



Research Article

OPEN ACCESS

Kharazmi Journal of Earth Sciences

Journal homepage <https://gnf.khu.ac.ir>

## Detection and classification of environmental anomalies using multi-temporal sentinel-2 satellite imagery and lightweight machine learning

Mohammad Javad Shojaei<sup>1</sup>, Asghar Milan<sup>2\*</sup>

1, 2. Faculty of Civil, Water and Environmental Engineering, Shahid Beheshti University, Tehran, Iran

### Article info

#### Article history

Received: 6 November 2025

Accepted: 26 December 2025

#### Keywords:

Environmental anomalies, spectral indices, machine learning, multi-temporal satellite imagery, lightweight algorithms.



### Abstract

Monitoring vegetation cover changes and environmental anomalies is essential for ecosystem management, precision agriculture, and early warning systems. Given the complexity of spatiotemporal environmental patterns, multi-temporal satellite data offer an efficient approach to track gradual and abrupt changes. In this study, time-series data from the Sentinel-2 sensor and spectral indices NDVI, EVI, and NBR are used as indicators sensitive to chlorophyll content and severe disturbances, along with lightweight, unsupervised algorithms Isolation Forest, Local Outlier Factor (LOF), and One-Class Support Vector Machine to identify anomalies. Model performance evaluation shows that the Isolation Forest algorithm provides the most balanced and robust performance (Accuracy = 0.886, Precision = 0.065, Recall = 0.250, F1 = 0.103). The Local Outlier Factor algorithm demonstrated higher sensitivity to localized patterns but had lower stability in noisy data (F1 = 0.069). The One-Class SVM adopted a more conservative labeling approach and was mostly effective at detecting severe disturbances, especially with the NBR index (F1 = 0.035). Overlap and Distinction Analysis of Indices that NDVI captures gradual chlorophyll decline, EVI performs better in dense vegetation, and NBR plays a crucial role in identifying severe events such as wildfires and droughts. The concurrent use of these indices broadens the detectable range from subtle fluctuations to large-scale disturbances. Findings show that integrating freely available Sentinel-2 data with lightweight machine learning models yields a scalable, reproducible, and efficient framework for large-scale environmental anomaly monitoring. This approach minimizes reliance on costly field data and enables practical applications in drought monitoring, environmental-crisis management, ecosystem health assessment, and smart agricultural planning. For future work, integrating moisture- and temperature-related spectral data, higher temporal resolution time series, and metaheuristic parameter tuning is recommended to further enhance robustness and accuracy.

### Introduction

In recent years, remote sensing and satellite image processing have become key tools in Earth sciences, with wide applications in precision agriculture, natural resource management, environmental monitoring, land-use change analysis, and climate assessment (Tempa et al., 2024). Owing to their broad spatial coverage, suitable temporal revisit, and acceptable spatial resolution, multi-temporal satellite images enable continuous monitoring and timely updating of environmental information (Abdalla Mahmoud et al., 2025). In agricultural studies, remote sensing techniques have demonstrated strong capabilities in crop growth monitoring, vegetation health

assessment, and anomaly detection (Hosseini et al., 2022). One of the major challenges in remote sensing data analysis is the identification of anomalies or abnormal patterns caused by factors such as pests, diseases, nutrient deficiencies, and sudden environmental disturbances. Timely detection of these anomalies plays a critical role in improving agricultural management and productivity (Zhao et al., 2022). In this context, machine learning methods have been increasingly adopted as automated and effective solutions for anomaly detection in satellite imagery (Schiller et al., 2024). Lightweight and unsupervised algorithms, including Isolation Forest, Local Outlier Factor, and One-Class Support Vector

DOI <http://doi.org/10.22034/KJES.2025.11.2.107623>

\*Corresponding author: Asghar Milan; E-mail: [a\\_milan@sbu.ac.ir](mailto:a_milan@sbu.ac.ir)

How to cite this article: Shojaei, M. J., Milan, A., 2025. Detection and classification of environmental anomalies using multi-temporal sentinel-2 satellite imagery and lightweight machine learning. Kharazmi Journal of Earth Sciences 11(2), 342- 365. <http://doi.org/10.22034/KJES.2025.11.2.107623>



BY NC

Machine, are particularly suitable for large-scale environmental analysis due to their low computational cost and minimal reliance on labeled data (Khan et al., 2023).

Recent advances in satellite missions have led to the availability of large volumes of high-resolution multi-temporal data. Sentinel-2 imagery from the European Space Agency, with spatial resolutions ranging from 10 to 60 m and a five-day revisit cycle, has become one of the most important data sources for vegetation monitoring (Ju et al., 2025; Descals et al., 2025). Compared to earlier optical missions such as Landsat, Sentinel-2 offers enhanced spectral and spatial capabilities, enabling more accurate detection of vegetation dynamics, soil conditions, and surface changes. Among spectral indices, the Normalized Difference Vegetation Index (NDVI) remains the most widely used indicator for vegetation condition assessment, while other indices such as EVI and NBR provide complementary information for dense vegetation and severe disturbances.

Despite existing studies on vegetation anomaly detection in Iran, most have focused on a single spectral index and lack a systematic comparison of machine learning algorithms. Therefore, there is a clear need for a comprehensive framework that integrates multiple spectral indices and evaluates the performance of

different unsupervised learning approaches. In this study, the Moghan Plain one of Iran's major agricultural regions, is selected as the study area. The main objective is to compare the performance of Isolation Forest, Local Outlier Factor, and One-Class Support Vector Machine for unsupervised detection of vegetation anomalies using multi-temporal Sentinel-2 data, contributing to the development of scalable and reproducible monitoring systems for precision agriculture.

## Materials and Methods

The Moghan Plain is one of Iran's major agricultural regions, located in the northwest of the country within Ardabil Province, along the border with the Republic of Azerbaijan. Bounded by the Aras River to the north and the Sabalan Mountains to the west, the plain covers approximately 2,500–3,000 km<sup>2</sup> and is characterized by a cold semi-arid climate. Annual precipitation averages 300–350 mm, with warm, dry summers and cold winters. Fertile soils, extensive irrigation infrastructure, and favorable agro-climatic conditions have made the region a key production zone for strategic crops such as wheat, barley, maize, and oilseeds. Due to its sensitivity to climatic variability, drought, pests, and water stress, the Moghan Plain represents a suitable case study for satellite-based vegetation monitoring.

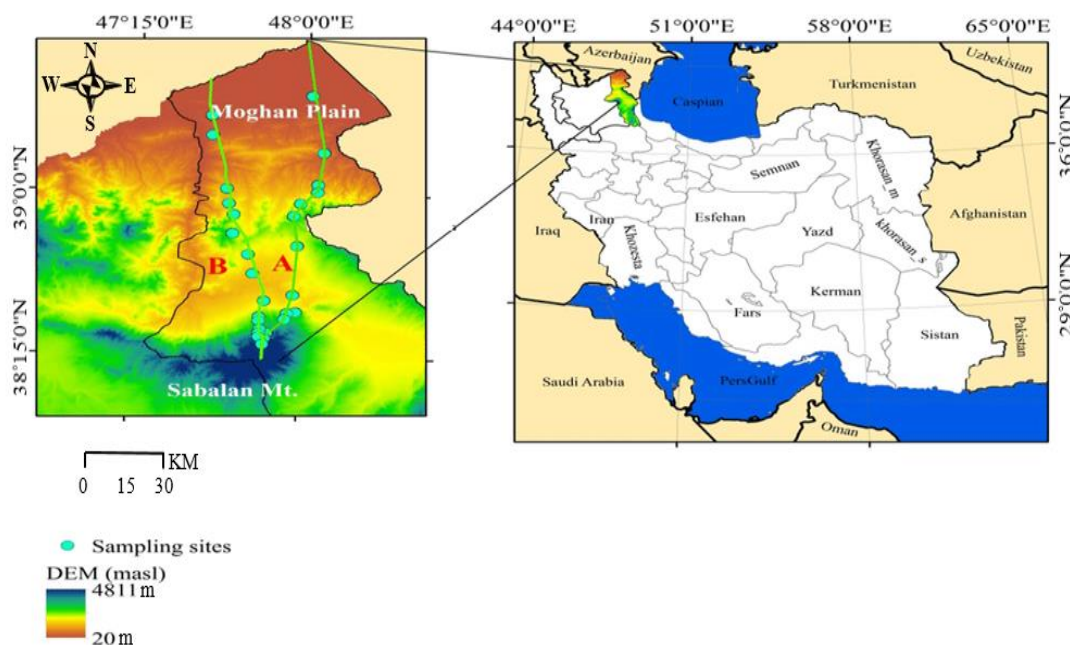


Fig. 1. Overview map of the study area (Moghan Plain).

This study utilizes multi-temporal Sentinel-2 Level-2A imagery (2020–2022) with a final spatial resolution of 10 m. After cloud masking, mosaicking, and band selection, three spectral indices NDVI, EVI, and NBR—were computed to capture complementary aspects of vegetation condition and disturbance. The SWIR2 band was resampled to ensure spatial consistency prior to index calculation. Unsupervised machine learning algorithms, including Isolation Forest, Local Outlier Factor, and One-Class Support Vector Machine, were implemented in Python to detect vegetation anomalies without labeled training data. A reference dataset was generated by integrating high-resolution imagery, local reports, and visual interpretation of index time series. Model performance was evaluated using Accuracy, Precision, Recall, and F1-score, with particular emphasis on F1 due to class imbalance in anomaly detection tasks.

## Results and Discussion

Temporal analysis of NDVI, EVI, and NBR series revealed distinct seasonal and interannual variations across the Moghan Plain. NDVI exhibited a clear cyclical

pattern corresponding to the agricultural growing season, peaking between May and August. EVI provided better discrimination in areas with dense vegetation cover, while NBR effectively captured extreme drought and burnt patches.

The three algorithms yielded varying sensitivity to spatial and temporal anomalies (Table 1). The Isolation Forest algorithm outperformed the others, achieving an F1-score of 0.103, with Precision = 0.065 and Recall = 0.250. It effectively detected both moderate and severe vegetation disturbances while maintaining robustness against noise. LOF was more responsive to local deviations but showed instability in noisy spectral regions, leading to higher false positives. OC-SVM displayed a conservative behavior with lower recall but fewer false alarms—making it suitable for applications where Type I errors are costly. A statistical comparison of the performance of the three algorithms is presented in Figure 2.

Table 1. Performance results of three algorithms

Algorithm	Accuracy	Precision	Recall	F1
Isolation Forest	0.886	0.065	0.250	0.103
One-Class SVM	0.879	0.022	0.083	0.035
Local Outlier Factor	0.882	0.043	0.167	0.069

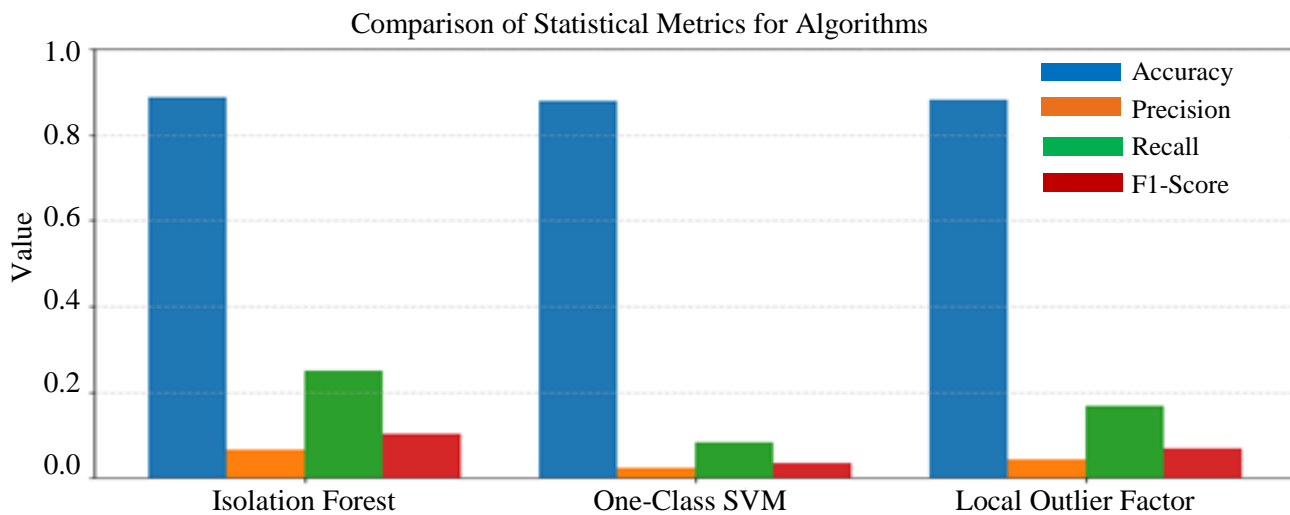


Fig. 2. Comparison chart of three algorithms

These results align with previous studies emphasizing the robustness of Isolation Forest in NDVI-based anomaly detection (Zhao et al., 2022; Khan et al., 2023). Moreover, the joint use of NDVI, EVI, and NBR improved sensitivity across vegetation conditions, reducing false positives caused by phenological variations.

