^{a*x_1+b*x_2+c*x_3}$	$R^2(Sentinel2) = 0.73$

در اغلب پژوهش‌ها برآورد میزان کدورت آب به دلیل خواص نوری عوامل ایجادکننده آن، با دقت بالایی صورت گرفته است. در بررسی کدورت آب در این مطالعه دقت برآورد و باندهای موثر تا حدودی به مطالعات در دریاچه‌های کم‌عمق و یا آب‌های داخلی کم‌عمق مشابهت دارد. از جمله به نتایج مطالعه نس و همکاران (Nas et al., 2010) بر روی دریاچه کم‌عمق در ترکیه نزدیک‌تر می‌باشد که در تحقیق نس و همکاران R^2 معادل ۰/۵۶ بوده که در تخمین کدورت همبستگی قابل قبولی تلقی شده است. همچنین در پژوهشی مشابه بر روی دریاچه‌ای کم‌عمق در پاکستان، معادلات رگرسیونی بین باندهای آبی، سبز و قرمز در ماهواره لندست-۵، همبستگی‌های R^2 بزرگتر از ۰/۷ را نشان داده است. و در نهایت باندهای ۲ و ۳ (سبز و قرمز) در این ماهواره بالاترین همبستگی را نشان داده است (Atif et al., 2018). در مطالعه‌ای که جهت برآورد کدورت در آب‌های کم‌عمق داخلی در تونس انجام شده است به اهمیت تاثیر عمق بر کدورت و نیز خطاهای ناشی از

بازتاب کف اشاره شده است (Katlane et al., 2020) در پژوهش ذکر شده از ماهواره سنتینل-۲ جهت برآورد استفاده شده که R^2 برابر 0.73 به دست آمد که با توجه به شرایط پهنه آبی همبستگی بالایی بوده است. در مطالعه حاضر نیز در ماهواره سنتینل-۲ ضریب تعیین $R^2=0.73$ و نتایج به دست آمده، مؤید دقت قابل قبولی در تخمین کدورت آب در دریاچه شهدای خلیج فارس چیتگر می باشد. از عوامل موثر بر دقت روابط، عوامل ایجاد خطا می باشند که از میان این خطاها، دقیق نبودن محل ایستگاه های نمونه برداری، خطای روش نمونه گیری، خطای تصحیح جوی تصاویر و استفاده از تصاویر با یک روز اختلاف با نمونه برداری میدانی حائز اهمیت هستند. لازم به ذکر است که عمق دریاچه به دلیل اثرات بازتاب امواج از کف از عوامل بسیار موثر ایجاد خطا در تخمین کدورت در پهنه های آبی می باشد (Moore, 1980). که با توجه به کم عمق بودن دریاچه چیتگر و این که سه ایستگاه نمونه برداری ۱ و ۳ و ۶ در قسمت کم عمق تر دریاچه قرار دارند، خطای بازتاب از کف در این دریاچه غیر قابل اجتناب می باشد.

۵. نتیجه گیری

در این پژوهش کارایی و دقت برآورد دو ماهواره لندست-۸ و سنتینل-۲ برای تخمین مقدار کدورت آب با استفاده از مدل های پیش بینی و با کمک روش جنگل تصادفی در دریاچه شهدای خلیج فارس چیتگر بررسی شد تا با مقایسه عملکرد این دو سنجنده در تخمین کدورت آب، بتوان امکان پایش این پارامتر کیفی با استفاده از سنجش از دور را در این دریاچه بررسی کرد. لازم به ذکر است پایش کدورت آب دریاچه چیتگر با سنجش از دور، به دلیل وسعت و عمق کم دریاچه و نیز مصنوعی بودن آن با چالش و محدودیت همراه است. نتایج تحقیق نشان داد که ماهواره سنتینل-۲ در مقایسه با ماهواره لندست-۸ عملکرد بهتری در تخمین کدورت آب در دریاچه را دارد. اما با توجه به دقت نسبتاً قابل قبول مدل پیش بینی به دست آمده در ماهواره لندست-۸ نیز می توان از این ماهواره برای پوشش زمانی بیشتر و بررسی روند، در پایش کدورت آب دریاچه استفاده کرد. همچنین با توجه به نتایج مطالعه، به کارگیری روش جنگل تصادفی به یافتن مدل های پیش بینی کمک کرد. استفاده از این روش در دریاچه شهدای خلیج فارس چیتگر می تواند اطلاعات مورد نیاز را در کل سطح دریاچه در اختیار قرار دهد و به دلیل پوشش زمانی بالاتر از پایش های میدانی، داده های به روز تر و به موقع تری جهت دستیابی به اهداف کیفی مدنظر و ارتقای فرایند نظارت و پایش، هم زمان با کاهش هزینه ها در دریاچه شهدای خلیج فارس چیتگر و در نتیجه کارایی و اثربخشی فرایند مدیریت را در پی داشته باشد.

