University of Tehran

*Earth Observation and Geomatics Engineering*

Online ISSN: 2588-4360

Homepage: https://eoge.ut.ac.ir/

# Semi-Automatic Monitoring of Miankale Wetland Drought Based on a Hybrid Anomaly Detection Method Using Time Series Satellite Images in Google Earth Engine

**Parisa Dodangeh**[1], **Reza Shah-Hosseini**[2]✉

1. School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran. E-mail: pdodangeh@ut.ac.ir
2. Corresponding author, School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran. E-mail: rshahosseini@ut.ac.ir

## Article Info

**Article type:**
Research Article

**Article history:**
Received 2024-03-30
Received in revised form 2024-07-04
Accepted 2024-09-11
Available online 2025-01-01

***Keywords***:
Anomaly detection,
Wetland,
Remote sensing,
Google Earth Engine,
Drought monitoring.

## ABSTRACT

Wetlands are invaluable ecosystems at risk of destruction due to drought. Continuous monitoring of drought over the years is crucial for environmental management. Traditional mapping methods are expensive and time-consuming. Remote sensing techniques provide a more efficient alternative, allowing for the survey of large geographic areas in a short period. This research aims to detect drought-affected areas in the Miankale wetland, Mazandaran province, Iran, from 2009 to 2021 using a semi-automatic hybrid approach combining multiple anomaly detection algorithms. Time series Landsat satellite images were used to identify effective drought indicators. Clustering-based methods identified anomalous temporal and spatial breakpoints, followed by statistical and machine learning techniques to produce an accurate wetland drought map. Evaluation using ground truth images yielded an overall accuracy of 97.89% and an F1 Score of 98.91%. This study fills a gap by utilizing anomaly detection methods for drought monitoring, presenting a fast and accurate approach that leverages the maximum capacity of satellite data and minimizes errors through a combination of different methods.

**Cite this article:** Dodangeh, P., Shah-hosseini, R. (2023). Semi-Automatic Monitoring of Miankale Wetland Drought Based on a Hybrid Anomaly Detection Method Using Time Series Satellite Images in Google Earth Engine. *Earth Observation and Geomatics Engineering*, Volume 7, Issue 2, Pages 9-23. http//doi.org/10.22059/eoge.2024.374541.1148



Publisher: University of Tehran.

DOI: http//doi.org/10.22059/eoge.2024.374541.1148

## 1. Introduction

Natural wetland ecosystems provide not only important habitats for many wildlife species but also food for migratory and resident animals (Lin et al., 2018). They generate a variety of benefits to society and nature, such as providing fertile soils for agriculture, food, and habitat for shorebirds, generating oxygen, adjusting climate, improving water quality, etc. Therefore, they have been regarded as one of the most valuable resources in the world (Barbier et al., 2011). However, the landscapes of wetlands have changed or even been lost due to negligent wetland management (Gong et al., 2010; Nicholls et al., 1999). Totally, because of the increasing population, industrialization, increasing dependence on irrigation, infrastructure deficiencies, and the inherent high variability of precipitation and discharge; water resource scarcity is already common in many regions of the world (Thakur et al., 2021). In the context of wetlands, drought refers to prolonged periods of insufficient precipitation resulting in reduced water levels, which affect both the land and water parts of the wetland ecosystem. On land, drought can lead to soil moisture depletion, vegetation stress, and habitat loss for terrestrial wildlife. In the water parts, it can cause reduced water quality, diminished aquatic habitats, and a decrease in the overall water volume. Therefore, it is crucial to monitor and manage these changes to maintain the ecological balance and health of wetlands. In any case, considering the importance of the mentioned topic, the continuous monitoring of wetlands from different perspectives should be the attention of researchers and managers in the field of environment. For example, it is very important to examine the process of changes in wetlands over time. Usually, changes in wetlands are the reduction of water volume and droughts that occur over a long period. Therefore, continuous monitoring of these changes will be very helpful to prevent their loss. In addition to that, in the issue of drought, there are various factors involved, whose identification can control the process of drought.

According to the mentioned necessities, several kinds of research have been conducted in the direction of wetlands drought monitoring. In 2018, Yua et al. presented a study on this issue. In their study, Vegetation Temperature Condition Index (VTCI) derived from NDVI and LST[1] was used to observe the drought status of the wetland in the Tumen River Basin from 1991 to 2016. For this purpose, the Landsat TM/ETM+ data for six periods were used for the analysis (Yu et al., 2019). Although the results of this research have had good quality, the number of factors considered for review was limited. Also, the preparation and pre-processing of the used time series required a lot of time. In 2019, Lv et al. focused on the drought in the Xiong'an New Area. The studied period was from 1980 to 2015, for which Landsat data was used. Their proposed approach was based on three indices NDVI, NDWI, SMMI[2], and Z_statistics. Their investigation showed a degradation, a slight improvement and a degradation trend in the study area, respectively (Lv et al., 2019). Li et al. used multiple drought indices (DIs), including Percent of Normal (PN), Standardized Precipitation Index (SPI), statistical Z-Score, and Effective Drought Index (EDI) at 18 different timesteps were employed to evaluate the drought condition in Wuyuer River Basin, Northeast China. Daily precipitation data for 50 years (1960–2010) from three meteorological stations were used in this study. They found DIs with intermediate time steps (7 to 18 months) to have the highest predictive values for identifying droughts and DIs showed a better similarity in the 12-month timestep. Their results showed among all the DIs, EDI exhibited the best correlation with other DIs for various timesteps (Li et al., 2019). Research had shown that in the field of identifying changes and droughts, the use of related indicators has been effective. In this regard, using and combining different approaches can provide better results. In the following, in 2021, Sarkar et al. (Sarkar et al., 2021) made an attempt at spatiotemporal change analysis of three floodplain wetlands of Eastern India using GIS tools. It was found that the surveyed wetlands have reduced in size by 37.20–57.68% coupled with a reduction in minimum and maximum depth reported from these wetlands. The analysis of data (1985–2018) indicated a considerable change in climate with an average temperature (1.9 °C) and rainfall anomaly (−698.1 mm) in the study area during the year 2018. Therefore, the factors used in this research were precipitation and temperature. One of the common points of most research is utilizing a large volume of remote sensing data in a long time series. This issue increases the processing time and, in some cases, the implementation of the algorithm faces hardware challenges. For this reason, using a web-based cloud system such as Google Earth Engine can significantly help to carry out research more easily.

Remote sensing satellite imagery provides long-term earth observations, accumulating a large amount of time series data and providing a possible means to learn the long-term spatiotemporal distribution of wetland land cover types (Wang et al., 2020). Among various remote sensing data, Landsat images providing the large-scale historical and current status of natural resources have been widely employed for examining long-term wetland dynamics (Ji et al., 2018; Ji et al., 2015). Recently, a system called Google Earth Engine (GEE) has made it easy to access a huge number of different images, including the Landsat satellite. Since its official establishment in 2010, Google Earth Engine (GEE) has developed rapidly and has played a

[1] Land Surface Temperature

[2] Soil Moisture Monitoring Index

significant role in the global remote sensing community (Pham-Duc et al., 2023). It enables cloud computation and is an operative tool for analyzing global geospatial big data (Zhao et al., 2021). It also supports more types of geospatial data (for example, Sentinel and Landsat data) and provides services accessible to all users. GEE is currently the most approved cloud computing platform in geoscience studies and is broadly used to process data related to a variety of fields, including agriculture (Shelestov et al., 2017; Teluguntla et al., 2018), water (Huang et al., 2018; Zhou et al., 2019), land cover/land use (Zurqani et al., 2018), disasters (DeVries et al., 2020), climate change (Hao et al., 2019), soil (Ivushkin et al., 2019), wetland (Amani et al., 2019), forest (Bullock et al., 2020) and urbanization (Gong et al., 2020) as well as other fields. GEE can provide users with public free Earth observation data at the petabyte scale, advanced algorithms for analyzing geographic big data, and an interactive programming environment. It also hosts long time series of Earth observation records and plays a vital role in environmental monitoring and analysis (Amani et al., 2020). For this purpose, in the present research, the mentioned system has been utilized to use the high volume of data in a time series format for study area monitoring.

Since wetland drought is an inconsistency in nature, anomaly detection algorithms can be used to monitor this issue. Anomaly detection is a prominent data analysis technique that detects anomalous or abnormal data from a given dataset (Ahmed & Mahmood, 2014). Anomalies are considered important because they indicate eloquently but out of ordinary events and can prompt critical actions to be taken in a wide range of application domains (Ahmed et al., 2016); Anomaly detection algorithms, identifying unusual data patterns, are valuable for monitoring inconsistencies like wetland droughts. Despite their potential, these algorithms are rarely applied in wetland monitoring. This research proposes a semi-automatic method combining simple and fast anomaly detection algorithms to monitor wetland droughts. Key contributions include:

- A combined method reducing error probability in drought monitoring while minimizing user intervention and enhancing automation.
- Inclusion of multivariate remote sensing factors to improve drought detection accuracy.
- Long-term Landsat data for periodic monitoring was implemented on the GEE cloud system to reduce processing time and hardware needs.
- Preparation of a ground truth map using Landsat and Google Earth images, validated by user knowledge.