### Conclusions

This study proposed a data-driven framework for environmental anomaly detection using multi-temporal Sentinel-2 imagery and unsupervised machine learning algorithms. The results demonstrate that integrating NDVI, EVI, and NBR with the Isolation Forest algorithm

provides a robust capability for detecting vegetation anomalies across different temporal scales. Compared to Local Outlier Factor and One-Class Support Vector Machine, Isolation Forest achieved a better balance between anomaly detection sensitivity and false alarm control, making it well suited for environmental monitoring, agricultural management, and early warning applications.

The use of freely available Sentinel-2 data and open-source, cloud-based processing environments enhances the reproducibility and scalability of the proposed framework. Although absolute values of performance metrics such as F1-score were relatively low due to



severe class imbalance and the inherent nature of unsupervised anomaly detection, the relative comparison of algorithms remains reliable. Model outputs were generated at the pixel level, enabling spatial representation of anomalies, while the primary focus of the study was on temporal analysis. The main limitation of this research is the lack of extensive ground-truth data, which can be addressed in future studies through the integration of field observations and multi-source reference data.

Overall, the findings indicate that combining multiple vegetation indices with lightweight machine learning models provides an efficient, reproducible, and cost-effective approach for environmental anomaly detection, particularly in regions with limited financial and infrastructural resources. Future research should explore the inclusion of additional spectral indices (e.g., moisture and thermal indicators), multi-source data fusion (e.g., Sentinel-1 and Sentinel-2), deep learning architectures for spatiotemporal modeling, and metaheuristic optimization methods for automatic parameter tuning, in order to further enhance the robustness and performance of the proposed framework.

## References

- Abdalla Mahmoud, A.M., Rayne, L., Bewley, R., 2025. EAMENA MLACD: Automated remote sensing tool based on machine learning in Google Earth Engine using Sentinel-2 imagery for detecting changes and threats around archaeological sites. *Remote Sensing Applications: Society and Environment* 37, 101396.
- Descals, A., Torres, K., Verger, A., Peñuelas, J., 2025. Evaluating Sentinel-2 for monitoring drought-induced crop failure in winter cereals. *Remote Sensing* 17 (2), 340.
- Hosseini, A., Karami, M., Moghaddam, R., 2022. Assessment of drought impacts on agricultural lands using Sentinel-2 NDVI in Nishabour plain, Iran. *Iranian Journal of Remote Sensing and GIS* 14 (2), 45–60.
- Ju, J., Zhou, Q., Freitag, B., Roy, D.P., Zhang, H.K., Sridhar, M., Mandel, J., Arab, S., Schmidt, G., Crawford, C.J., Gascon, F., Strobl, P.A., Masek, J.G., Neigh, C.S.R., 2025. Harmonized Landsat and Sentinel-2 surface reflectance data (HLS V2.0): Consistent multispectral observations for large-scale terrestrial monitoring. *Remote Sensing of Environment* 324, 114723.
- Khan, M., Singh, R., Ahmad, I., 2023. Detection of unusual vegetation changes using local outlier factor in Punjab, India. *Environmental Monitoring and Assessment* 195 (7), 849.
- Schiller, C., Költzow, J., Schwarz, S., Schiefer, F., Fassnacht, F.E., 2024. Transformer-based Forest disturbance detection using Sentinel-2 time series. *Remote Sensing of Environment* 315, 114475.
- Tempa, K., Ilunga, M., Agarwal, A., Tashi, P., 2024. Land use/land cover change dynamics and NDVI analysis using Sentinel-2 in Gelephu, Bhutan. *Applied Sciences* 14 (4), 1578.
- Zhao, X., Wang, Y., Liu, Y., 2022. Vegetation anomaly detection using NDVI time series and isolation forest algorithm. *Journal of Applied Remote Sensing* 16 (3), 036510.

## CRediT authorship contribution statement

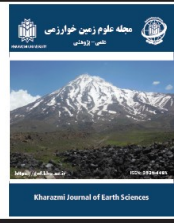
 <p><b>Mohammad Javad Shojaei</b></p>	<p>Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data Curation, Writing -Original Draft, Writing - Review &amp; Editing, Supervision, Project administration</p>
 <p><b>Asghar Milan</b></p>	<p>Writing - Original Draft, Writing -Review &amp; Editing, Supervision, Project administration</p>



مقاله پژوهشی

دسترسی آزاد

مجله علوم زمین خوارزمی

Journal homepage <https://gnf.khu.ac.ir>

## تشخیص و طبقه‌بندی ناهنجاری‌های محیطی با استفاده از تصاویر چندزمانه ماهواره‌ای سنتینل-۲ و یادگیری ماشین سبک

محمد جواد شجاعی<sup>۱</sup>، اصغر میلان<sup>۲\*</sup>

۱. دانشکده مهندسی عمران، آب و محیط زیست دانشگاه شهید بهشتی، تهران، ایران

اطلاعات مقاله	چکیده
<b>تاریخچه مقاله</b> دریافت: ۱۴۰۴/۰۸/۱۵ پذیرش: ۱۴۰۴/۱۰/۰۵	پایش تغییرات پوشش گیاهی و ناهنجاری‌های محیطی برای مدیریت اکوسیستم‌ها، کشاورزی دقیق و سامانه‌های هشدار زودهنگام ضروری است. با توجه به پیچیدگی الگوهای زمانی-مکانی تغییرات محیطی، داده‌های ماهواره‌ای چندزمانه رویکردی کارآمد برای ردیابی تغییرات تدریجی و ناگهانی ارائه می‌دهند. در این پژوهش، از داده‌های سری‌زمانی سنجنده سنتینل-۲ و شاخص‌های طیفی NDVI، EVI و NBR به‌عنوان شاخص‌های حساس به کلروفیل و اختلالات شدید و الگوریتم‌های سبک و بدون ناظر جنگل جداساز، عامل ناهنجاری محلی و ماشین بردار پشتیبان تک‌کلاسه، برای شناسایی ناهنجاری‌ها بهره‌گرفته شده است. ارزیابی عملکرد مدل‌ها نشان داد که الگوریتم جنگل جداساز، متعادل‌ترین و پایدارترین کارایی را ارائه می‌دهد (Accuracy = 0.886, Precision = 0.065, Recall = 0.250, F1 = 0.103). الگوریتم عامل ناهنجار محلی در شناسایی الگوهای موضعی حساسیت بیشتری نشان داد، اما در داده‌های دارای نویز ثبات پایین‌تری داشت (F1 = 0.069). الگوریتم ماشین بردار تک‌کلاسه رویکرد محافظه‌کارانه‌تری در برچسب‌گذاری ناهنجاری‌ها اتخاذ کرد و عمدتاً در تشخیص اختلالات شدید، به‌ویژه در شاخص NBR، مؤثرتر بود (F1 = 0.035). تحلیل هم‌پوشانی و تمایز شاخص‌ها نشان می‌دهد NDVI کاهش تدریجی کلروفیل را ثبت می‌کند، EVI در پوشش گیاهی متراکم و NBR در شناسایی رخدادهای شدید نظیر آتش‌سوزی و خشکسالی نقش مکمل ایفا می‌کنند. استفاده هم‌زمان از این شاخص‌ها دامنه قابل تشخیص از نوسانات ظریف تا اختلالات گسترده را پوشش می‌دهد. یافته‌ها نشان می‌دهد که تلفیق داده‌های سنتینل-۲ با مدل‌های سبک یادگیری ماشین، چارچوبی قابل‌توسعه، بازتولیدپذیر و کارا برای پایش ناهنجاری‌های محیطی در مقیاس‌های وسیع فراهم می‌سازد. این رویکرد وابستگی به داده‌های میدانی پرهزینه را به حداقل رسانده و امکان استفاده عملی در پایش خشکسالی، مدیریت بحران‌های محیطی، پایش سلامت اکوسیستم‌ها و برنامه‌ریزی کشاورزی هوشمند را فراهم می‌کند. برای تحقیقات آینده، تلفیق داده‌های طیفی مرتبط با رطوبت و حرارت، سری‌های زمانی با تفکیک زمانی بالاتر و به‌کارگیری تنظیم پارامترهای فراابتکاری پیشنهاد می‌شود تا پایداری و دقت بیشتر حاصل گردد.
<b>واژه‌های کلیدی</b> ناهنجاری‌های محیطی، شاخص‌های طیفی، یادگیری ماشین، تصاویر ماهواره‌ای چندزمانه، الگوریتم‌های سبک.	



### مقدمه

کشاورزی دقیق، مدیریت منابع طبیعی، نظارت بر محیط‌زیست، مطالعه تغییرات کاربری اراضی و ارزیابی شرایط اقلیمی مورد توجه قرار گرفته است. صفا و همکاران (Safa et al., 2023) و عبدالله محمود و همکاران (Abdalla Mahmoud et al., 2025) بیان کردند که تصاویر ماهواره‌ای به دلیل گستره پوشش وسیع، تکرار زمانی مناسب و وضوح

در سال‌های اخیر، تمپا و همکاران (Tempa et al., 2024) و فرانکلین و همکاران (Farahnakian et al., 2024) نشان دادند که سنجش از دور و پردازش تصاویر ماهواره‌ای به عنوان یکی از ابزارهای کلیدی در علوم زمین، به‌طور فزاینده‌ای در زمینه‌هایی همچون

DOI <http://doi.org/10.22034/KJES.2025.11.2.107623>

\*نویسنده مسئول: اصغر میلان a\_milan@sbu.ac.ir

استناد به این مقاله: شجاعی، م. ج.، میلان، ا. (۱۴۰۴). تشخیص و طبقه‌بندی ناهنجاری‌های محیطی با استفاده از تصاویر چندزمانه ماهواره‌ای سنتینل-۲ و یادگیری ماشین

سبک. مجله علوم زمین خوارزمی. جلد ۱۱، شماره ۲، صفحه ۳۴۲ تا ۳۶۵. <http://doi.org/10.22034/KJES.2025.11.2.107623>

پدیده‌هایی هستند که از الگوهای طبیعی فاصله می‌گیرند و می‌توانند ناشی از عوامل طبیعی مانند خشکسالی، سیلاب و آتش‌سوزی یا عوامل انسانی نظیر تغییر کاربری اراضی، جنگل‌زدایی و توسعه شهری باشند. تشخیص به‌موقع این ناهنجاری‌ها برای مدیریت منابع طبیعی اهمیت بالایی دارد. پیشرفت‌های اخیر در فناوری سنجنش از دور منجر به تولید حجم عظیمی از داده‌های چندزمانه با وضوح بالا شده است. داده‌های سنتینل-۲ آژانس فضایی اروپا (ESA) با وضوح مکانی ۱۰ تا ۶۰ متر و بازه بازدید پنج‌روزه، یکی از مهم‌ترین منابع برای پیش‌بینی تغییرات پوشش گیاهی به‌شمار می‌روند. برخلاف داده‌های اپتیک قدیمی‌تر نظیر لندست، داده‌های سنتینل-۲ به دلیل ترکیب باندهای طیفی متنوع (مرئی، مادون قرمز نزدیک و مادون قرمز کوتاه) و رزولوشن بالاتر، قابلیت تشخیص دقیق‌تری از پوشش گیاهی، منابع آبی و تغییرات سطح خاک دارند (Ju et al., 2025). این ویژگی‌ها، سنتینل-۲ را به ابزاری توانمند برای مطالعات زیست‌محیطی و آشکارسازی ناهنجاری‌ها تبدیل کرده است. از میان شاخص‌های طیفی، شاخص تفاضلی نرمال‌شده پوشش گیاهی (NDVI) پرکاربردترین شاخص برای تحلیل وضعیت پوشش گیاهی است. ژائو و همکاران (Zhao et al., 2022) بیان کردند که این شاخص با استفاده از تفاوت بازتاب طیفی در باندهای قرمز و مادون قرمز نزدیک محاسبه می‌شود و به‌طور گسترده در پایش بیوماس، خشکسالی، رشد گیاه، آتش‌سوزی جنگل‌ها و بهره‌وری اکولوژیکی مورد استفاده قرار می‌گیرد.

در سال‌های اخیر، سری زمانی NDVI برای تشخیص تغییرات غیرمعمول توسعه یافته و روش‌هایی نظیر میانگین متحرک، تحلیل روند، مدل‌های ARIMA و روش‌های مبتنی بر آزمون‌های آماری برای تحلیل این سری‌های زمانی به‌کار رفته‌اند. با این حال، با افزایش پیچیدگی داده‌های محیطی، این روش‌های سنتی قادر به تشخیص ناهنجاری‌های پیچیده یا غیرخطی نیستند؛ از این رو الگوریتم‌های یادگیری ماشین، به‌ویژه روش‌های بدون ناظر همچون جنگل جداساز، عامل ناهنجاری محلی و ماشین بردار پشتیبان تک‌کلاسه، به‌طور روزافزون در مطالعات مرتبط استفاده می‌شوند. این الگوریتم‌ها بدون

مکانی قابل قبول امکان پیش‌بینی مستمر و به‌روزرسانی مداوم اطلاعات محیطی را فراهم می‌کنند. در زمینه کشاورزی، واسهون و همکاران (Wasehun et al., 2025) گزارش کردند که فناوری‌های سنجنش از دور قابلیت‌های گسترده‌ای از پایش رشد محصولات تا ارزیابی سلامت گیاهان و شناسایی ناهنجاری‌ها دارند. همچنین بابایی‌پور و میلان (Babaepour and Milan, 2024) نشان دادند که با استفاده از داده‌های سنجنش از دور در سکوه‌های مختلف می‌توان نقشه‌های دقیق تغییرات کاربری اراضی تهیه کرد.