۶. تشکر و قدردانی

بدین وسیله از همکاری مجموعه آزمایشگاه و تصفیه خانه دریاچه چیتگر، بالاخص جناب آقای مهندس جواد بیات در انجام این تحقیق قدردانی می گردد.

۷. منابع

- Aghighi, H., Alimohamma, A., Reza Sarad, M., & Ashourloo, D. (2008). Estimation of Water Turbidity in Gorgan Bay, South-East of Caspian Sea by Using IRS-LISS-III Images. *Pakistan Journal of Biological Sciences*, 11(5), 711–718. <https://doi.org/10.3923/pjbs.2008.711.718>
- Akbar, T. A., Hassan, Q. K., & Achari, G. (2010). A remote sensing-based framework for predicting water quality of different water sources. *Remote Sensing and Spatial Information Sciences*, 34(xxx).
- Atif, S., Syed Jamil Hasan, K., Suhaib bin, F., Saima, S., Adnan, A., Hafiz Uzair Ahmed, K., Aimen Fatima, A., & Fahad, A. (2018). Mapping Turbidity Levels in the Lake's Water Using Satellite Remote Sensing Technique. *International Journal of Economic and Environment Geology*, 9(3).
- Baughman, C., Jones, B., Bartz, K., Young, D., & Zimmerman, C. (2015). Reconstructing Turbidity in a Glacially Influenced Lake Using the Landsat TM and ETM+ Surface Reflectance Climate Data Record Archive, Lake Clark, Alaska. *Remote Sensing*, 7(10), 13692–13710. <https://doi.org/10.3390/rs71013692>