According to the items mentioned above, this research is divided into different sections, including materials and methods, experimental results, discussion, and conclusion.

## 2. Material and methods

This section of the research describes the details of the various parts of the implementation approach, such as the study area, the data used, the proposed approach, and its different stages. For this reason, the material and method section is divided into three sub-sections, including the study area, data, and proposed method.

### 2.1. Study area

The study area in this research is the Miankale wetland located in the southeast extremity of the Caspian Sea in Mazandaran province of Iran (52° 25′ – 54° 02′ E longitude and 36° 46′ – 36° 53′ N latitude), which has suffered from a drought crisis in recent years. Its water level has decreased due to various reasons, such as climatic changes, a decrease in rainfall, human factors, etc. Miankale Wetland is one of the most productive ecological harbors in West Asia and maybe the world. The elongated wetland is 48 kilometers (30 miles) long and between 1.3 and 3.2 kilometers (1,400 and 3,500 yards) wide. The mean annual precipitation of this area is about 600 mm, and the region's mean annual temperature is 17 °C with a climate ranging from warm semi-humid to temperate. The wetland's maximum, minimum, and mean elevations are −27, −20, and −24 m, respectively (Gholami et al., 2020). An overview of the study area can be seen in Fig.1.

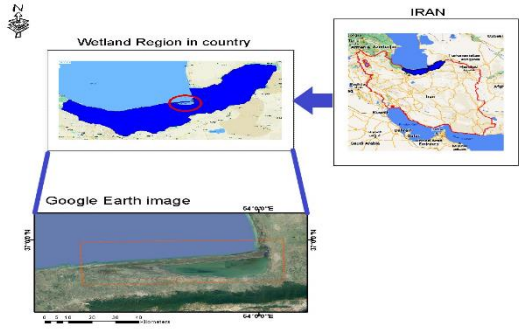



Figure 1. Study area.

### 2.2. Data and Pre-processing

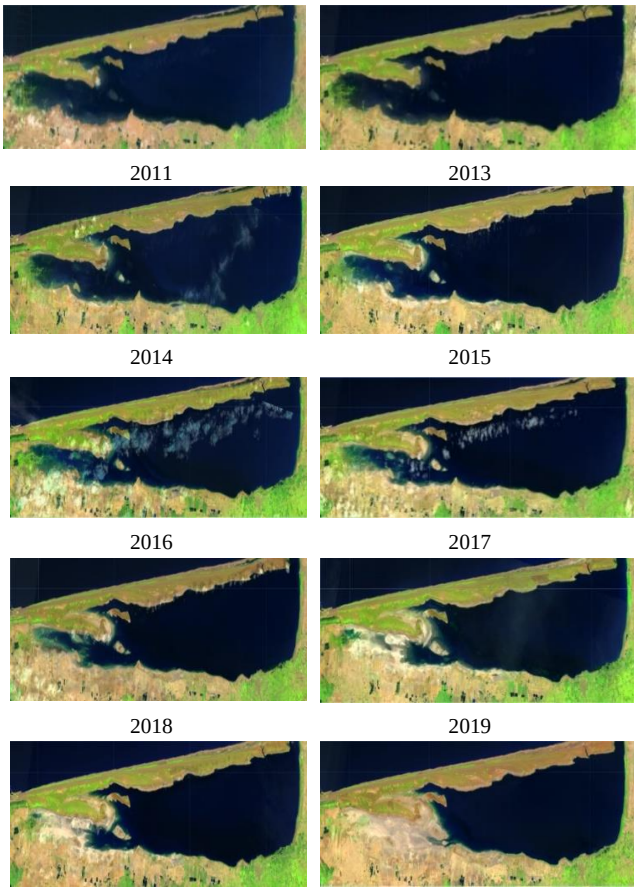
Given that drought occurs over a long period, a compact time series is needed. Table 1 describes the time series prepared from Landsat satellite images with a spatial resolution of 30 meters.

**Table 1. Dataset preparation**.

| Platform | Date | Number of images |
|---|---|---|
| Landsat 5 TM | 2009-05-22, 2009-09-23 | 23 images |
| Landsat 5 TM | 2010-05-22, 2010-09-23 | |
| Landsat 5 TM | 2011-05-22, 2011-09-23 | |
| Landsat 8 OLI | 2013-05-22, 2013-09-23 | 95 images |
| Landsat 8 OLI | 2014-05-22, 2014-09-23 | |
| Landsat 8 OLI | 2015-05-22, 2015-09-23 | |
| Landsat 8 OLI | 2016-05-22, 2016-09-23 | |
| Landsat 8 OLI | 2017-05-22, 2017-09-23 | |
| Landsat 8 OLI | 2018-05-22, 2018-09-23 | |
| Landsat 8 OLI | 2019-05-22, 2019-09-23 | |
| Landsat 8 OLI | 2020-05-22, 2020-09-23 | |
| Landsat 8 OLI | 2021-05-22, 2021-09-23 | |

The time series consists of 12 years, all of which the summer season is considered. In this way, implementations are possible under the same conditions for each year. In addition, since rainfall does not usually occur in summer, this season is better for examining drought. On the other hand, instead of Landsat 7 images, Landsat 5 is used to prevent further calculation errors. It should be noted that due to the use of the GEE platform and Landsat TOA reflectance images that have already undergone pre-processing, no other corrections have been made, accelerating the process of future processing. Fig.2 shows the mean image of each year mentioned in the table above.

2009 2010



2011 2013 2014 2015 2016 2017 2018 2019 2020 2021

Figure 2. Mean of images in different years

The accuracy assessment step can be considered one of the main steps in developing algorithms and implementing various research operations. This action requires having a ground truth map so that the accuracy evaluation criteria can be estimated with its help. This research provided the ground truth map based on images from Landsat and Google Earth. This map can be seen in Fig.3.

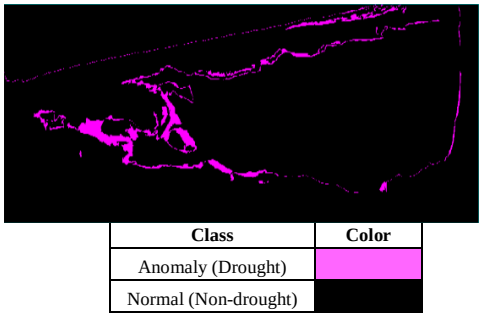



Figure 3. Ground truth map

**2.3. Methodology**

This part presents the proposed approach, which ultimately forms a hybrid and semi-automatic method. Fig.4

shows the processing flowchart by running the methods for detecting anomalies in Google Earth Engine.

We adopt a comprehensive approach integrating remote sensing data and anomaly detection techniques to identify drought conditions in the Miankale wetland. The analysis spans a time series of Landsat data from 2009 to 2021, processed using Google Earth Engine (GEE) for efficient image analysis. To assess drought comprehensively, we consider multiple factors influencing wetland health: the Normalized Difference Vegetation Index (NDVI), Modified Normalized Difference Water Index (MNDWI), Land Surface Temperature (LST), and Soil Salinity Index (SI). NDVI and MNDWI assess vegetation and water dynamics, respectively, while LST and SI provide insights into temperature and soil conditions affecting both land and water areas. The anomaly detection process begins with spectral and spatial breakpoint identification. Spatial breakpoints are identified using K-means clustering on a comprehensive stack of all Landsat data bands. This clustering approach categorizes the wetland into distinct clusters based on spectral characteristics, identifying areas with anomalous behavior indicative of potential drought conditions. Subsequently, histograms of time series values within each cluster pinpoint temporal breakpoints, marking years of significant environmental change. Automatic training sample preparation for drought mapping involves statistical methods such as Z-score analysis applied to the anomalous clusters identified earlier. The Z-score measures each factor's deviation within anomalous clusters from its mean, facilitating the selection of training samples that accurately represent drought and non-drought conditions. We employ a Random Forest classifier on Landsat 8 Top of Atmosphere (TOA) Reflectance images to generate the final drought map. Random Forest is chosen for its ability to handle large datasets with multiple features, robustness against overfitting, and high accuracy in classification tasks. Training samples derived from the anomaly detection phase guide the classifier in delineating drought-affected areas within the wetland. To validate the method's accuracy, the resulting drought map is compared against a ground truth map prepared using Landsat and Google Earth images. Evaluation metrics, including F1 score, overall accuracy, and Kappa coefficient, are computed to assess the reliability and performance of the classification model.

In summary, this methodological approach ensures a comprehensive assessment of drought conditions in the Miankale wetland, addressing both land and water aspects crucial for effective wetland management under changing environmental conditions.

In the following, all the steps of the proposed approach are described in detail.