یکی از چالش‌های اساسی در تحلیل داده‌های سنجنش از دور، شناسایی و استخراج آنومالی‌ها یا داده‌های غیرطبیعی است که ممکن است ناشی از عواملی نظیر آفات، بیماری‌ها، کمبود عناصر غذایی یا اختلالات ناگهانی محیطی باشد. تشخیص دقیق و به‌موقع این ناهنجاری‌ها نقش مهمی در بهبود مدیریت کشاورزی و افزایش بهره‌وری محصولات دارد. در این راستا، پاچکو-پاسکاگازا و همکاران (Pacheco and Pascagaza et al., 2022) استفاده از الگوریتم‌های یادگیری ماشین را به عنوان روشی قدرتمند و خودکار برای شناسایی ناهنجاری‌ها در تصاویر ماهواره‌ای معرفی کرده‌اند. به گفته الشری و گوالی (Alshari and Gawali, 2022)، الگوریتم‌های یادگیری ماشین سبک به دلیل نیاز کمتر به داده‌های آموزش بزرگ و هزینه محاسباتی پایین، برای تحلیل داده‌های محیطی در مقیاس‌های وسیع بسیار مناسب هستند. علاوه بر این، مطالعات خان و همکاران (Khan et al., 2023) نشان داد که الگوریتم‌هایی مانند جنگل جداساز، عامل ناهنجاری محلی و ماشین بردار پشتیبان تک‌کلاسه به دلیل سادگی، سرعت و دقت قابل قبول در داده‌های بدون برچسب به‌طور گسترده برای تشخیص آنومالی‌ها به‌کار گرفته می‌شوند. این الگوریتم‌ها با تحلیل ویژگی‌های طیفی و مکانی تصاویر ماهواره‌ای قادرند مناطق مشکوک را به صورت خودکار شناسایی کنند، که این امر می‌تواند در کاهش هزینه‌ها و زمان نظارت‌های میدانی مؤثر باشد. همچنین تحلیل و شناسایی ناهنجاری‌های محیطی یکی از موضوعات مهم در علوم محیطی، سنجنش از دور و تحلیل داده‌های مکانی است. ناهنجاری‌ها شامل

شناسایی ناهنجاری‌های پوشش گیاهی را به صورت مؤثر و مقیاس‌پذیر فراهم سازد. در ایران نیز، به گفته حسینی و همکاران (Hosseini et al., 2022) اگرچه پژوهش‌هایی در زمینه شناسایی ناهنجاری‌های پوشش گیاهی انجام شده است، اغلب آن‌ها به یک شاخص خاص مانند NDVI محدود بوده و مقایسه جامع بین الگوریتم‌ها کمتر دیده می‌شود. پژوهش حاضر در پاسخ به این نیاز، با تمرکز بر دشت مغان به عنوان یکی از قطب‌های کشاورزی ایران، تلاش می‌کند این خلأ پژوهشی را پوشش داده و چارچوبی عملی برای پایش ناهنجاری‌های محیطی ارائه دهد. از اهداف اصلی مطالعه حاضر، مقایسه سه الگوریتم سبک یادگیری ماشین، جنگل جداساز، عامل ناهنجاری محلی و ماشین بردار پشتیبان تک‌کلاسه در تشخیص ناهنجاری‌های کشاورزی با استفاده از داده‌های چندزمانه سنتینل-۲ است. نتایج این بررسی می‌تواند به بهبود سامانه‌های نظارت کشاورزی و توسعه کاربردهای عملی سنجش از دور و یادگیری ماشین در کشاورزی دقیق کمک کند. همچنین در این پژوهش، هدف تشخیص ناهنجاری‌های پوشش گیاهی به صورت بدون ناظر با استفاده از تصاویر چندزمانه ماهواره‌ای سنتینل-۲ و الگوریتم‌های یادگیری ماشین بدون ناظر است.

## مواد و روش‌ها

### منطقه مورد مطالعه

دشت مغان یکی از پهنه‌های کشاورزی کلیدی ایران بوده و در شمال غرب کشور و در محدوده استان اردبیل واقع شده است. این دشت در همجواری مرز ایران و جمهوری آذربایجان قرار گرفته و از شمال به رودخانه ارس، از شرق به اراضی جمهوری آذربایجان، از غرب به دامنه‌های شرقی کوهستان سبلان و از جنوب به جلگه‌های منتهی به شهرستان پارس‌آباد محدود می‌شود. وسعت تقریبی این دشت در منابع مختلف بین ۲۵۰۰ تا ۳۰۰۰ کیلومتر مربع گزارش شده است و در محدوده‌ای با مختصات جغرافیایی حدود ۳۹ تا ۳۹/۷ عرض شمالی و ۴۷/۵ تا ۴۸/۳ طول شرقی جای می‌گیرد. این موقعیت جغرافیایی، دشت مغان را در یکی از حاصلخیزترین نواحی نیمه‌خشک ایران قرار

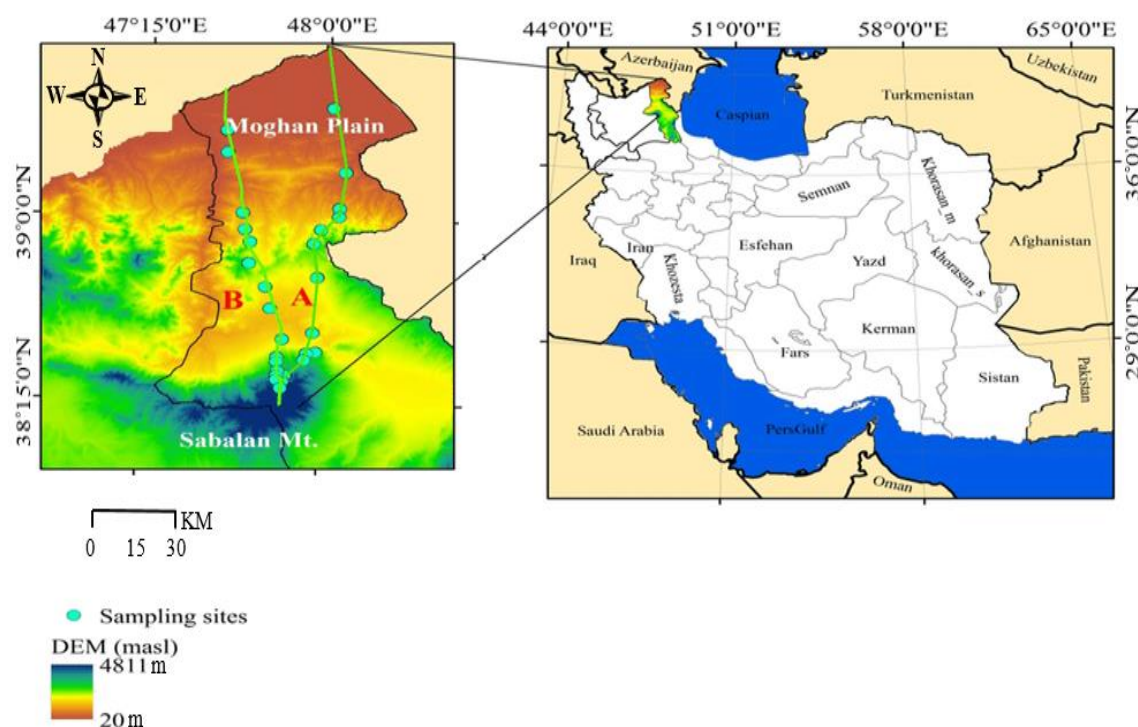
نیاز به داده‌های برچسب‌خورده، الگوهای غیرعادی را از داده‌های سری‌زمانی استخراج کرده و در مطالعات متعددی عملکرد قابل توجهی نشان داده‌اند (Kovárník and Janová, 2025).

در سال‌های ۲۰۲۴ و ۲۰۲۵، مطالعات جدیدی منتشر شده‌اند که این مسیر را توسعه داده‌اند. به عنوان نمونه، شیلر و همکاران (Schiller et al., 2024) از مدل‌های انتقال بر روی سری‌زمانی سنتینل-۲ برای تشخیص اختلالات جنگلی استفاده کردند. همچنین دسکالس و همکاران (Descals et al., 2025) کارایی سنتینل-۲ را در پایش شکست محصول ناشی از خشکسالی در غلات زمستانی بررسی کردند. لیانگ و همکاران (Liang et al., 2025) شاخص طیفی جدیدی تحت عنوان Forest Anomaly Comprehensive Index (FACI) معرفی کردند که قادر است چندین نوع ناهنجاری جنگلی از جمله آفات، آتش‌سوزی و تخریب را تشخیص دهد.

با وجود پیشرفت‌های قابل توجه در بهره‌گیری از داده‌های چندزمانه سنجش از دور و الگوریتم‌های یادگیری ماشین برای پایش تغییرات پوشش گیاهی، مرور مطالعات پیشین نشان می‌دهد که همچنان برخی خلأهای پژوهشی قابل توجه وجود دارد. بخش عمده‌ای از تحقیقات انجام‌شده، به استفاده از یک شاخص طیفی منفرد، به‌ویژه NDVI، محدود بوده و نقش شاخص‌های مکمل نظیر EVI و NBR در تشخیص انواع مختلف ناهنجاری‌ها کمتر به صورت هم‌زمان مورد بررسی قرار گرفته است. از سوی دیگر، بسیاری از مطالعات بر روش‌های خاص یا مدل‌های یادگیری ماشین منفرد تمرکز داشته و مقایسه نظام‌مند عملکرد الگوریتم‌های سبک و بدون ناظر در چارچوبی یکپارچه کمتر گزارش شده است، به‌ویژه در مطالعات مرتبط با مناطق کشاورزی ایران. همچنین، وابستگی برخی روش‌ها به داده‌های برچسب‌خورده یا مدل‌های محاسباتی پیچیده، کاربردپذیری آن‌ها را در مقیاس‌های وسیع و در شرایط محدودیت داده‌های میدانی کاهش می‌دهد. از این رو، نیاز به توسعه چارچوبی بازتولیدپذیر و کارا احساس می‌شود که با بهره‌گیری از داده‌های آزاد چندزمانه سنتینل-۲، ترکیب چند شاخص طیفی و مقایسه الگوریتم‌های سبک یادگیری ماشین بدون ناظر، امکان

از نظر ویژگی‌های خاک، بیشتر اراضی دشت دارای خاک‌های عمیق رسی-شنی با غنای بالای مواد آلی و ساختار فیزیکی مناسب برای زراعت مکانیزه هستند. همین ویژگی‌ها سبب شده که منطقه مغان به یکی از قطب‌های اصلی تولید محصولات استراتژیک کشور نظیر گندم، جو، ذرت، پنبه، دانه‌های روغنی و محصولات باغی تبدیل شود. پوشش گیاهی غالب منطقه در طول فصل‌های رشد شامل کشت‌های زراعی یک‌ساله، باغات، مزارع علوفه‌ای و اراضی آبی است، در حالی که در نواحی غیرزراعی، گونه‌های مرتعی و نیمه‌طبیعی مشاهده می‌شود. حساسیت اکولوژیکی دشت مغان نسبت به نوسانات اقلیمی، خشکسالی، کمبود منابع آب، آفات و بیماری‌های گیاهی، این پهنه را به منطقه‌ای مناسب برای مطالعات مبتنی بر سنجش از دور به‌ویژه پایش زمانی و مکانی تغییرات پوشش گیاهی و کاربری اراضی تبدیل کرده است. شکل ۱ نقشه منطقه مورد مطالعه را نشان می‌دهد.

داده است. اقلیم منطقه مغان بر اساس تقسیم‌بندی‌های اقلیمی ایران، در گروه نیمه‌خشک سرد طبقه‌بندی می‌شود و تحت تأثیر همزمان جریان‌های مرطوب سواحل دریای خزر و توده‌های هوایی سرد قفقاز قرار دارد (Encyclopaedia Iranica, n.d. Moḡān., 2015). میانگین بارش سالانه در این دشت حدود ۳۰۰ تا ۳۵۰ میلی‌متر و میانگین دمای سالانه در حدود ۱۳ تا ۱۵ درجه سانتی‌گراد گزارش شده که نشان‌دهنده تابستان‌های گرم و خشک و زمستان‌های سرد و نسبتاً مرطوب است. الگوی توزیع بارش نیز عمدتاً در فصل‌های پاییز و زمستان متمرکز بوده و بخش قابل توجهی از بارش‌های مؤثر در تغذیه سفره‌های آب زیرزمینی در این بازه زمانی رخ می‌دهد. علاوه بر این، وزش بادهای شدید در فصل بهار یکی از ویژگی‌های اقلیمی برجسته منطقه به‌شمار می‌آید که می‌تواند بر فرآیندهای فرسایش بادی و مدیریت کشاورزی تأثیرگذار باشد. منابع آب دشت مغان شامل رودخانه ارس، شبکه‌ای از رودخانه‌های فصلی و دائمی فرعی، و سامانه گسترده آبیاری منشعب از سد میل-مغان است که این زیرساخت‌ها امکان توسعه کشاورزی آبی در بخش‌های وسیعی از دشت را فراهم کرده‌اند (Solgi et al., 2018).