- Bayat, J., Hashemi, S. H., Zolfagharian, M., Emam, A., & Nooshabadi, E. Z. (2019). Water quality management of an artificial lake, case study: The lake of the Martyrs of the Persian Gulf. In *Sustainable and Safe Dams Around the World* (pp. 1442–1449). CRC Press. <https://doi.org/10.1201/9780429319778-127>
- Blix, K., Pálffy, K., Tóth, V. R., & Eltoft, T. (2018). Remote sensing of water quality parameters over Lake Balaton by using Sentinel-3 OLCI. *Water (Switzerland)*, *10*(10). <https://doi.org/10.3390/w10101428>
- Bohn, V. Y., Carmona, F., Rivas, R., Lagomarsino, L., Diovisalvi, N., & Zagarese, H. E. (2018). Development of an empirical model for chlorophyll-a and Secchi Disk Depth estimation for a Pampean shallow lake (Argentina). *The Egyptian Journal of Remote Sensing and Space Science*, *21*(2), 183–191. <https://doi.org/10.1016/j.ejrs.2017.04.005>
- Buma, W. G., & Lee, S.-I. (2020). Evaluation of Sentinel-2 and Landsat 8 Images for Estimating Chlorophyll-a Concentrations in Lake Chad, Africa. *Remote Sensing*, *12*(15), 2437. <https://doi.org/10.3390/rs12152437>
- Chen, X., Chen, W., Bai, Y., & Wen, X. (2022). Changes in turbidity and human activities along Haihe River Basin during lockdown of COVID-19 using satellite data. *Environmental Science and Pollution Research*, *29*(3), 3702–3717. <https://doi.org/10.1007/s11356-021-15928-6>
- Chu, H.-J., He, Y.-C., Chusnah, W. N., Jaelani, L. M., & Chang, C.-H. (2021). Multi-Reservoir Water Quality Mapping from Remote Sensing Using Spatial Regression. *Sustainability*, *13*(11), 6416. <https://doi.org/10.3390/su13116416>
- Ellero M. (2018). *Water Quality Assessment using Landsat 8 and Sentinel-2: A case study of the Umdloti Estuary, KwaZulu-Natal, South Africa*. KwaZulu Natal.
- El-Zeiny, A., & El-Kafrawy, S. (2017). Assessment of water pollution induced by human activities in Burullus Lake using Landsat 8 operational land imager and GIS. *The Egyptian Journal of Remote Sensing and Space Science*, *20*, S49–S56. <https://doi.org/10.1016/j.ejrs.2016.10.002>
- ESRI. (2022). *Forest-based Classification and Regression (Spatial Statistics)*. <https://pro.arcgis.com/en/pro-app/2.8/tool-reference/spatial-statistics/forestbasedclassificationregression.htm>
- Garg, V., Aggarwal, S. P., & Chauhan, P. (2020). Changes in turbidity along Ganga River using Sentinel-2 satellite data during lockdown associated with COVID-19. *Geomatics, Natural Hazards and Risk*, *11*(1), 1175–1195. <https://doi.org/10.1080/19475705.2020.1782482>
- Gholizadeh, M. H., & Melesse, A. M. (2017). Study on Spatiotemporal Variability of Water Quality Parameters in Florida Bay Using Remote Sensing. *Journal of Remote Sensing & GIS*, *06*(03). <https://doi.org/10.4172/2469-4134.1000207>
- Gholizadeh, M., Melesse, A., & Reddi, L. (2016). A Comprehensive Review on Water Quality Parameters Estimation Using Remote Sensing Techniques. *Sensors*, *16*(8), 1298. <https://doi.org/10.3390/s16081298>
- Ha, N. T. T., Thao, N. T. P., Koike, K., & Nhuan, M. T. (2017). Selecting the Best Band Ratio to Estimate Chlorophyll-a Concentration in a Tropical Freshwater Lake Using Sentinel 2A Images from a Case Study of Lake Ba Be (Northern Vietnam). *ISPRS International Journal of Geo-Information*, *6*(9), 290. <https://doi.org/10.3390/ijgi6090290>
- Hafeez, S., Sing Wong, M., Abbas, S., Y. T., Kwok, C., Nichol, J., Ho Lee, K., Tang, D., & Pun, L. (2019). Detection and Monitoring of Marine Pollution Using Remote Sensing Technologies. In *Monitoring of Marine Pollution*. IntechOpen. <https://doi.org/10.5772/intechopen.81657>
- Harrington, J. A., Schiebe, F. R., & Nix, J. F. (1992). Remote sensing of Lake Chicot, Arkansas: Monitoring suspended sediments, turbidity, and Secchi depth with Landsat MSS data. *Remote Sensing of Environment*, *39*(1), 15–27. [https://doi.org/10.1016/0034-4257\(92\)90137-9](https://doi.org/10.1016/0034-4257(92)90137-9)
- He, Y., Jin, S., & Shang, W. (2021). Water Quality Variability and Related Factors along the Yangtze River Using Landsat-8. *Remote Sensing*, *13*(12), 2241. <https://doi.org/10.3390/rs13122241>
- Hossain, A. K. M. A., Mathias, C., & Blanton, R. (2021). Remote Sensing of Turbidity in the Tennessee River Using Landsat 8 Satellite. *Remote Sensing*, *13*(18), 3785. <https://doi.org/10.3390/rs13183785>
- Huang J, Guo H, Guo X, & Singh V.P. (2020). Retrieval of Non-Optically Active Parameters for Small Scale Urban Waterbodies by a Machine Learning-Based Strategy. *Preprints*. <https://doi.org/10.20944/preprints202004.0111.v1>
- IBM. (2020, December 7). *Random forest*. IBM. <https://www.ibm.com/cloud/learn/random-forest>
- Kalele, A. (2019). *ESTIMATION AND MAPPING OF TURBIDITY IN THE LOWER CHARLES RIVER USING LANDSAT 8 OLI SATELLITE IMAGERY*. Northeastern University Boston, Massachusetts.
- Kapalanga, T. S., Hoko, Z., Gumindoga, W., & Chikwiramakomo, L. (2021). Remote-sensing-based algorithms for water quality monitoring in Olushandja Dam, north-central Namibia. *Water Supply*, *21*(5), 1878–1894. <https://doi.org/10.2166/ws.2020.290>
- Karch, J. (2020). Improving on adjusted R-squared. *Collabra: Psychology*, *6*(1). <https://doi.org/10.1525/collabra.343>