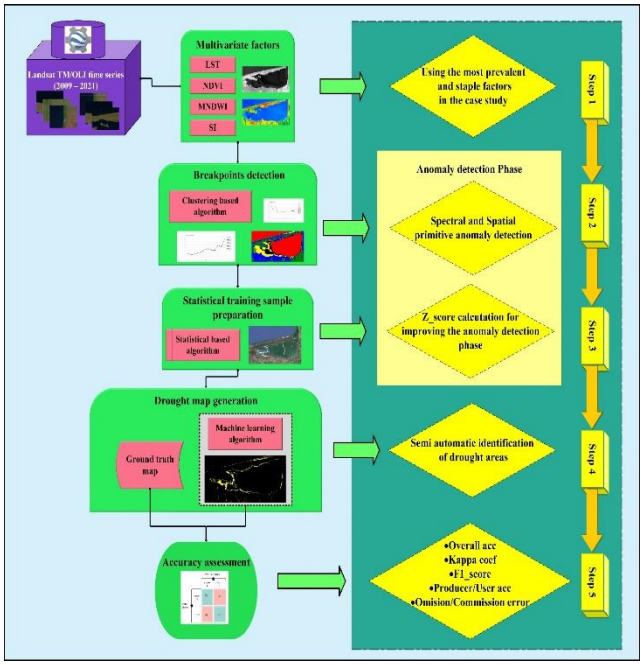



Figure 4. Proposed method's flowchart

2.3.1. Multivariate remote sensing factors

Many factors cause wetland changes. For example, climate change-induced temperature and precipitation alterations affect wetland distribution (spatial extent and health of wetland areas, indicating how well the wetland is maintaining its ecological functions) by changing hydrological and biochemical cycles (Junk et al., 2013). Soil properties are one of the critical factors in wetland distribution (Meng et al., 2020). Research has shown that areas with high soil organic matter and low soil salinity are more favorable for maintaining healthy wetland ecosystems. In addition, anthropogenic activities, like cultivation for agricultural land use, are the most influential factors attributed to wetland loss (Bolca et al., 2007). In addition, these factors often interact with wetland changes (McLaughlin & Cohen, 2013), and the dominant factors resulting in wetland change are inconsistent during various periods (Li et al., 2021; Werkmeister et al., 2018). According to the investigations (Ghosh et al., 2018; Hou et al., 2020; Walter & Mondal, 2019; Zhang et al., 2021), four critical factors have been extracted as the basis of anomaly detection from the time series in each year. These factors include the Normalized difference vegetation index **(**NDVI) (Jiang et al., 2006), Modified normalized difference water index **(**MNDWI) (Xu, 2006), Land surface temperature **(**LST) (Ermida et al., 2020), and Salinity index (SI) (Aksoy et al., 2022), the details of which can be found in the references. In the following, processing has been done on these factors due to their strong impact on drought detection.

2.3.2. Spectral and Spatial Breakpoints Detection

In the second step, it is necessary to identify breakpoints

during the existing time series from different years. In terms of time, the breakpoint in this study means the years when the values of the factors in the region have changed drastically. In terms of location, it is also necessary to recognize areas with severe changes. At first, spatial breakpoints are identified by clustering and image processing algorithms. Therefore, the Kmeans clustering (Sinaga & Yang, 2020) approach, which can automatically classify the image into different clusters, is applied to the factors. This clustering is done on a complete stack of all data including 44 bands.

The rationale for using this stack of 44 bands is to:

- **Facilitate Time-Series Analysis:** Clustering algorithms and anomaly detection methods often require time-series data to identify significant changes or anomalies over time. By stacking NDVI, MNDWI, LST, and SI for each year, the dataset provides a comprehensive temporal profile of environmental conditions in the wetland area.
- **Enhance Spatial and Temporal Resolution:** Combining multiple indices over a long period (11 years) ensures a robust dataset with sufficient temporal resolution. This allows for more detailed analysis and detection of anomalies related to wetland dynamics, such as drought events or changes in vegetation and water content.

After that, on each cluster, a histogram of the time series values of the factors was drawn to identify the probability of anomalous occurrence from the range of values and the intensity of its changes. In this case, the location of the anomalies occurring on the clusters is detected. After identifying the anomaly in terms of location, the time of occurrence of breakpoints will be defined from those histograms. Up to this stage, the approximate temporal and spatial ranges of the anomaly and drought events have been identified, and it is necessary to improve the results.

### 2.3.3. Training sample preparation

To prepare training samples automatically and without user intervention, it is not possible to rely only on the results of the clustering-based method. Therefore, to increase the accuracy of the results and also to ensure the correctness in extracting automatic samples, statistical methods were also used in this research. In this step, the Z-score of each factor in the Anomalous cluster and drought year was calculated. Z-score is Another criterion to show the significant differences in accuracy between the methods (Tamimi et al., 2017). It is used to compare each of the optimized methods with other basic methods to assess whether the other methods significantly differed in terms of accuracy or not. It is more appropriate as it also is more precise and sensitive (Mushore et al., 2017; Tamimi et al., 2017). McNamar's chi squared is computed as (DeVries et al., 2020):

$$Z = \frac{x - \mu}{\sigma} \quad (1)$$

- X is the value of the observation.
- μ is the mean of the population.
- σ is the standard deviation of the population.

The difference in accuracies was tested at a 95% significant level and deemed different if Z > 1.96 (Mushore et al., 2017; Tamimi et al., 2017). Next, all the Z-scores were aggregated to form a single image. The remaining values in the aggregated image will be used to extract samples of anomaly and non-anomaly classes for the classification step.

### 2.3.4. Drought map generation

In this research, the type of Random Forest described in (Breiman, 2001), which is one of the machine learning methods, was considered. The application of machine learning in surveying fields related to earth has a history of about 7 decades. Machine learning can be defined as the ability of computers to recognize patterns without being explicitly programmed (Alférez et al., 2022).

By using the training sample extracted automatically from the last step, Random Forest classification is implemented on Landsat 8 TOA Reflectance image for identified drought. Theoretical results (Breiman, 2001) demonstrate that Random Forest do not overfit when more trees are added. In addition, Random Forest produces high accuracy for many datasets; they can process data with a large number of features where each feature is weak, that is, carries a small amount of information; they are relatively robust to mixed variable types, missing data, outliers, and noisy data; constructing Random Forest is relatively fast (faster than bagging and boosting). In brief, a Random Forest is a classifier that consists of decision trees, each of which provides a vote for a certain class. Combining a large number of trees in a Random Forest can lead to more reliable predictions, while a single decision tree may overfit the data (Devetyarov & Nouretdinov, 2010).

### 2.3.5. Accuracy assessment

Validation plays an important role in the performance of different algorithms which confirms the accuracy of the proposed approach (Vazini Ahghar, 2023). In order to assess the proposed method, the classification map can be compared against the reference data (ground truth). The ground truth map of the region was prepared by the researcher from Landsat images at the time of classification and with the help of Google Earth images. Then, the F1_score, overall accuracy, and Kappa coefficient (De Leeuw et al., 2006) and some other assessment criteria can be calculated (Kiani et al., 2019).

## 3. Experimental Results

In this section, the output of different implementation parts has been presented. Subsections included multivariate factors, breakpoint detection, training sample preparation, drought map, and accuracy assessment.

### 3.1. Multivariate factors

First, the various factors estimated for each year are shown in Fig.5, 6, 7, and 8. Fig.5 shows the NDVI of each year.

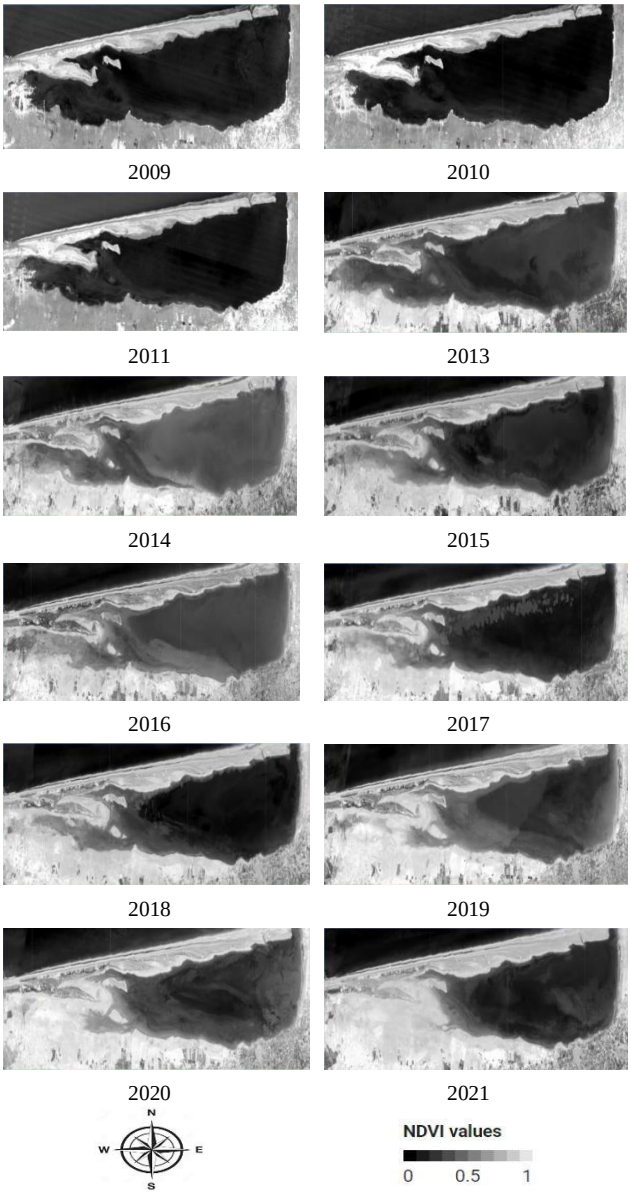



Figure 5. NDVI images of each year.