شکل ۱- نقشه منطقه مورد مطالعه (دشت مغان)

Fig. 1. Map of the study area (Moghan Plain)

۹۲ تصویر چندزمانه برای هر باند در تحلیل استفاده شده است. این اطلاعات به منظور شفاف‌سازی ویژگی‌های داده‌های ورودی و افزایش قابلیت بازتولیدپذیری پژوهش ارائه شده‌اند. از آنجا که باند مادون قرمز کوتاه (SWIR2) در تصاویر ماهواره‌ای سنتینل-۲ دارای قدرت تفکیک مکانی ۲۰ متر است، در حالی که سایر باندهای مورد استفاده در این پژوهش دارای تفکیک مکانی ۱۰ متر می‌باشند، به منظور یکسان‌سازی داده‌ها، باند SWIR2 پیش از محاسبه شاخص NBR به تفکیک مکانی ۱۰ متر بازنمونه‌برداری شد. این فرآیند در محیط گوگل ارث انجام و با استفاده از روش‌های استاندارد بازنمونه‌برداری انجام گرفت. در نهایت، تمامی شاخص‌های طیفی و تحلیل‌های بعدی بر اساس داده‌هایی با تفکیک مکانی نهایی ۱۰ متر انجام شدند.

## داده‌های مورد استفاده

در این پژوهش از تصاویر ماهواره‌ای سنتینل-۲ متعلق به آژانس فضایی اروپا (ESA) بهره‌برداری شد. این تصاویر با قدرت تفکیک مکانی بالا (۱۰ تا ۶۰ متر) و تنوع طیفی مناسب (۱۳ باند) از منابع قابل اتکا در پایش تغییرات پوشش سطح زمین و شناسایی ناهنجاری‌های زیست‌محیطی به‌شمار می‌آیند. باندهای مورد استفاده در جدول ۱ نشان داده شده است. در جدول ۱، علاوه بر نام باندها و طول موج مرکزی، اطلاعات مربوط به تفکیک مکانی هر باند، واحد طول موج و تعداد تصاویر مورد استفاده در تحلیل نیز ارائه شده است. تمامی باندهای انتخابی دارای قدرت تفکیک مکانی ۱۰ متر بوده و در بازه زمانی بهار ۲۰۲۰ تا پایان ۲۰۲۲، پس از اعمال فیلتر ابری، در مجموع

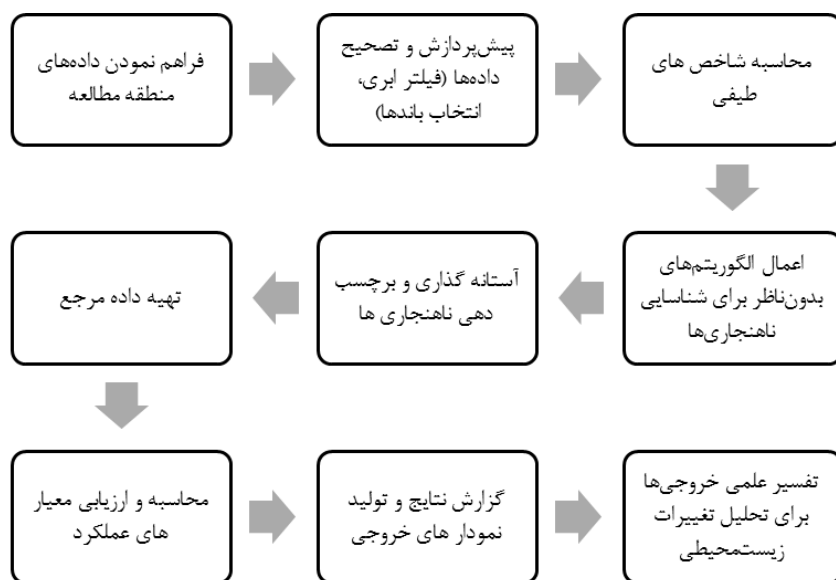
جدول ۱- باند های ماهواره سنتینل-۲

باند	نام باند	طول موج (نانومتر)	تفکیک مکانی (متر)	تعداد تصاویر	کاربرد
B2	آبی (Blue)	490(nm)	10	92	بررسی کیفیت آب و تصحیح اتمسفری
B4	قرمز (Red)	665(nm)	10	92	شناسایی تغییرات کلروفیل
B8	مادون قرمز نزدیک (NIR)	842(nm)	10	92	حساس به تراکم و سلامت پوشش گیاهی
B12	مادون قرمز کوتاه (SWIR2)	2190(nm)	10	92	حساس به رطوبت پوشش گیاهی و خاک

## روش تحقیق

گرفته شد. این بازه امکان مقایسه الگوهای فصلی و بین‌سالی پوشش گیاهی را فراهم کرده و برای آشکارسازی ناهنجاری‌های تدریجی و ناگهانی مناسب است. همچنین با توجه به اینکه مجموعه داده با تصحیح اتمسفری سطح-بازتاب (L2A) جلوی خطاهای ناشی از جو را می‌گیرد، این سری داده انتخاب گردید. بعد از فراهم نمودن داده‌های مورد نیاز فرایند پیش‌پردازش و تصحیح داده‌ها صورت گرفت. مراحل پیش‌پردازش شامل فیلتر ابری، موزاییک‌سازی و انتخاب باندهای مؤثر می‌باشد. در مرحله پیش‌پردازش داده‌ها برای حذف اثرات نویز و بهبود کیفیت داده‌ها، از فیلتر ابری استفاده شد که در این فیلتر از باند SCL برای حذف پیکسل‌های ابر، سایه ابر و برف بهره گرفته شد. سپس در مرحله موزاییک‌سازی زمانی اقدام به ترکیب تصاویر یک دوره زمانی به منظور تهیه تصویر یکپارچه از کامل منطقه گردید در ادامه انتخاب باندهای مؤثر صورت گرفت که برای این منظور باندهای قرمز (Red) و نزدیک فروسرخ (NIR) برای محاسبه شاخص‌های پوشش گیاهی انتخاب گردید.

در این پژوهش، از الگوریتم‌های یادگیری ماشین بدون ناظر برای تفکیک شرایط نرمال از ناهنجاری‌های محیطی استفاده شده است و هیچ داده برچسب‌خورده‌ای در مرحله آموزش مدل‌ها به کار نرفته است. مراحل انجام تحقیق صورت گرفته در شکل ۲ نمایش داده شده است که در ادامه مراحل صورت گرفته توضیح داده خواهد شد. مطابق شکل ۲ در گام اول اقدام به تهیه داده‌های مورد نیاز تحقیق از منبع گوگل ارث انجام گردید در انتخاب داده‌ها به منظور افزایش دقت تحلیل‌ها، داده‌ها به صورت چندزمانه از بازه‌ای مشخص گردآوری شدند تا امکان استخراج تغییرات زمانی فراهم شود. انتخاب زمان‌های تصویربرداری بر مبنای فصول کم‌ابر و دوره‌های بحرانی تغییرات محیطی صورت گرفت. با توجه به هدف پژوهش که تمرکز بر شناسایی ناهنجاری‌های محیطی کوتاه‌مدت تا میان‌مدت و تغییرات بین‌سالی پوشش گیاهی است، بازه زمانی بهار ۲۰۲۰ تا پایان سال ۲۰۲۲ در نظر



شکل ۲ - فلوجارت فرایند صورت گرفته در تحقیق

Fig. 2. Flowchart of research process

در مرحله بعد شاخص‌های طیفی منتخب محاسبه شدند که این شاخص‌ها عبارتند از شاخص NDVI، شاخص EVI و شاخص NBR جدول شماره ۲ اطلاعات این شاخص‌ها را نشان می‌دهد.

جدول ۲ - روابط شاخص‌های طیفی منتخب

Table 2. Equation for selected spectral indices

شاخص	فرمول	منبع
NDVI	$NDVI = \frac{NIR - RED}{NIR + RED}$	(Rouse et al., 1974)
EVI	$EVI = \frac{NIR - RED}{NIR + C_1 \times RED - C_2 \times BLUE + L} \times G$	(Huete et al., 2002)
NBR	$NBR = \frac{NIR - SWIR2}{NIR + SWIR2}$	(Key and Benson, 2006)

در روابط بالا:

NIR، بازتاب طیفی باند مادون قرمز نزدیک (Near Infrared) خواهد بود گیاهان سالم در این باند بازتاب بالایی دارند چون ساختار سلولی برگ‌ها نور NIR را به شدت بازتاب می‌دهد. RED، بازتاب طیفی باند قرمز (Red) می‌باشد گیاهان سالم در این باند بازتاب کمی دارند چون رنگدانه کلروفیل، نور قرمز را برای فتوسنتز جذب می‌کند. BLUE بازتاب طیفی باند آبی (Blue) را نشان می‌دهد که به دلیل حساسیت بالا به پراکندگی اتمسفری و جذب توسط ذرات معلق، عمدتاً در تصحیح اثرات جوی و محاسبه شاخص EVI به منظور بهبود تفکیک پوشش گیاهی در مناطق با تراکم بالا مورد استفاده قرار می‌گیرد. ضرایب به کاررفته در فرمول شاخص EVI شامل ضریب بهره‌برداری (G)، ضریب تصحیح پس‌زمینه خاک (L) و ضرایب تصحیح اثرات جوی  $C_1$  و  $C_2$  بر اساس مقادیر پیشنهادی در مطالعه مرجع هونت و همکاران (Huete et al., 2002) انتخاب شده‌اند. این مقادیر به طور گسترده در مطالعات سنجش از دور و پایش پوشش گیاهی مورد استفاده قرار گرفته و عملکرد پایدار آن‌ها در شرایط مختلف اقلیمی و پوشش گیاهی تأیید شده است. از این‌رو، در این پژوهش نیز مقادیر استاندارد  $G = 2.5$ ،  $L = 1$ ،  $C_1 = 6$  و  $C_2 = 7.5$  به کار گرفته شده‌اند. در این پژوهش، سه شاخص NDVI، EVI و NBR به عنوان ورودی‌های اصلی انتخاب شدند، زیرا هر یک حساسیت متفاوتی نسبت به انواع خاصی از ناهنجاری‌ها دارند؛ به گونه‌ای که NDVI کاهش کلی پوشش گیاهی را به خوبی نشان می‌دهد، EVI برای مناطق با پوشش متراکم عملکرد دقیق‌تری دارد و NBR نسبت به سوختگی و تغییرات شدید آلی حساس‌تر است. استفاده هم‌زمان از چند شاخص سبب افزایش حساسیت و کاهش مثبت‌های کاذب می‌شود و امکان پایش دقیق‌تر تغییرات محیطی را فراهم می‌آورد. پس از پایش پردازش داده‌ها و استخراج شاخص‌ها، سه الگوریتم یادگیری ماشین بدون ناظر شامل جنگل جداساز، عامل ناهنجاری محلی و ماشین بردار پشتیبان تک‌کلاسه در محیط پایتون پیاده‌سازی شدند. برای بهبود عملکرد،

پارامترهای اصلی هر الگوریتم با استفاده از روش آزمون و خطا و تحلیل نتایج اولیه تنظیم شدند تا دقت تفکیک ناهنجاری‌ها افزایش یابد. از آنجا که داده‌های میدانی گسترده در منطقه مورد مطالعه در دسترس نبود، داده مرجع با ترکیب چند منبع اطلاعاتی تولید شد. در این راستا، از تصاویر وضوح‌بالای ماهواره‌ای موجود در گوگل ارث برای بررسی تغییرات مکانی، گزارش‌های محلی منتشرشده توسط سازمان‌های مرتبط با کشاورزی و منابع طبیعی استان، و تفسیر بصری سری‌های زمانی شاخص‌های NDVI، EVI و NBR استفاده گردید. تفسیر بصری شامل بررسی الگوهای غیرعادی کاهش یا افزایش ناگهانی شاخص‌ها در بازه‌های زمانی متوالی و تطبیق آن‌ها با شواهد مکانی و زمانی موجود بود.

تفسیر رفتار شاخص‌های NDVI، EVI و NBR بر اساس ویژگی‌های فیزیکی-طیفی هر شاخص و مطالعات پیشین انجام شده است. مطابق گزارش‌های ارائه‌شده توسط هونت و همکاران (۲۰۰۲)، شاخص EVI در مناطق با پوشش گیاهی متراکم نسبت به NDVI دچار اشباع کمتری شده و عملکرد پایدارتری دارد. همچنین کی و بنسون (Key and Benson, 2006) نشان دادند که شاخص NBR حساسیت بالایی به تغییرات شدید ناشی از آتش‌سوزی و تنش‌های شدید گیاهی دارد. بنابراین، تحلیل‌های ارائه‌شده در این پژوهش مبتنی بر چارچوب نظری مطالعات پیشین بوده و با نتایج آن‌ها همخوانی دارد.

$$|NDVI_t - NDVI_{t-1}| = \Delta NDVI_t \quad (1)$$

$$|EVI_t - EVI_{t-1}| = \Delta EVI_t \quad (2)$$

$$|NBR_t - NBR_{t-1}| = \Delta NBR_t \quad (3)$$

سپس بر اساس شواهد کتابخانه‌ای و تحلیل تجربی، آستانه‌هایی برای هر شاخص تعریف شدند:

$$\text{if } \Delta NDVI_t > 0.2 \quad NDVI_{change_t} = 1 \quad (4)$$

$$\text{if } \Delta EVI_t > 0.15 \quad EVI_{change_t} = 1 \quad (5)$$

است؛ در این حالت هشدارها عمدتاً صحیح‌اند اما بخش قابل توجهی از ناهنجاری‌های واقعی از دست می‌روند.