- Katlane, R., Dupouy, C., Kilani, B. el, & Berges, J. C. (2020). Estimation of Chlorophyll and Turbidity Using Sentinel 2A and EO1 Data in Kneiss Archipelago Gulf of Gabes, Tunisia. *International Journal of Geosciences*, 11(10), 708–728. <https://doi.org/10.4236/ijg.2020.1110035>
- Li, X., Ding, J., & Ilyas, N. (2021). Machine learning method for quick identification of water quality index (WQI) based on Sentinel-2 MSI data: Ebinur Lake case study. *Water Science and Technology: Water Supply*, 21(3). <https://doi.org/10.2166/ws.2020.381>
- Lim, J., & Choi, M. (2015). Assessment of water quality based on Landsat 8 operational land imager associated with human activities in Korea. *Environmental Monitoring and Assessment*, 187(6), 384. <https://doi.org/10.1007/s10661-015-4616-1>
- Moore, G. K. (1980). Satellite remote sensing of water turbidity. *Hydrological Sciences Bulletin*, 25(4). <https://doi.org/10.1080/02626668009491950>
- Nas, B., Ekercin, S., Karabörk, H., Berktay, A., & Mulla, D. J. (2010). An Application of Landsat-5TM Image Data for Water Quality Mapping in Lake Beysehir, Turkey. *Water, Air, & Soil Pollution*, 212(1–4), 183–197. <https://doi.org/10.1007/s11270-010-0331-2>
- Pavelsky, T. M., & Smith, L. C. (2009). Remote sensing of suspended sediment concentration, flow velocity, and lake recharge in the Peace-Athabasca Delta, Canada. *Water Resources Research*, 45(11). <https://doi.org/10.1029/2008WR007424>
- Quang, N., Sasaki, J., Higa, H., & Huan, N. (2017). Spatiotemporal Variation of Turbidity Based on Landsat 8 OLI in Cam Ranh Bay and Thuy Trieu Lagoon, Vietnam. *Water*, 9(8), 570. <https://doi.org/10.3390/w9080570>
- Ritchie, J. C., Zimba, P. V., & Everitt, J. H. (2003). Remote Sensing Techniques to Assess Water Quality. *Photogrammetric Engineering & Remote Sensing*, 69(6), 695–704. <https://doi.org/10.14358/PERS.69.6.695>
- Rubin, H. J., Lutz, D. A., Steele, B. G., Cottingham, K. L., Weathers, K. C., Ducey, M. J., Palace, M., Johnson, K. M., & Chipman, J. W. (2021). Remote Sensing of Lake Water Clarity: Performance and Transferability of Both Historical Algorithms and Machine Learning. *Remote Sensing*, 13(8), 1434. <https://doi.org/10.3390/rs13081434>
- Sebastiá-Frasquet, M.-T., Aguilar-Maldonado, J. A., Santamaría-Del-Ángel, E., & Estornell, J. (2019). Sentinel 2 Analysis of Turbidity Patterns in a Coastal Lagoon. *Remote Sensing*, 11(24), 2926. <https://doi.org/10.3390/rs11242926>
- Topp, S. N., Pavelsky, T. M., Jensen, D., Simard, M., & Ross, M. R. V. (2020). Research trends in the use of remote sensing for inland water quality science: Moving towards multidisciplinary applications. In *Water (Switzerland)* (Vol. 12, Issue 1). <https://doi.org/10.3390/w12010169>
- Verzani, J. (2004). Using R for Introductory Statistics. In *Using R for Introductory Statistics*. New York: Chapman and Hall/CRC. <https://doi.org/10.4324/9780203499894>
- Wang, L., Xu, M., Liu, Y., Liu, H., Beck, R., Reif, M., Emery, E., Young, J., & Wu, Q. (2020). Mapping Freshwater Chlorophyll-a Concentrations at a Regional Scale Integrating Multi-Sensor Satellite Observations with Google Earth Engine. *Remote Sensing*, 12(20), 3278. <https://doi.org/10.3390/rs12203278>
- Wass, P. D., Marks, S. D., Finch, J. W., Leeks, G. J. L., & Ingram, J. K. (1997). Monitoring and preliminary interpretation of in-river turbidity and remotely sensed imagery for suspended sediment transport studies in the Humber catchment. *Science of The Total Environment*, 194–195, 263–283. [https://doi.org/10.1016/S0048-9697\(96\)05370-3](https://doi.org/10.1016/S0048-9697(96)05370-3)
- Zhang, S., & Gao, H. (2020). Using the digital elevation model (DEM) to improve the spatial coverage of the MODIS-based reservoir monitoring network in South Asia. *Remote Sensing*, 12(5). <https://doi.org/10.3390/rs12050745>