NDVI and its values always ranges from −1 to+1. Thus, the value of +1 indicates a high possibility of dense vegetation, while the values of NDVI close to Zero mostly refers to urbanization and water extent (Abu El-Magd et al., 2023). In the above pictures, the process of changes and the amount of water reduction can be seen to some extent. But this index has noises that alone can cause problems in clustering performance.

In the following, MNDWI images of each year are shown in Fig.6.

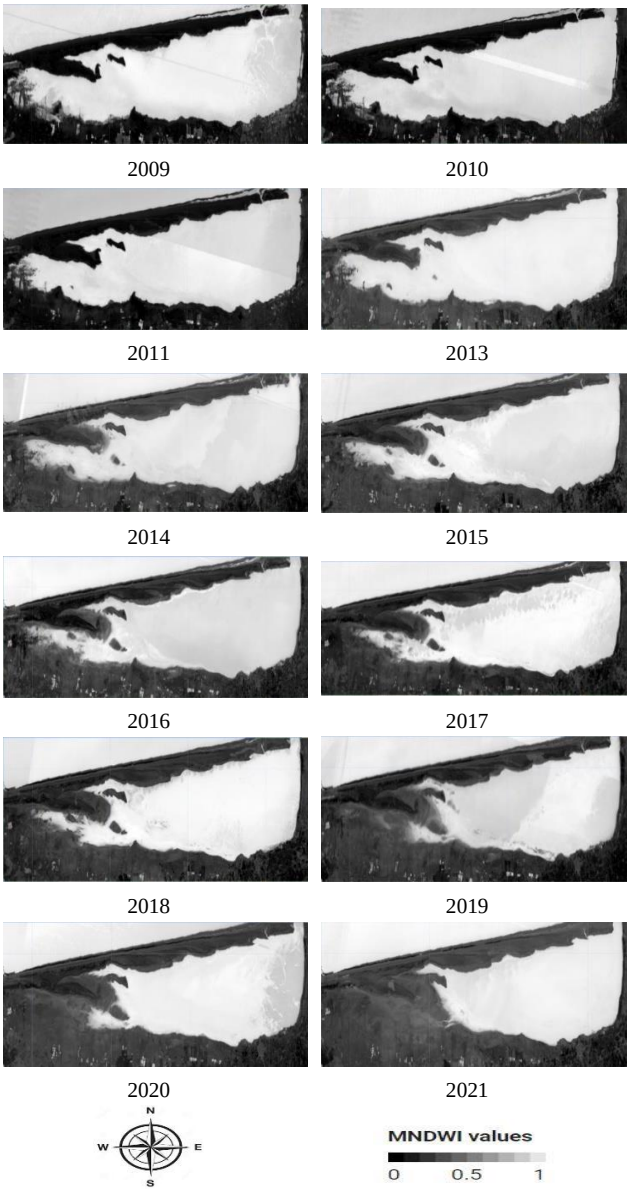
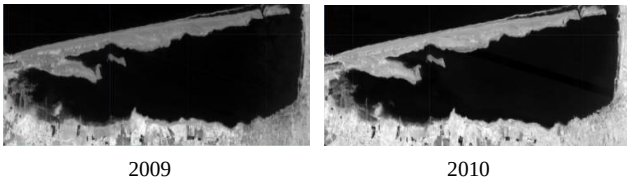



Figure 6. MNDWI images of each year.

MNDWI index has more capability in terms of efficiency in identifying water areas. But it seems that it will face problems in identifying other classes. For this purpose, the soil index, which has received less attention in other research, has been used. SI images of each year are shown in Fig.7.



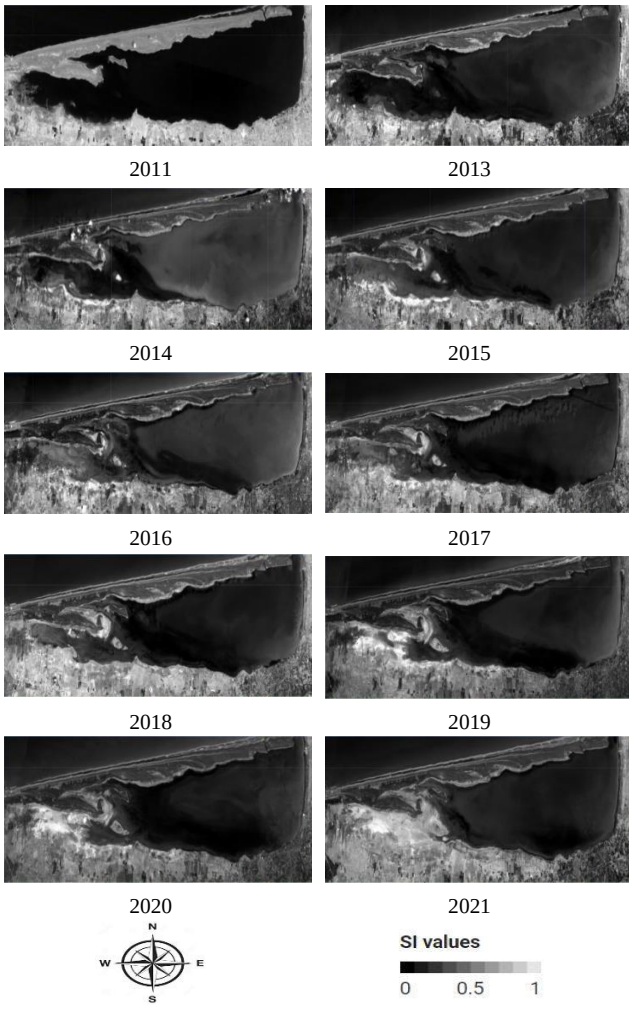
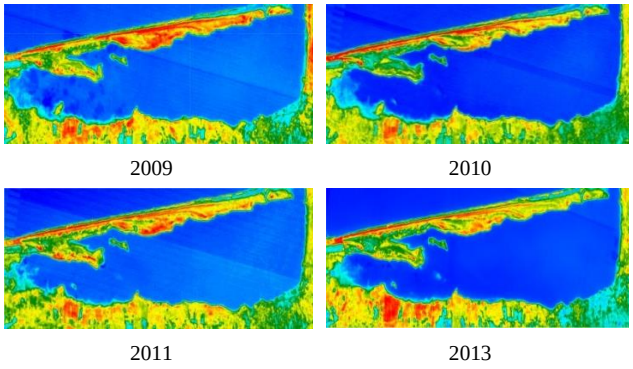



Figure 7. SI images of each year.

The SI index has good efficiency in separating and displaying the areas that suffered from drought. In addition, water areas seem to have a good identification capability.
In the following, the last factor, LST images of each year are shown in Fig.8.



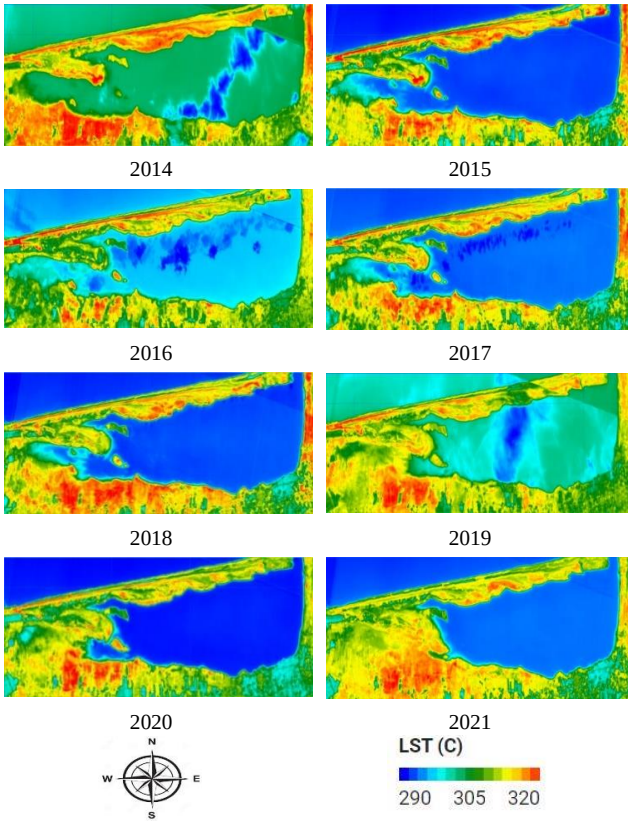



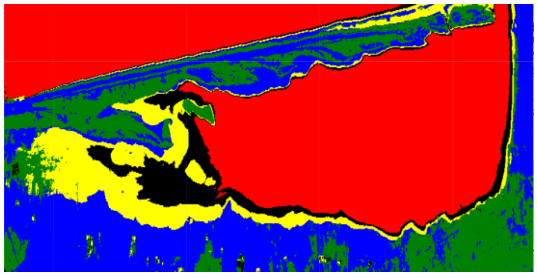

Figure 8. LST images of each year.