این رویکرد برای کاربردهایی که هزینه هشدار کاذب بالاست مطلوب‌تر است. در مقابل، Recall بالا همراه با Precision پایین به معنای حساسیت بالای مدل است؛ یعنی بیشتر رخ داده‌های واقعی شناسایی می‌شوند اما هشدارهای کاذب زیادی نیز تولید می‌گردد، و این وضعیت برای کاربردهایی مانند پایش خطر آتش‌سوزی یا خشکسالی مناسب‌تر است. در صورتی که مقدار F1 بالا باشد، می‌توان نتیجه گرفت که بین Precision و Recall توازن مناسبی برقرار شده است. همچنین در مواردی که Accuracy بالا ولی F1 پایین باشد، این مسئله نشانه وجود نامتوازن‌ی شدید است؛ به عبارت دیگر، مدل عمدتاً در شناسایی شرایط نرمال موفق بوده اما در کشف ناهنجاری‌ها توان کافی نداشته است.

### پیاده سازی و ارائه نتایج

در این پژوهش از ترکیب دو بستر قدرتمند برای پردازش داده‌های سنجش‌ازدور و اجرای مدل‌های یادگیری ماشین استفاده شد. گوگل ارث انجین به‌عنوان یک پلتفرم پردازش ابری برای دریافت تصاویر سنتینل-۲، اعمال فیلترهای ابری و استخراج شاخص‌های طیفی NDVI، EVI و NBR به‌کار گرفته شد. پس از استخراج شاخص‌ها، الگوریتم‌های یادگیری ماشین بدون‌ناظر شامل جنگل‌جدا‌ساز، عامل ناهنجاری محلی و ماشین بردار پشتیبان تک‌کلاسه در محیط پایتون (نسخه ۳/۱۰) پیاده‌سازی شدند. کتابخانه‌های NumPy و Pandas برای پردازش داده‌ها، Matplotlib برای ترسیم نمودارها، scikit-learn برای پیاده‌سازی الگوریتم‌ها، و GeoPandas و Rasterio برای مدیریت داده‌های مکانی مورد استفاده قرار گرفتند. تمامی اسکریپت‌ها با رویکرد بازتولیدپذیری توسعه داده شدند و خروجی‌ها به‌صورت مستند ذخیره گردیدند. سه شاخص مورد استفاده هرکدام حساسیت متفاوتی به تغییرات محیطی نشان دادند. شاخص NDVI بیشترین کارایی را در شناسایی کاهش تراکم پوشش سبز داشت (شکل ۳)، در حالی که EVI

$$if \Delta NBR_t > 0.25 \quad NBR_{change_t} = 1 \quad (6)$$

برای کاهش اثر نویز و جلوگیری از برچسب‌گذاری تغییرات اتفاقی به‌عنوان ناهنجاری، یک قاعده ترکیبی تعریف شد که بر اساس آن تنها تغییرات هم‌زمان چند شاخص به‌عنوان تغییر واقعی محسوب شدند.

$$GT = \begin{cases} 1 & if (NDVI_{change_t} + EVI_{change_t} + NBR_{change_t}) \geq 2 \\ 0 & otherwise \end{cases} \quad (7)$$

در گام بعد، خروجی هر الگوریتم با داده مرجع ترکیبی مقایسه شد و برای هر شاخص و هر الگوریتم، یک ماتریس اغتشاش تشکیل گردید. در این ماتریس، TP به نمونه‌هایی اشاره دارد که واقعاً ناهنجر بوده و مدل نیز آن‌ها را درست تشخیص داده است؛ FP نمونه‌های نرمالی هستند که مدل به اشتباه به‌عنوان ناهنجاری شناسایی کرده است؛ FN نمونه‌های ناهنجر واقعی‌اند که مدل موفق به شناسایی آن‌ها نشده است؛ و TN نمونه‌های نرمالی‌اند که مدل به درستی آن‌ها را عادی شناسایی کرده است. بر اساس این مقادیر، چهار معیار اصلی دقت محاسبه شدند. روابط این معیارها به ترتیب زیر در متن درج می‌شوند:

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (8)$$

$$Precision = \frac{TP}{TP+FP+FN+TN} \quad (9)$$

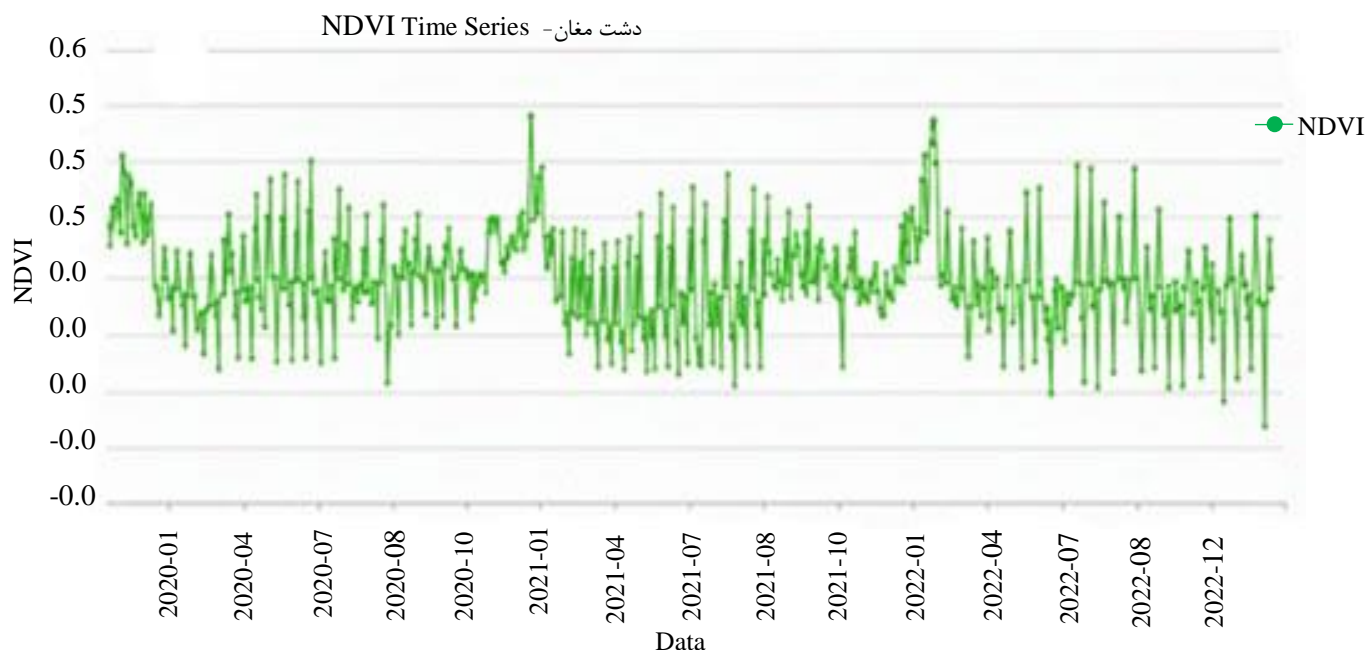
$$Recall = \frac{TP}{TP+Fn} \quad (10)$$

$$F1 = \frac{Precision \times Recall}{Precision + Recall} \times 2 \quad (11)$$

با توجه به اینکه در مسائل شناسایی ناهنجاری، توزیع کلاس‌ها به‌شدت نامتوازن است و ناهنجاری‌ها بخش بسیار کوچکی از داده‌ها را تشکیل می‌دهند، معیار Accuracy می‌تواند گمراه‌کننده باشد، زیرا به دلیل تعداد زیاد TN مقدار آن بزرگ می‌شود. در چنین شرایطی، استفاده از Precision، Recall و به‌ویژه F1 ارزیابی قابل‌اعتمادتری از عملکرد مدل‌ها ارائه می‌دهد. تحلیل تفسیری این شاخص‌ها نشان داد که Precision بالا همراه با Recall پایین به معنای محافظه‌کاری مدل

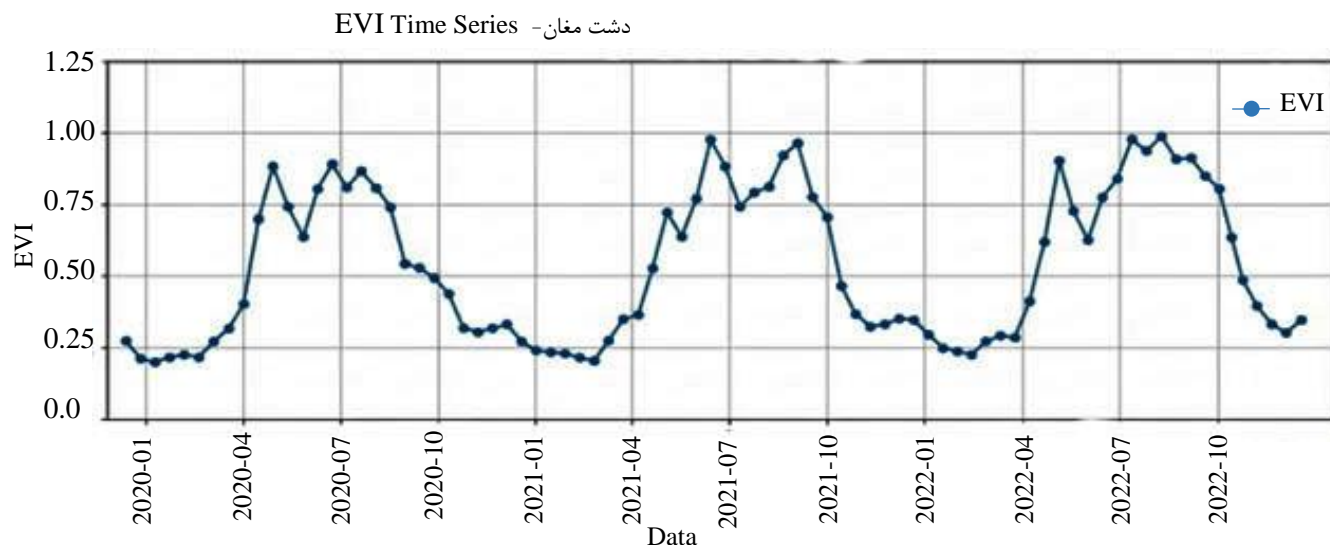
حساسیت را به تغییرات شدید ناشی از آتش‌سوزی و خشکسالی داشت و توانست این تغییرات را با دقت بیشتری ثبت کند (شکل ۵).

به دلیل مقاومت بیشتر در برابر اشباع طیفی در مناطق با پوشش متراکم عملکرد بهتری نشان داد (شکل ۴). در مقابل، NBR بیشترین



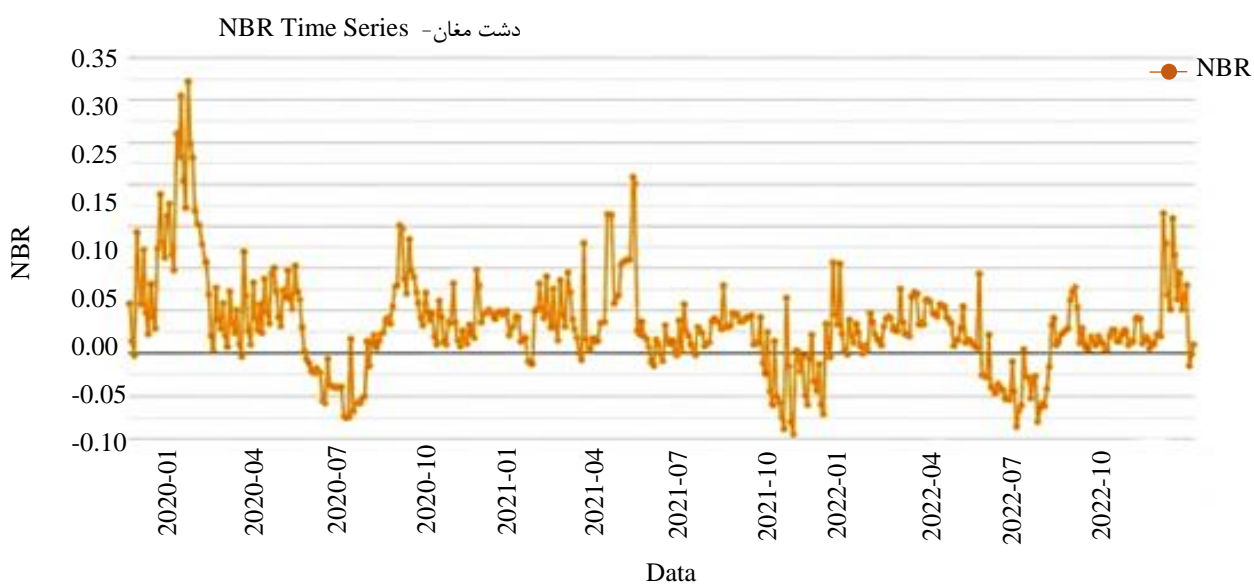
شکل ۳- نمودار سری زمانی NDVI دشت مغان (۲۰۲۰-۲۰۲۲)