Land surface temperature is mainly used for determining the surface radiation and the energy exchange. It is also essential for determining the dynamics of the earth's surface, which impact-feedback loops that occur over a wide range of temporal and spatial scales (Tariq et al., 2020). Here, the decrease in precipitation and the increase in temperature will accelerate the process of drought in water areas. The temperature increases or decreases do not have a specific trend over time. But to be more precise, in the above time series, an increase in temperature can be seen in the drought-affected area. Therefore, this factor will also be very effective in identifying drought.

### 3.2. Breakpoints detection

Breakpoints are identified using a combination of clustering and image processing algorithms applied to a stack of 44 bands representing NDVI, MNDWI, LST, and SI for each year from 2009 to 2021 (excluding 2012). Clustering, specifically Kmeans clustering, is employed to classify the multi-dimensional data into clusters based on spectral similarities. These clusters help identify spatial anomalies or areas where significant changes occur. Subsequently, temporal breakpoints are determined by analyzing the temporal behavior of these clusters over the study period. Histograms of the time series values for each cluster are examined to pinpoint years with anomalous values indicative of potential drought events or other significant changes in wetland conditions.

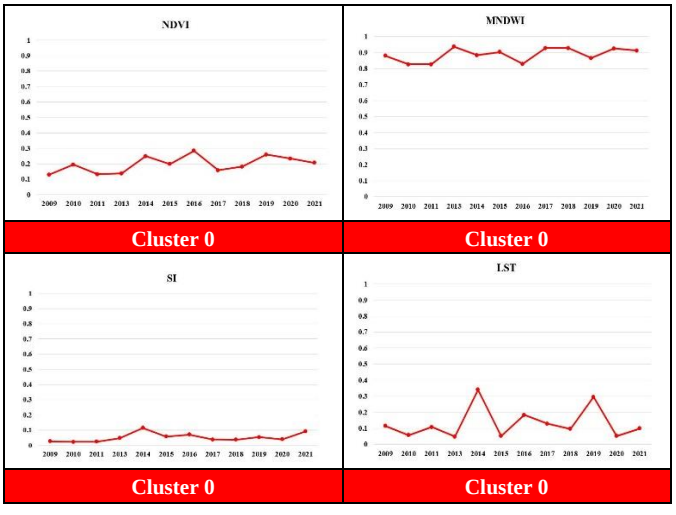
It should be noted that according to the environmental conditions, the number of land uses in the image, and Trial and error, the number of clusters was considered 5. Due to the disproportion of the values of the images, normalization was done before clustering to set the range of values between 0 and 1. Fig.9 shows the result of clustering.

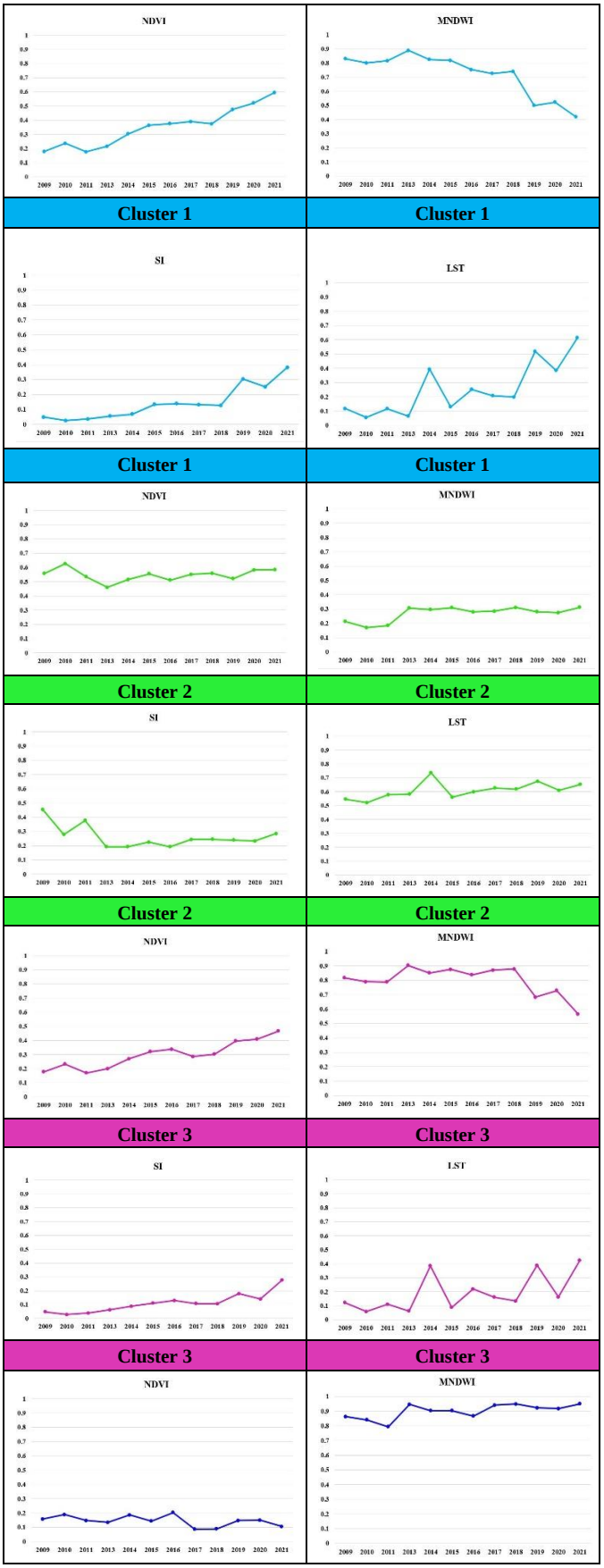


| Cluster | Name | Color |
|---|---|---|
| C0 | Agriculture | |
| C1 | Shallow water (low intensity) | |
| C2 | Deep water (high intensity) | |
| C3 | Barren land 1 | |
| C4 | Barren land 2 | |

Figure 9. Clustering result.

The purpose of implementing the clustering algorithm is to identify the anomalies spatially. In general, a cluster that has a low density can be identified as an anomalous cluster. Also, if the difference in the values of remote sensing factors is high in a cluster, then that cluster can be considered an anomaly. To check this, the histogram of each factor's values has been drawn for each cluster in the time series. A look at Fig.10 reveals the histograms of each cluster.



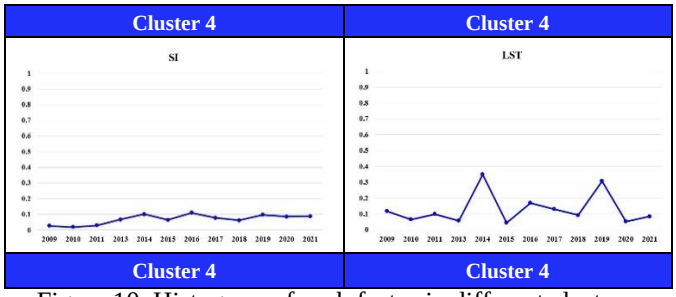
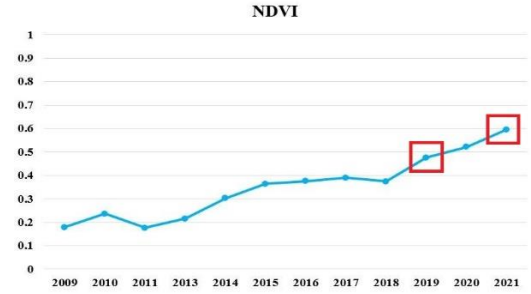
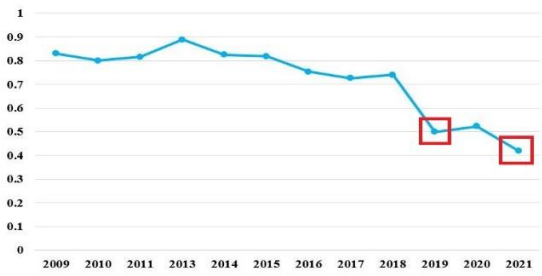



Figure 10. Histogram of each factor in different clusters.

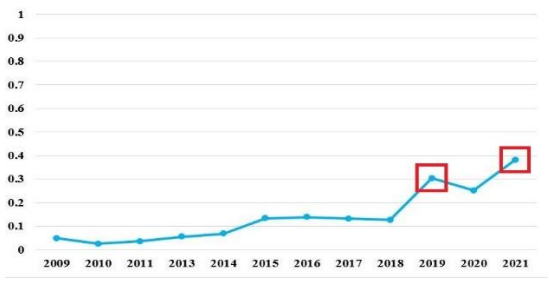
Based on the above histograms, the range of values for cluster 1 in all factors is much larger than the others, which increases the possibility of anomaly and detection of breakpoint location. In the following, the identical histograms of cluster 1 are used to identify the time break point.



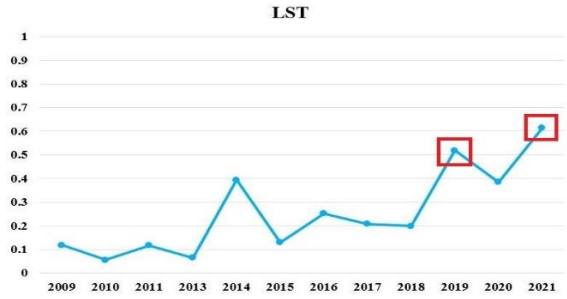

C1 – NDVI




C1 - MNDWI




C1 – SI




C1 - LST

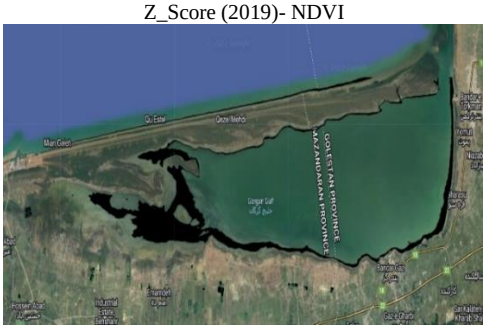
Figure 11. Detecting anomalous year.

As seen in the above figures, in 2019 and 2021, all the factors have undergone sudden changes. Therefore, until this stage, 2019 is recognized as the first year of drought occurrence in the time series.

### 3.3. Training sample preparation

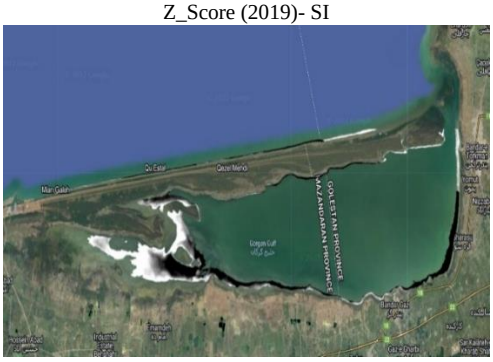
Next, the primary anomaly is calculated using the Z-Score technique to generate training data. This section displays the results calculated on cluster 1 in 2019 by equation 1.

Z_Score (2019)- NDVI



Z_Score (2019)- MNDWI


Z_Score (2019)- SI




Z_Score (2019)- LST


Figure 12. Z_score on each factor for 2019.

Regarding the above maps, the lighter or darker the color, the greater the intensity of the difference in 2019 from the time series. The light color indicates the positive direction, and the dark color indicates the negative direction of this difference. Finally, all the obtained Z_Scores are limited by a threshold and aggregated to produce a single image. The final result is presented in Fig.13.


| Class | Color |
|---|---|
| Aggregated Z_Score | |

Figure 13. Aggregated Z_Score.

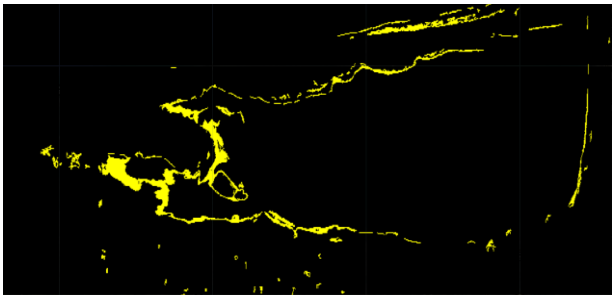
Since training samples are very effective in implementing supervised classification algorithms, the researcher selects training points at this stage on an aggregated Z_Score image. In this way, the sample selection process is semi-automatic but accurate. Here, the desired classes were divided into anomalous (drought) and non-anomalous (other). It should be noted that the number of training samples was considered very small to retain the automation level. Fig.14 shows these training samples.


| Sample's Class | Color |
|---|---|
| Anomaly (Drought) | |
| Normal (Non-drought) | |

Figure 14. Training samples on Z_Score (2019).

**3.4. Drought maps and accuracy assessment**

Random Forest classification was implemented to complete the anomaly detection process in the Miankale wetland using the semi-automatic hybrid method. The number of trees used in this classification was considered 100, but it is worth considering that changing this number to 150 and 200 did not affect the final classified image much. Therefore, their unnecessary and repetitive display was avoided. Since the applied classification was pixel-based, noise-like errors are common in the final map. Morphological operations are also performed to remove these noises and visually improve the final map. The final anomaly map is shown in Fig.15.



| Class | Color |
|---|---|
| Anomaly (Drought) | |
| Normal (Non-drought) | |

Figure 15. Final anomaly map.

Finally, the two drought maps obtained with the proposed algorithm were evaluated by the ground truth map in Table

2.

**Table 2. Accuracy Assessment**

| Map / Criteria | RF (100 trees) | RF (100 trees) with morphology |
|---|---|---|
| **Overall accuracy** | 97.75 | 97.89 |
| **Kappa coefficient** | 67.21 | 68.43 |
| **F1_score** | 98.84 | 98.91 |
| **Producer accuracy** | 97.75 | 97.89 |
| **User accuracy** | 97.72 | 97.82 |
| **Commission error** | 2.27 | 2.18 |
| **Omission error** | 2.24 | 2.11 |

As seen from the above table, the results are highly accurate for identification, which indicates the high efficiency of the proposed method in this research. Also, the low impact of the morphology algorithm on the accuracy of the initial map can be seen. As mentioned earlier, this operation was mainly due to the visual improvement of the final map to remove some noise from the pixel-based classifier.

## 4. Discussion

In this research, drought areas in one of Iran's most important wetlands were identified by analyzing time series images. The aim was to develop a simple, fast, and accurate method using a cloud processing system—Google Earth Engine—that enables swift and effortless image analysis of large datasets without the need for powerful hardware. A set of Landsat images from 2000 to 2021 during the summer season was selected for this purpose.

In this context, drought is defined as an unusual event occurring in the wetland's water zone. Anomaly detection methods, seldom used in wetland drought identification, were employed to detect these events. Using these methods, the time and place of anomalies/drought can be identified simultaneously, providing a comprehensive view of the subject.

Given that multiple factors contribute to the drought process, several indices were calculated annually: vegetation, water, soil salinity, and temperature. The selection of these indices was based on their direct impact on wetland health and drought conditions:

1. **Vegetation and NDVI**: Vegetation health, indicated by the Normalized Difference Vegetation Index (NDVI), is crucial because it reflects the overall ecological status of the wetland. Healthy vegetation usually implies adequate water availability and good soil conditions. Conversely, declining NDVI values may signal water stress or drought conditions.
2. **Soil Salinity Index**: Soil salinity is included because it significantly affects vegetation and water quality in wetlands. Research has indicated that areas with high soil organic matter and low soil salinity are more conducive to robust wetland distribution and health. High soil salinity can exacerbate drought conditions by reducing soil moisture retention and adversely impacting plant growth. Thus, monitoring soil salinity helps understand how drought conditions are evolving in the wetland's land and water zones.
3. **Water Indices**: Water indices directly measure the presence and extent of water bodies within the wetland, directly indicating drought when significant reductions are observed.
4. **Temperature Indices**: Temperature variations affect evaporation rates and water availability, which are critical in drought dynamics.

The integrated analysis of these indices through statistical and clustering-based anomaly detection algorithms allowed for creating a preliminary drought map of the wetland. This map was refined using a machine learning supervised classification algorithm, with training samples provided by user intervention to ensure high accuracy. The final map, evaluated against ground truth data, achieved an accuracy of 97.89% and an F1 score of 98.91%.

Despite the robust approach, there are limitations. The process remains semi-automatic, with user intervention needed for selecting training samples and identifying breakpoints. Future research should focus on fully automating these steps and incorporating related indices such as the Standardized Precipitation Evapotranspiration Index (SPEI) to improve anomaly identification. Additionally, using deep learning algorithms could further enhance the accuracy of the final drought map.

In conclusion, the combined use of various indices, including soil salinity and NDVI, provided a comprehensive assessment of drought conditions in the Miankale wetland, addressing land and water aspects crucial for understanding and managing wetland health under changing climatic conditions.