Fig. 3. NDVI Time series over Moghan Plain (2020-2022)



کل ۴- نمودار سری زمانی EVI دشت مغان (۲۰۲۰-۲۰۲۲)

Fig. 4. EVI Time series over Moghan Plain (2020-2022)

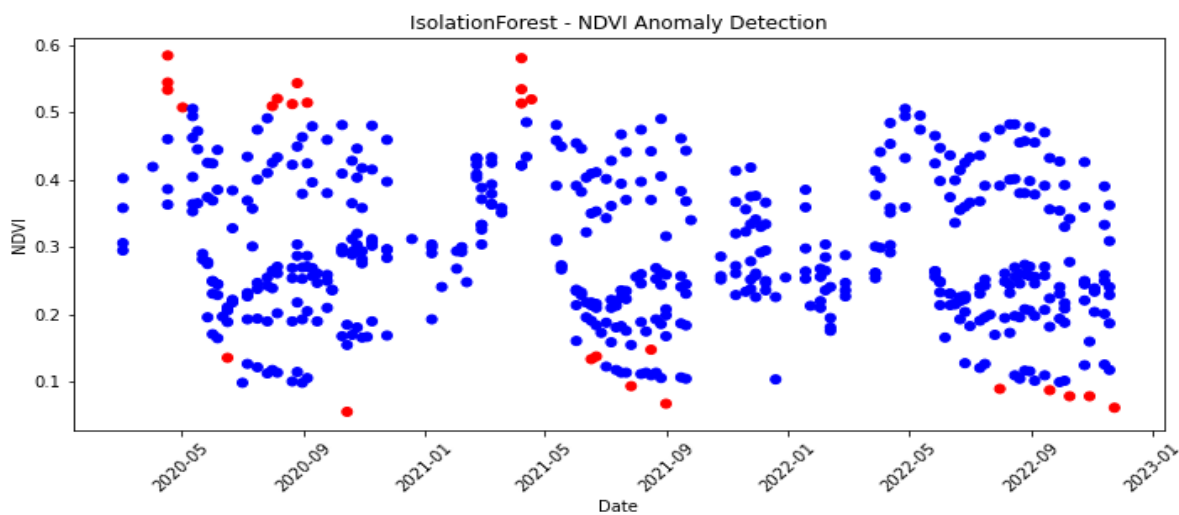


شکل ۵- نمودار سری زمانی NBR دشت مغان (۲۰۲۰-۲۰۲۲)

Fig. 5. NBR Time series over Moghan Plain (2020-2022)

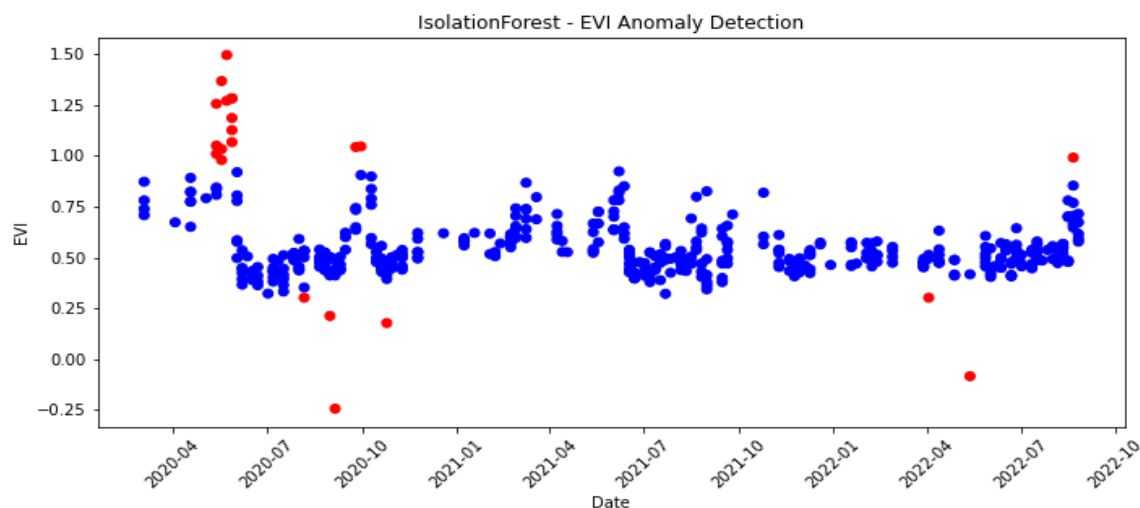
بهتر الگوریتم جنگل جداساز را می‌توان به ماهیت ایزوله‌سازی این روش نسبت داد؛ به گونه‌ای که این الگوریتم بدون وابستگی به توزیع داده‌ها، قادر است تغییرات غیرعادی در سری‌های زمانی شاخص‌های طیفی را با حساسیت بالاتری شناسایی کند. این ویژگی باعث شده است که جنگل جداساز نسبت به نویز طیفی و تغییرات فصلی معمول، پایداری بیشتری نشان دهد (شکل‌های ۶ تا ۸).

پس از استخراج شاخص‌ها، الگوریتم‌های بدون‌ناظر بر روی سری‌های زمانی هر شاخص اعمال شدند. الگوریتم جنگل جداساز با استفاده از درخت‌های تصادفی، توانست نواحی دچار تغییر شدید را به خوبی شناسایی کند. نتایج این الگوریتم در سه شاخص نشان داد که خروجی NDVI و EVI از دقت بالاتری نسبت به NBR برخوردار است. هرچند NBR نیز تغییرات بحرانی را ثبت کرده است. عملکرد



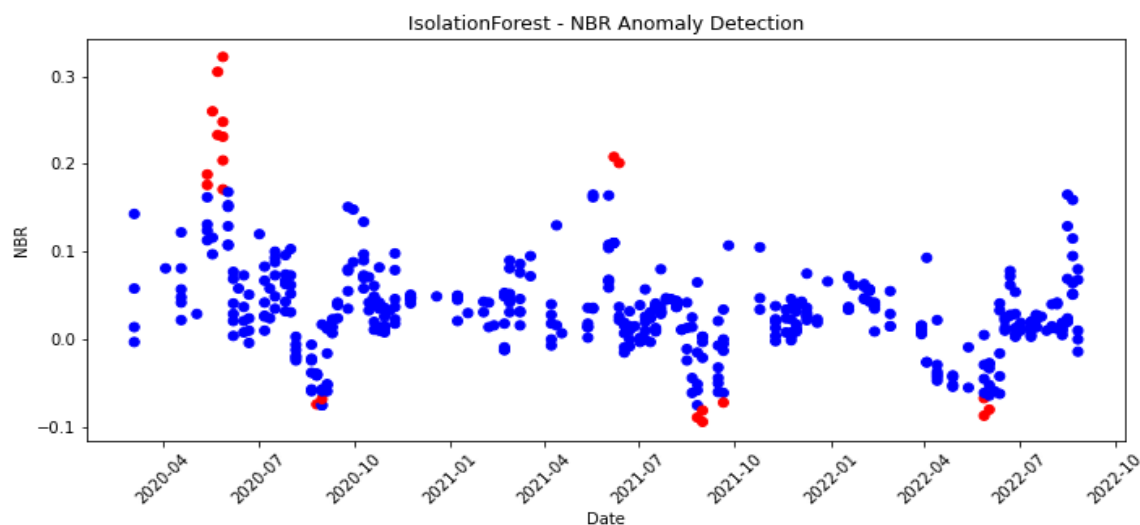
شکل ۶- نمودار تشخیص ناهنجاری NDVI با روش جنگل جداساز

Fig. 6. NDVI Anomaly detection chart with isolation forest algorithm



شکل ۷- نمودار تشخیص ناهنجاری EVI با روش جنگل جداساز

Fig. 7. EVI Anomaly detection chart with isolation forest algorithm

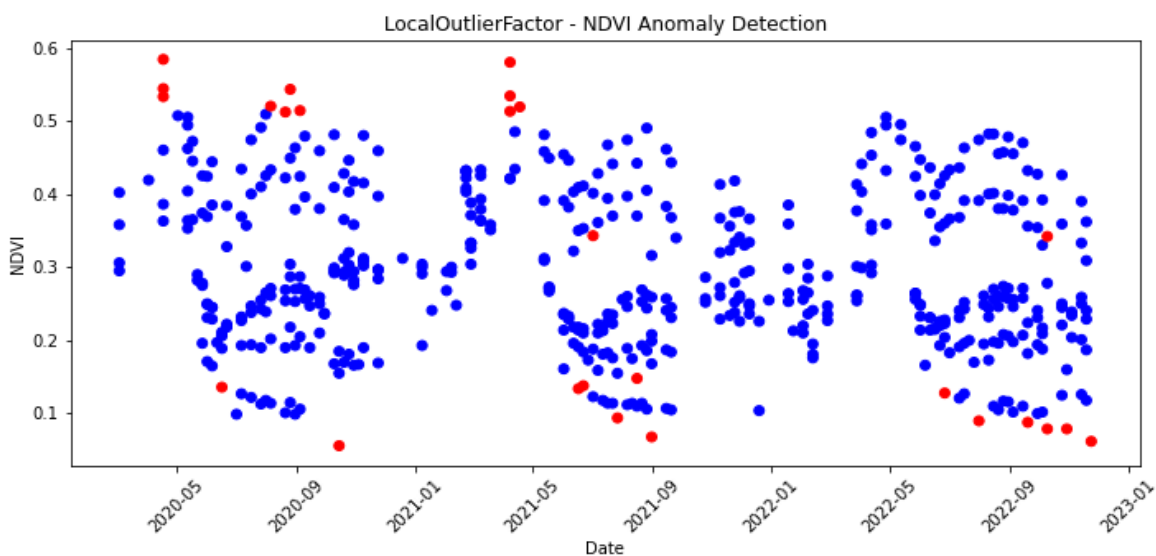


شکل ۸- نمودار تشخیص ناهنجاری NBR با روش جنگل جداساز

Fig. 8. NBR Anomaly detection chart with isolation forest algorithm

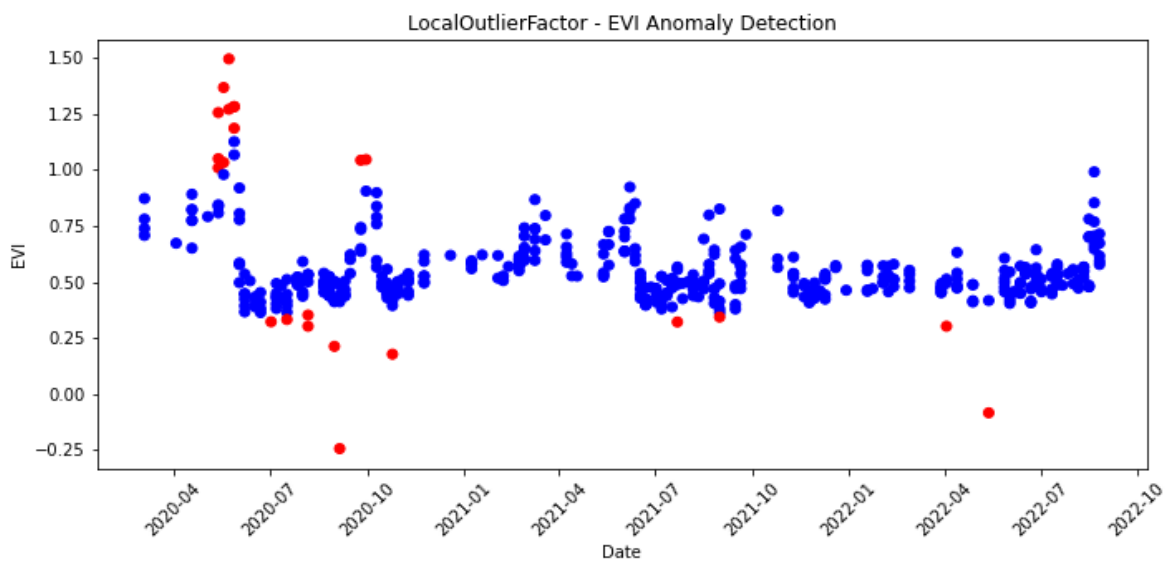
داشت. حساسیت بالاتر الگوریتم عامل ناهنجاری محلی به تغییرات موضعی، ناشی از ماهیت چگالی‌محور این روش است. با این حال، وابستگی شدید LOF به همسایگی محلی سبب می‌شود که در مناطق دارای نویز طیفی یا تغییرات ناگهانی غیرمرتبط با ناهنجاری واقعی، میزان مثبت کاذب افزایش یابد (شکل‌های ۹ تا ۱۱).