## 5. Conclusion

The importance of preserving and maintaining wetlands is always emphasized worldwide as one of the critical natural habitats. Due to the trend of temperature changes, global warming, and a decrease in precipitation, the drought in these areas is increasing rapidly. Therefore, it is essential to provide quick and accurate methods to identify sudden changes in land use, which can be referred to as anomalies. In this research, a hybrid method was proposed that combined simple and standard anomaly detection algorithms. In addition to simplicity and ease of implementation, it is highly accurate in presenting results. These methods included clustering-based algorithms, statistical methods with light calculations, and classification methods, which were used for quick implementation using the Google Earth Engine system. This system allowed us to use and process many satellite images for a long time series in a short time, which is essential for anomaly detection

techniques. Finally, the evaluation of the obtained map's accuracy shows the presented method's quality. However, like other research, there have been some challenges in this research that can be improved in future research. For example, although an attempt has been made to automate the executive process fully, this level of automation has decreased in determining the year of the anomaly in the time series. In future research, efforts will be made to automate this part. In addition, in the Z_score section, techniques should be implemented to improve the final aggregated image to increase the accuracy of this map and to automatically obtain training samples for classification. Lastly, using kernel-based methods can greatly improve the final map.

**Funding** No funding was received to assist with the preparation of this manuscript.

**Data availability** The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

## References

Abu El-Magd, S., Soliman, G., Morsy, M., & Kharbish, S. (2023). Environmental hazard assessment and monitoring for air pollution using machine learning and remote sensing. *International Journal of Environmental Science and Technology*, *20*(6), 6103-6116. https://doi.org/10.1007/s13762-022-04367-6

Ahmed, M., & Mahmood, A. N. (2014). Network traffic analysis based on collective anomaly detection. 2014 9th IEEE Conference on Industrial Electronics and Applications, https://doi.org/10.1109/ICIEA.2014.6931337

Ahmed, M., Mahmood, A. N., & Hu, J. (2016). A survey of network anomaly detection techniques. *Journal of Network and Computer Applications*, *60*, 19-31. https://doi.org/10.1016/j.jnca.2015.11.016

Aksoy, S., Yildirim, A., Gorji, T., Hamzehpour, N., Tanik, A., & Sertel, E. (2022). Assessing the performance of machine learning algorithms for soil salinity mapping in Google Earth Engine platform using Sentinel-2A and Landsat-8 OLI data. *Advances in Space Research*, *69*(2), 1072-1086. https://doi.org/10.1016/j.asr.2021.10.024

Alférez, G. H., Esteban, O. A., Clausen, B. L., & Ardila, A. M. M. (2022). Automated machine learning pipeline for geochemical analysis. *Earth Science Informatics*, *15*(3), 1683-1698. https://doi.org/10.1007/s12145-022-00821-8

Amani, M., Ghorbanian, A., Ahmadi, S. A., Kakooei, M., Moghimi, A., Mirmazloumi, S. M., Moghaddam, S. H. A., Mahdavi, S., Ghahremanloo, M., & Parsian, S. (2020). Google earth engine cloud computing platform for remote sensing big data applications: A comprehensive review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *13*, 5326-5350. https://doi.org/10.1109/JSTARS.2020.3021052

Amani, M., Mahdavi, S., Afshar, M., Brisco, B., Huang, W., Mohammad Javad Mirzadeh, S., White, L., Banks, S., Montgomery, J., & Hopkinson, C. (2019). Canadian wetland inventory using Google Earth Engine: The first map and preliminary results. *Remote Sensing*, *11*(7), 842. https://doi.org/10.3390/rs11070842

Barbier, E. B., Hacker, S. D., Kennedy, C., Koch, E. W., Stier, A. C., & Silliman, B. R. (2011). The value of estuarine and coastal ecosystem services. *Ecological monographs*, *81*(2), 169-193. https://doi.org/10.1890/10-1510.1

Bolca, M., Turkyilmaz, B., Kurucu, Y., Altinbas, U., Esetlili, M. T., & Gulgun, B. (2007). Determination of impact of urbanization on agricultural land and wetland land use in Balçovas' Delta by remote sensing and GIS technique. *Environmental monitoring and assessment*, *131*(1), 409-419. https://doi.org/10.1007/s10661-006-9486-0

Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5-32.

Bullock, E. L., Woodcock, C. E., & Olofsson, P. (2020). Monitoring tropical forest degradation using spectral unmixing and Landsat time series analysis. *Remote Sensing of Environment*, *238*, 110968. https://doi.org/10.1016/j.rse.2018.11.011

De Leeuw, J., Jia, H., Yang, L., Liu, X., Schmidt, K., & Skidmore, A. (2006). Comparing accuracy assessments to infer superiority of image classification methods. *International Journal of Remote Sensing*, *27*(1), 223-232. https://doi.org/10.1080/01431160500275762

Devetyarov, D., & Nouretdinov, I. (2010). Prediction with confidence based on a random forest classifier. IFIP international conference on artificial intelligence applications and innovations,

DeVries, B., Huang, C., Armston, J., Huang, W., Jones, J. W., & Lang, M. W. (2020). Rapid and robust monitoring of flood events using Sentinel-1 and Landsat data on the Google Earth Engine. *Remote Sensing of Environment*, *240*, 111664. https://doi.org/10.1016/j.rse.2020.111664

Ermida, S. L., Soares, P., Mantas, V., Göttsche, F.-M., & Trigo, I. F. (2020). Google Earth Engine Open-Source Code for Land Surface Temperature Estimation from the Landsat Series. *Remote Sensing*, *12*(9), 1471. https://www.mdpi.com/2072-4292/12/9/1471 https://doi.org/10.3390/rs12091471

Gholami, V., Khalili, A., Sahour, H., Khaleghi, M., & Tehrani, E. N. (2020). Assessment of environmental water requirement for rivers of the Miankaleh wetland drainage basin. *Applied Water Science*, *10*(11), 1-14. https://doi.org/10.1007/s13201-020-01319-8

Ghosh, S., Dinda, S., Chatterjee, N. D., & Das, K. (2018). Analyzing risk factors for shrinkage and transformation of East Kolkata Wetland, India. *Spatial Information Research*, *26*(6), 661-677. https://doi.org/10.1007/s41324-018-0212-0

Gong, P., Li, X., Wang, J., Bai, Y., Chen, B., Hu, T., Liu, X., Xu, B., Yang, J., & Zhang, W. (2020). Annual maps of

global artificial impervious area (GAIA) between 1985 and 2018. *Remote Sensing of Environment*, *236*, 111510. https://doi.org/10.1016/j.rse.2019.111510

Gong, P., Niu, Z., Cheng, X., Zhao, K., Zhou, D., Guo, J., Liang, L., Wang, X., Li, D., & Huang, H. (2010). China's wetland change (1990–2000) determined by remote sensing. *Science China Earth Sciences*, *53*(7), 1036-1042. https://doi.org/10.1007/s11430-010-4002-3

Hao, B., Ma, M., Li, S., Li, Q., Hao, D., Huang, J., Ge, Z., Yang, H., & Han, X. (2019). Land use change and climate variation in the three gorges reservoir catchment from 2000 to 2015 based on the Google Earth Engine. *Sensors*, *19*(9), 2118. https://doi.org/10.3390/s19092118

Hou, M., Ge, J., Gao, J., Meng, B., Li, Y., Yin, J., Liu, J., Feng, Q., & Liang, T. (2020). Ecological risk assessment and impact factor analysis of alpine wetland ecosystem based on LUCC and boosted regression tree on the Zoige Plateau, China. *Remote Sensing*, *12*(3), 368. https://doi.org/10.3390/rs12030368

Huang, Q., Long, D., Du, M., Zeng, C., Qiao, G., Li, X., Hou, A., & Hong, Y. (2018). Discharge estimation in high-mountain regions with improved methods using multisource remote sensing: A case study of the Upper Brahmaputra River. *Remote Sensing of Environment*, *219*, 115-134. https://doi.org/10.1016/j.rse.2018.10.008

Ivushkin, K., Bartholomeus, H., Bregt, A. K., Pulatov, A., Kempen, B., & De Sousa, L. (2019). Global mapping of soil salinity change. *Remote Sensing of Environment*, *231*, 111260. https://doi.org/10.1016/j.rse.2019.111260

Ji, C., Zhang, Y., Cheng, Q., Tsou, J., Jiang, T., & San Liang, X. (2018). Evaluating the impact of sea surface temperature (SST) on spatial distribution of chlorophyll-a concentration in the East China Sea. *International Journal of Applied Earth Observation and Geoinformation*, *68*, 252-261. https://doi.org/10.1016/j.jag.2018.01.020

Ji, L., Gong, P., Geng, X., & Zhao, Y. (2015). Improving the accuracy of the water surface cover type in the 30 m FROM-GLC product. *Remote Sensing*, *7*(10), 13507-13527. https://doi.org/10.3390/rs71013507

Jiang, Z., Huete, A. R., Chen, J., Chen, Y., Li, J., Yan, G., & Zhang, X. (2006). Analysis of NDVI and scaled difference vegetation index retrievals of vegetation fraction. *Remote Sensing of Environment*, *101*(3), 366-378. https://doi.org/10.1016/j.rse.2006.01.003

Jin, X., Qiang, H., Zhao, L., Jiang, S., Cui, N., Cao, Y., & Feng, Y. (2020). SPEI-based analysis of spatio-temporal variation characteristics for annual and seasonal drought in the Zoige Wetland, Southwest China from 1961 to 2016. *Theoretical and Applied Climatology*, *139*, 711-725. https://doi.org/10.1007/s00704-019-02981-y

Junk, W. J., An, S., Finlayson, C., Gopal, B., Květ, J., Mitchell, S. A., Mitsch, W. J., & Robarts, R. D. (2013). Current state of knowledge regarding the world's wetlands and their future under global climate change: a synthesis. *Aquatic sciences*, *75*(1), 151-167. https://doi.org/10.1007/s00027-012-0278-z

Kiani, A., Ebadi, H., & Farnood Ahmadi, F. (2019). Development of an object-based interpretive system based on weighted scoring method in a multi-scale manner. *ISPRS International Journal of Geo-Information*, *8*(9), 398. https://doi.org/10.3390/ijgi8090398

Li, B., Hu, Y., Chang, Y., Liu, M., Wang, W., Bu, R., Shi, S., & Qi, L. (2021). Analysis of the factors affecting the long-term distribution changes of wetlands in the Jing-Jin-Ji region, China. *Ecological Indicators*, *124*, 107413. https://doi.org/10.1016/j.ecolind.2021.107413

Li, F., Li, H., Lu, W., Zhang, G., & Kim, J.-C. (2019). Meteorological drought monitoring in Northeastern China using multiple indices. *Water*, *11*(1), 72. https://doi.org/10.3390/w11010072

Lin, Y., Yu, J., Cai, J., Sneeuw, N., & Li, F. (2018). Spatio-temporal analysis of wetland changes using a kernel extreme learning machine approach. *Remote Sensing*, *10*(7), 1129. https://doi.org/10.3390/rs10071129

Lv, J., Jiang, W., Wang, W., Wu, Z., Liu, Y., Wang, X., & Li, Z. (2019). Wetland loss identification and evaluation based on landscape and remote sensing indices in Xiong'an new area. *Remote Sensing*, *11*(23), 2834. https://doi.org/10.3390/rs11232834

McLaughlin, D. L., & Cohen, M. J. (2013). Realizing ecosystem services: wetland hydrologic function along a gradient of ecosystem condition. *Ecological Applications*, *23*(7), 1619-1631. https://doi.org/10.1890/12-1489.1

Meng, Y., He, Z., Liu, B., Chen, L., Lin, P., & Luo, W. (2020). Soil salinity and moisture control the processes of soil nitrification and denitrification in a riparian wetlands in an extremely arid regions in Northwestern China. *Water*, *12*(10), 2815. https://doi.org/10.3390/w12102815

Mushore, T. D., Mutanga, O., Odindi, J., & Dube, T. (2017). Assessing the potential of integrated Landsat 8 thermal bands, with the traditional reflective bands and derived vegetation indices in classifying urban landscapes. *Geocarto International*, *32*(8), 886-899. https://doi.org/10.1080/10106049.2016.1188168

Nicholls, R. J., Hoozemans, F. M., & Marchand, M. (1999). Increasing flood risk and wetland losses due to global sea-level rise: regional and global analyses. *Global Environmental Change*, *9*, S69-S87. https://doi.org/10.1016/S0959-3780(99)00019-9

Pham-Duc, B., Nguyen, H., Phan, H., & Tran-Anh, Q. (2023). Trends and applications of google earth engine in remote sensing and earth science research: a bibliometric analysis using scopus database. *Earth Science Informatics*, 1-17. https://doi.org/10.1007/s12145-023-01035-2

Sarkar, U. K., Das Ghosh, B., Puthiyottil, M., Das, A. K., Lianthuamluaia, L., Karnatak, G., Acharya, A., & Das, B. K. (2021). Spatio-temporal change analysis of three floodplain wetlands of eastern India in the context of

climatic anomaly for sustainable fisheries management. *Sustainable Water Resources Management*, *7*(3), 1-16. https://doi.org/10.1007/s40899-021-00529-5

Shelestov, A., Lavreniuk, M., Kussul, N., Novikov, A., & Skakun, S. (2017). Exploring Google Earth Engine platform for big data processing: Classification of multi-temporal satellite imagery for crop mapping. *frontiers in Earth Science*, *5*, 17. https://doi.org/10.3389/feart.2017.00017

Sinaga, K. P., & Yang, M.-S. (2020). Unsupervised K-means clustering algorithm. *IEEE Access*, *8*, 80716-80727.

Tamimi, E., Ebadi, H., & Kiani, A. (2017). Evaluation of different metaheuristic optimization algorithms in feature selection and parameter determination in SVM classification. *Arabian Journal of Geosciences*, *10*(22), 478. https://doi.org/10.1007/s12517-017-3254-z

Tariq, A., Riaz, I., Ahmad, Z., Yang, B., Amin, M., Kausar, R., Andleeb, S., Farooqi, M. A., & Rafiq, M. (2020). Land surface temperature relation with normalized satellite indices for the estimation of spatio-temporal trends in temperature among various land use land cover classes of an arid Potohar region using Landsat data. *Environmental Earth Sciences*, *79*, 1-15. https://doi.org/10.1007/s12665-019-8766-2

Teluguntla, P., Thenkabail, P. S., Oliphant, A., Xiong, J., Gumma, M. K., Congalton, R. G., Yadav, K., & Huete, A. (2018). A 30-m landsat-derived cropland extent product of Australia and China using random forest machine learning algorithm on Google Earth Engine cloud computing platform. *ISPRS Journal of Photogrammetry and Remote Sensing*, *144*, 325-340. https://doi.org/10.1016/j.isprsjprs.2018.07.017

Thakur, P. K., Garg, V., Kalura, P., Agrawal, B., Sharma, V., Mohapatra, M., Kalia, M., Aggarwal, S. P., Calmant, S., & Ghosh, S. (2021). Water level status of Indian reservoirs: A synoptic view from altimeter observations. *Advances in Space Research*, *68*(2), 619-640. https://doi.org/10.1016/j.asr.2020.06.015

Vazini Ahghar, E. S.-H., R.; Nazari, B.; Dodangeh, P.; Mousavi, S. . (2023). *Assessment of drought in agricultural areas by combining meteorological data and remote sensing data* in Proceedings of the 4th International Electronic Conference on Geosciences, 1–15 December 2022, MDPI: , Basel, Switzerland,. https://doi.org/10.3390/IECG2022-13960

Walter, M., & Mondal, P. (2019). A rapidly assessed wetland stress index (RAWSI) using Landsat 8 and Sentinel-1 radar data. *Remote Sensing*, *11*(21), 2549. https://doi.org/10.3390/rs11212549

Wang, S., Zhang, L., Zhang, H., Han, X., & Zhang, L. (2020). Spatial–temporal wetland landcover changes of poyang lake derived from Landsat and HJ-1A/B data in the dry season from 1973–2019. *Remote Sensing*, *12*(10), 1595. https://doi.org/10.3390/rs12101595

Werkmeister, C., Jacob, D. L., Cihacek, L., & Otte, M. L. (2018). Multi-element composition of prairie pothole wetland soils along depth profiles reflects past disturbance to a depth of at least one meter. *Wetlands*, *38*(6), 1245-1258. https://doi.org/10.1007/s13762-022-04367-6

Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing*, *27*(14), 3025-3033. https://doi.org/10.1080/01431160600589179

Yu, H., Li, L., Zhu, W., Piao, D., Cui, G., Kim, M., Jeon, S. W., & Lee, W.-K. (2019). Drought monitoring of the wetland in the Tumen River Basin between 1991 and 2016 using Landsat TM/ETM+. *International Journal of Remote Sensing*, *40*(4), 1445-1459. https://doi.org/10.1080/01431161.2018.1524604

Zhang, M., Lin, H., Long, X., & Cai, Y. (2021). Analyzing the spatiotemporal pattern and driving factors of wetland vegetation changes using 2000-2019 time-series Landsat data. *Science of the Total Environment*, *780*, 146615. https://doi.org/10.1016/j.scitotenv.2021.146615

Zhao, Q., Yu, L., Li, X., Peng, D., Zhang, Y., & Gong, P. (2021). Progress and trends in the application of Google Earth and Google Earth Engine. *Remote Sensing*, *13*(18), 3778. https://doi.org/10.3390/rs13183778

Zhou, Y., Dong, J., Xiao, X., Liu, R., Zou, Z., Zhao, G., & Ge, Q. (2019). Continuous monitoring of lake dynamics on the Mongolian Plateau using all available Landsat imagery and Google Earth Engine. *Science of the Total Environment*, *689*, 366-380. https://doi.org/10.1016/j.scitotenv.2019.06.341

Zurqani, H. A., Post, C. J., Mikhailova, E. A., Schlautman, M. A., & Sharp, J. L. (2018). Geospatial analysis of land use change in the Savannah River Basin using Google Earth Engine. *International Journal of Applied Earth Observation and Geoinformation*, *69*, 175-185. https://doi.org/10.1016/j.jag.2017.12.006