الگوریتم عامل ناهنجاری محلی به دلیل ماهیت چگالی‌محور خود، نسبت به تغییرات موضعی حساس‌تر بود و توانست برخی تغییرات خفیف را که سایر روش‌ها نادیده گرفته بودند، آشکار کند. با این حال، در مناطق دارای نویز طیفی بالا نتایج این الگوریتم پراکندگی بیشتری



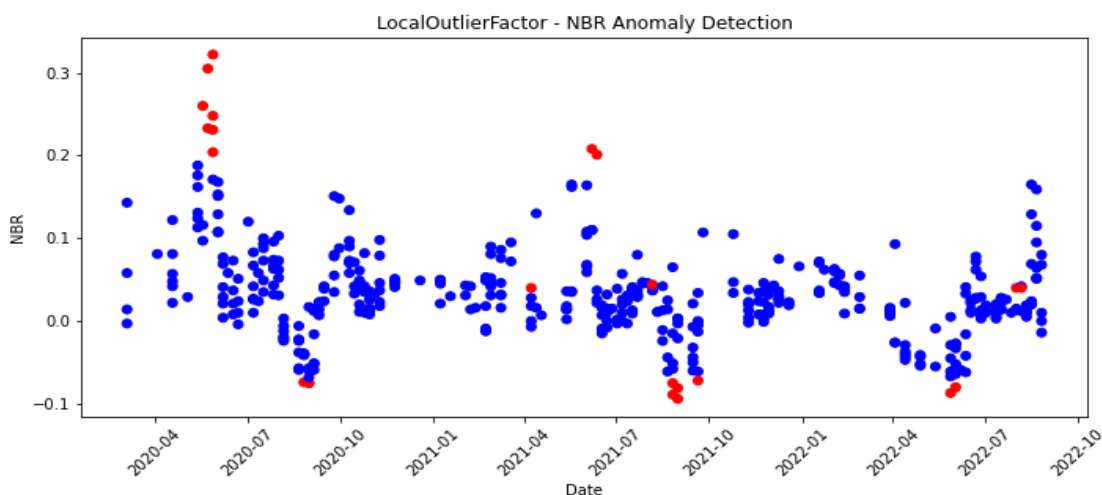
شکل ۹- نمودار تشخیص ناهنجاری NDVI با روش عامل ناهنجار محلی

Fig. 9. NVDI Anomaly detection chart with local outlier factor algorithm



شکل ۱۰- نمودار تشخیص ناهنجاری EVI با روش عامل ناهنجار محلی

Fig. 10. EVI Anomaly detection chart with local outlier factor algorithm

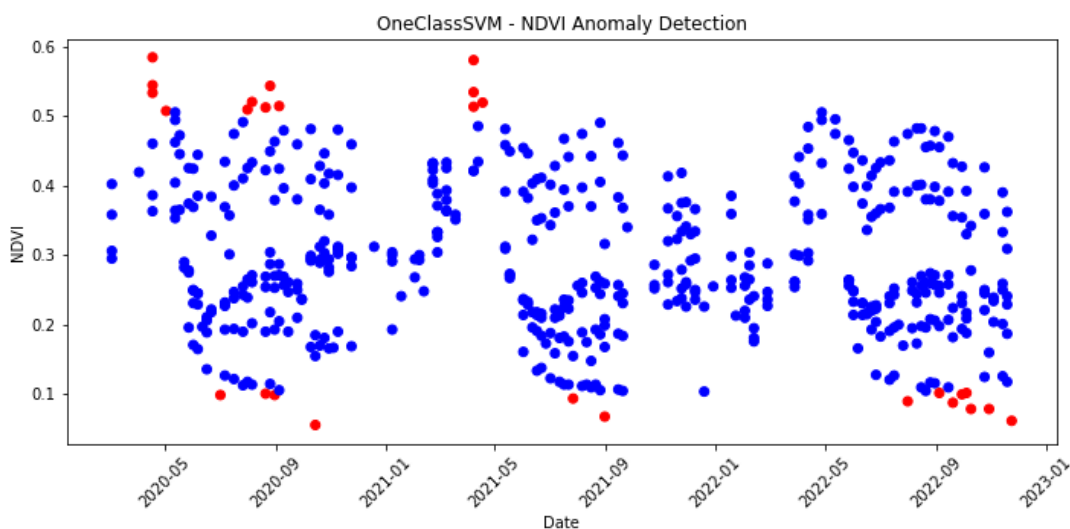


شکل ۱۱- نمودار تشخیص ناهنجاری NBR با روش عامل ناهنجار محلی

Fig. 11. NBR Anomaly detection chart with local outlier factor algorithm

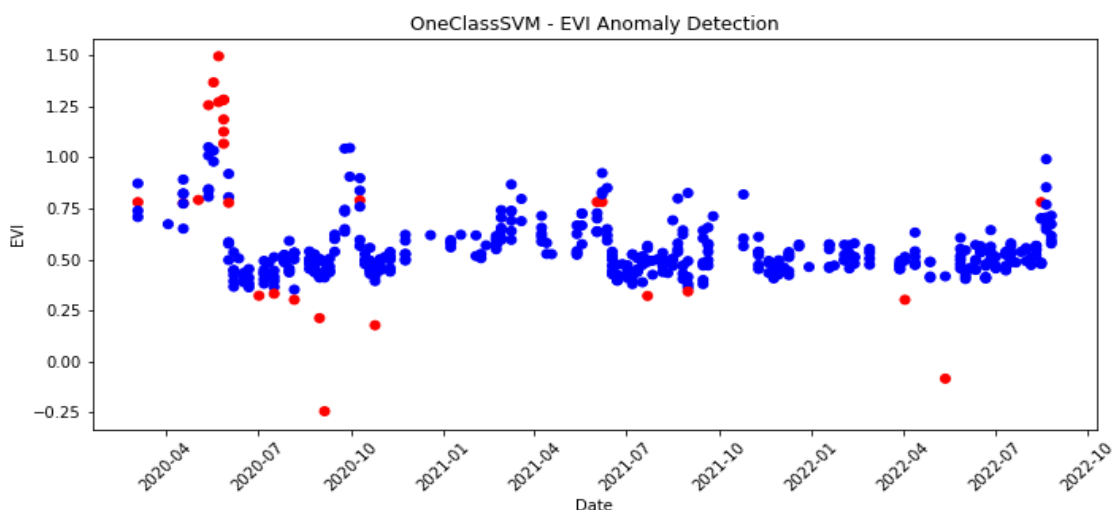
پشتیبان تک کلاسه به دلیل تلاش این الگوریتم برای تعریف یک مرز سخت پیرامون داده‌های نرمال است. این ویژگی موجب کاهش نرخ هشدار کاذب شده، اما در مقابل، بخشی از ناهنجاری‌های خفیف یا تدریجی شناسایی نشده‌اند که خود را به صورت کاهش مقدار Recall نشان می‌دهد (شکل‌های ۱۲ تا ۱۴).

ماشین بردار پشتیبان تک کلاسه عملکردی محافظه کارانه‌تر داشت؛ به این معنا که عمدتاً تغییراتی با اطمینان بالا را برجسب‌گذاری می‌کرد و بسیاری از تغییرات جزئی را نادیده می‌گرفت. این موضوع موجب کاهش Recall شد، اما در عوض نرخ مثبت کاذب کمتری تولید گردید. این ویژگی باعث شد که این مدل برای کاربردهایی که خطای نوع اول هزینه‌بر است، گزینه مناسب‌تری باشد. رفتار محافظه کارانه ماشین بردار



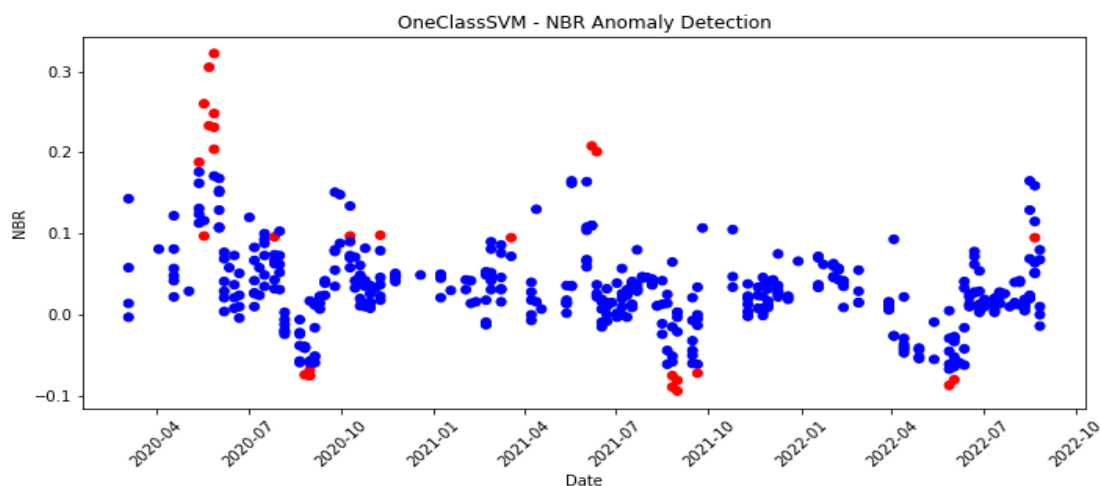
شکل ۱۲- نمودار تشخیص ناهنجاری NDVI با ماشین بردار پشتیبان تک کلاسه

Fig. 12. NDVI Anomaly detection chart with local outlier factor algorithm



شکل ۱۳- نمودار تشخیص ناهنجاری EVI با ماشین بردار پشتیبان تک کلاسه

Fig. 13. EVI Anomaly detection chart with local outlier factor algorithm



شکل ۱۴- نمودار تشخیص ناهنجاری NBR با ماشین بردار پشتیبان تک کلاسه

Fig. 14. NBR Anomaly detection chart with local outlier factor algorithm

جداساز با مقدار  $0/25$  برتری قابل توجهی نسبت به دو الگوریتم دیگر داشت. در نهایت، معیار F1 که نشان‌دهنده تعادل میان Precision و Recall است، نشان داد که جنگل جداساز با مقدار  $0/103$  نسبت به دو روش دیگر عملکرد متوازن‌تری ارائه داده است. تحلیل تفسیری نتایج معیارها نشان داد که جنگل جداساز با وجود حساسیت بالاتر به ناهنجاری‌ها و مقاومت در برابر نویز، بهترین گزینه برای داده‌های این پژوهش است. در حالی که ماشین بردار پشتیبان

برای ارزیابی عملکرد سه الگوریتم، معیارهای Accuracy، Precision، Recall و F1 محاسبه شدند (جدول ۳). نتایج نشان داد که هر سه مدل Accuracy بالاتر از  $0/87$  داشتند، اما این معیار به دلیل نامتوازن بودن داده‌ها به تنهایی کافی نبود. مقایسه بر اساس Precision نشان داد که جنگل جداساز با مقدار  $0/065$  بهترین عملکرد را داشت، در حالی که ماشین بردار پشتیبان تک کلاسه با مقدار  $0/022$  بیشترین مثبت کاذب را تولید کرد. در معیار Recall، جنگل

سه الگوریتم در شکل ۱۵ ارائه شده است که نشان می‌دهد جنگل جداساز به دلیل تعادل نسبی بین Precision و Recall و پایداری در شناسایی نواحی دچار تغییر، کاراترین روش برای کاربردهای پیش محیطی است. نتایج به صورت آماری در جدول ۳ آورده شده است:

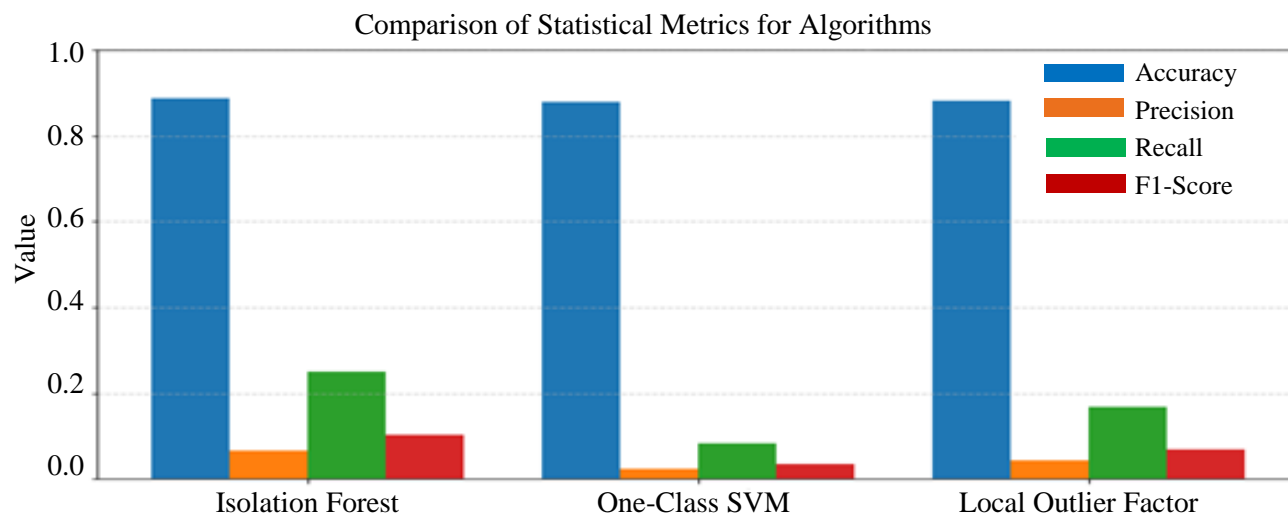
تک‌کلاسه در شناسایی تغییرات جزئی ضعف داشت، اما در شناسایی تغییرات گسترده در شاخص NBR بهتر عمل کرد. از سوی دیگر، عامل ناهنجاری محلی برای شناسایی تغییرات موضعی مناسب‌تر بود، اما ثبات نتایج آن کمتر از جنگل جداساز بود. در نهایت، مقایسه آماری عملکرد جدول ۳- نتایج عملکرد سه الگوریتم

Table 3. Performance results of three algorithms

الگوریتم	Accuracy	Precision	Recall	F1
جنگل جداساز	0.886	0.065	0.250	0.103
ماشین بردار پشتیبان تک کلاسه	0.879	0.022	0.083	0.035
عامل ناهنجاری محلی	0.882	0.043	0.167	0.069

جنگل جداساز به دلیل تعادل نسبی بین Precision و Recall و پایداری در شناسایی نواحی دچار تغییر، کاراترین روش برای کاربردهای پیش محیطی است. اگرچه هر سه الگوریتم مقدار Accuracy بالایی نشان دادند، اما به دلیل نامتوازن بودن داده‌ها، این معیار به تنهایی بیانگر عملکرد واقعی مدل‌ها نیست. برتری الگوریتم جنگل جداساز در معیارهای Recall و F1 نشان می‌دهد که این روش توازن مناسب‌تری میان شناسایی ناهنجاری‌ها و کنترل هشدارهای کاذب برقرار کرده است، که آن را برای کاربردهای پیش محیطی مناسب‌تر می‌سازد.

تحلیل تفسیری نتایج معیارها نشان داد که جنگل جداساز با وجود حساسیت بالاتر به ناهنجاری‌ها و مقاومت در برابر نویز، بهترین گزینه برای داده‌های این پژوهش است. در حالی که ماشین بردار پشتیبان تک‌کلاسه در شناسایی تغییرات جزئی ضعف داشت، اما در شناسایی تغییرات گسترده در شاخص NBR بهتر عمل کرد. از سوی دیگر، عامل ناهنجاری محلی برای شناسایی تغییرات موضعی مناسب‌تر بود، اما ثبات نتایج آن کمتر از جنگل جداساز بود. در نهایت، مقایسه آماری عملکرد سه الگوریتم در شکل ۱۵ ارائه شده است که نشان می‌دهد



شکل ۱۵- نمودار مقایسه آماری سه الگوریتم

Fig. 15. Comparison chart of three algorithms

## نتیجه گیری

در برابر نویز و تغییرات فصلی، توانست هم ناهنجاری‌های تدریجی و هم اختلالات شدید پوشش گیاهی را به صورت مؤثر شناسایی کند، که این ویژگی آن را برای کاربردهای عملی در پایش محیطی، مدیریت کشاورزی و سامانه‌های هشدار سریع مناسب می‌سازد.

استفاده از داده‌های آزاد و در دسترس سنتینل-۲ در کنار پیاده‌سازی مدل‌ها در محیط‌های پردازش ابری و متن‌باز، موجب افزایش قابلیت توسعه‌پذیری، مقیاس‌پذیری و بازتولیدپذیری چارچوب پیشنهادی شده است. هرچند مقادیر مطلق برخی معیارهای ارزیابی مانند F1 در این پژوهش نسبتاً پایین هستند، این موضوع ناشی از نامتوازن بودن شدید داده‌ها و ماهیت ذاتی مسئله تشخیص ناهنجاری بدون ناظر است و بنابراین، مقایسه نسبی عملکرد الگوریتم‌ها به‌عنوان معیار اصلی ارزیابی در نظر گرفته شده است. همچنین، خروجی مدل‌ها در سطح پیکسل تولید شده و امکان تحلیل و نمایش مکانی ناهنجاری‌ها را فراهم می‌کند، اگرچه تمرکز اصلی این مطالعه بر تحلیل زمانی تغییرات پوشش گیاهی بوده است.

در مجموع، نتایج این پژوهش نشان می‌دهد که تلفیق شاخص‌های چندگانه پوشش گیاهی با الگوریتم‌های سبک یادگیری ماشین

در این پژوهش، یک چارچوب داده‌محور و بازتولیدپذیر برای شناسایی ناهنجاری‌های محیطی با استفاده از تصاویر چندزمانه ماهواره‌ای سنتینل-۲ و الگوریتم‌های یادگیری ماشین بدون ناظر ارائه شد. نتایج حاصل از ارزیابی کمی مدل‌ها نشان داد که ترکیب شاخص‌های طیفی NDVI، EVI و NBR با الگوریتم جنگل جداساز، عملکرد برتری نسبت به سایر روش‌های بررسی شده دارد. به‌طور مشخص، الگوریتم جنگل جداساز با دستیابی به مقدار F1 برابر با ۰/۱۰۳ و Recall معادل ۰/۲۵۰، در مقایسه با عامل ناهنجاری محلی (F1 = 0.069) و ماشین بردار پشتیبان تک‌کلاسه (F1 = 0.035)، توازن مناسب‌تری میان شناسایی ناهنجاری‌ها و کنترل هشدارهای کاذب برقرار کرده است.

مقایسه عملکرد الگوریتم‌ها نشان داد که عامل ناهنجاری محلی اگرچه در شناسایی تغییرات موضعی حساس‌تر است، اما در داده‌های دارای نویز پایداری کمتری دارد، در حالی که ماشین بردار پشتیبان تک‌کلاسه رویکردی محافظه‌کارانه اتخاذ کرده و عمدتاً ناهنجاری‌های شدید را شناسایی می‌کند. در مقابل، جنگل جداساز با مقاومت بیشتر

توصیه می‌شود علاوه بر NDVI، شاخص‌هایی نظیر NDWI، SAVI و GNDVI نیز در فرآیند شناسایی ناهنجاری به کار گرفته شوند، زیرا ترکیب آن‌ها می‌تواند حساسیت مدل را نسبت به تغییرات رطوبتی، تراکم پوشش گیاهی و ساختار تاج گیاه افزایش دهد.

استفاده از داده‌های چندمنبعی

ترکیب داده‌های اپتیکی سنتینل - ۲ با داده‌های راداری سنتینل - ۱، به‌ویژه در مناطق ابری یا فصل‌های بارانی، می‌تواند پوشش زمانی و مکانی پایش را افزایش داده و اثرات مفقود شدن داده‌ها بر دقت مدل را کاهش دهد.

به‌کارگیری مدل‌های یادگیری عمیق

استفاده از معماری‌های یادگیری عمیق مانند شبکه‌های کانولوشنی (CNN) و شبکه‌های بازگشتی (RNN/LSTM)، با قابلیت یادگیری ویژگی‌های مکانی-زمانی پیچیده، می‌تواند موجب افزایش چشمگیر دقت شناسایی تغییرات شود.

تحلیل بلندمدت اثرات تغییرات اقلیمی

بررسی داده‌های چندساله و مقایسه نتایج با پارامترهای اقلیمی (دما، بارش، شاخص خشکسالی) می‌تواند کمک کند تا نقش عوامل اقلیمی در بروز ناهنجاری‌های پوشش گیاهی به‌طور دقیق‌تری مشخص شود.

## References

- Abdalla Mahmoud, A.M., Rayne, L., Bewley, R., 2025. EAMENA MLACD: Automated remote sensing tool based on machine learning in Google Earth Engine using Sentinel-2 imagery for detecting changes and threats around archaeological sites. *Remote Sensing Applications: Society and Environment* 37, 101396.
- Alshari, E.A., Gawali, B.W., 2022. Analysis of machine learning techniques for land use/land cover classification using Sentinel-2A satellite imagery. *Journal of Electrical and Computer Engineering* 9092299.
- Babaeepour, A., Milan, A., 2024. Landuse change mapping based on drone and ultracam aerial images. *Kharazmi Journal of Earth Sciences* 10 (2), 443-464
- Descals, A., Torres, K., Verger, A., Peñuelas, J., 2025. Evaluating Sentinel-2 for monitoring drought-induced crop failure in winter cereals. *Remote Sensing* 17 (2), 340.

بدون‌ناظر، رویکردی کارآمد، مقرون‌به‌صرفه و قابل‌اتکا برای شناسایی ناهنجاری‌های محیطی در مناطق کشاورزی فراهم می‌سازد، به‌ویژه در شرایطی که دسترسی به داده‌های میدانی محدود است. این چارچوب می‌تواند به‌عنوان مبنایی عملی برای توسعه سامانه‌های پایش هوشمند و تصمیم‌یار در مدیریت منابع طبیعی و کشاورزی دقیق مورد استفاده قرار گیرد.

به‌طور کلی، یافته‌های این پژوهش نشان می‌دهند که بهره‌گیری از شاخص‌های چندگانه پوشش گیاهی در کنار مدل‌های سبک یادگیری ماشین، ابزاری کارآمد و بازتولیدپذیر برای شناسایی ناهنجاری‌های محیطی فراهم می‌سازد. این رویکرد به‌ویژه در کشورهایی با محدودیت منابع مالی و زیرساختی ارزش بالایی دارد. برای توسعه آتی، پیشنهاد می‌شود ترکیب شاخص‌های طیفی بیشتر (مانند شاخص‌های رطوبتی و حرارتی) و بهینه‌سازی خودکار پارامترهای الگوریتم‌ها با روش‌های فراابتکاری مورد بررسی قرار گیرد تا کارایی و پایداری چارچوب پیشنهادی افزایش یابد. با توجه به یافته‌ها و محدودیت‌های تحقیق حاضر، پیشنهادات زیر جهت توسعه مطالعات آتی و افزایش کارایی چارچوب پیشنهادی ارائه می‌شود:

افزایش تنوع شاخص‌های گیاهی

- Farahnakian, F., Luodes, N., Karlsson, T., 2024. Machine learning algorithms for acid mine drainage Mapping using Sentinel-2 and WorldView-3. *Remote Sensing* 16 (24), 4680.
- Hosseini, A., Karami, M., Moghaddam, R., 2022. Assessment of drought impacts on agricultural lands using Sentinel-2 NDVI in Nishabour plain, Iran. *Iranian Journal of Remote Sensing and GIS* 14 (2), 45-60.
- Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., Ferreira, L.G., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment* 83 (1-2), 195-213.
- Ju, J., Zhou, Q., Freitag, B., Roy, D.P., Zhang, H.K., Sridhar, M., Mandel, J., Arab, S., Schmidt, G., Crawford, C.J., Gascon, F., Strobl, P.A., Masek, J.G., Neigh, C.S.R., 2025. Harmonized Landsat and Sentinel-2 surface reflectance

- data (HLS V2.0): Consistent multispectral observations for large-scale terrestrial monitoring. *Remote Sensing of Environment* 324, 114723.
- Key, C. H., Benson, N.C., 2006. Landscape Assessment: Ground Measure of Severity, the Composite Burn Index; and Remote Sensing of Severity, the Normalized Burn Ratio. in: Lutes, D.C., Keane, R.E., Caratti, J.F., Key, C.H., Benson, N.C., Sutherland, S., Gangi, L.J. (Eds.), FIREMON: Fire Effects Monitoring and Inventory System. Ogden, UT: USDA Forest Service, Rocky Mountain Research Station, Gen. Tech. Rep. RMRS-GTR-164-CD, pp. LA1–LA51.
- Khan, M., Singh, R., Ahmad, I., 2023. Detection of unusual vegetation changes using local outlier factor in Punjab, India. *Environmental Monitoring and Assessment* 195 (7), 849.
- Kovárník, R., Janová, J., 2025. Validation of machine learning models based on Sentinel-2 for national forest inventory. *Ecological Informatics* 87, 103133.
- Liang, D., Li, Y., Zhang, F., 2025. Forest anomaly comprehensive index (FACI) based on Sentinel-2 imagery for detecting various forest anomalies. *Forests* 16 (3), 497.
- Pacheco-Pascagaza, A.M., Ruiz, L.A., García, M., 2022. Near real-time change detection system using Sentinel-2 and machine learning for forest monitoring. *Remote Sensing* 14 (4), 707.
- Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., 1974. Monitoring vegetation systems in the Great Plains with ERTS. *NASA Special Publications* 351, 309–317.
- Safa, R., Ziaian Firouzabadi, P., Tokoli Sabour, S. M., 2023. Detection of changes in land cover in satellite imagery using the discrete fractional Fourier transform. *Kharazmi Journal of Earth Sciences* 9(1), 116–149.
- Schiller, C., Költzow, J., Schwarz, S., Schiefer, F., Fassnacht, F.E., 2024. Transformer-based Forest disturbance detection using Sentinel-2 time series. *Remote Sensing of Environment* 315, 114475.
- Tempa, K., Ilunga, M., Agarwal, A., Tashi, P., 2024. Land use/land cover change dynamics and NDVI analysis using Sentinel-2 in Gelephu, Bhutan. *Applied Sciences* 14 (4), 1578.
- Wasehun, E.T., Kenea, T., Abera, W., 2025. Comparative analysis of Sentinel-2 and PlanetScope imagery for predicting chlorophyll-a using machine learning. *Ecological Informatics* 85, 102988.
- Zhao, X., Wang, Y., Liu, Y., 2022. Vegetation anomaly detection using NDVI time series and isolation forest algorithm. *Journal of Applied Remote Sensing* 16 (3), 036510.
- Encyclopaedia Iranica, n.d. Moğān., 2015. Available at: <https://www.iranicaonline.org/articles/mogan-parent-i-ii>
- Solgi, M., Kamkar, B., Ghorbani, M., 2018. Assessment of agricultural characteristics of Moghan plain (Ardabil province, Iran). *Scientific Papers Series Management, Economic Engineering in Agriculture and Rural Development* 18 (4), 289–296.