An Environmental Support System for Predicting the Normalized Difference Vegetation Index (NDVI) in the Green Mountain Region, Libya

Afaf Abdullah Aljibali
Higher Institute of Science and Technology (HISTM), Al Marj, Libya ,
efaf_abdullah@histm.edu.ly

Naima Abdel Aziz Abdel Hadi
Dept.of Computer Science Comp.Voc.Int.Inst.for Girls Ajdabya Libya ,
naemaabdulazizi@gmail.com

Zainab Abdullah Mohmmed
Faculty of Computing and Information Technology – University of Ajdabiya
zainabbuzenin@gmail.com

Doaa Jamal Ahmad
Faculty of Computing and Information Technology – University of Ajdabiya
doaajamal508@gmail.com

Danya Abdulhakim Amer
Faculty of Computing and Information Technology – University of Ajdabiya
doniasuleiman6@gmail.com

ABSTRACT

This study presents an intelligent environmental decision support system for monitoring vegetation cover and desertification indicators in Libya's Green Mountain region, by integrating remote sensing, GIS and machine learning techniques into a single platform. The study relied on seasonal data for the period 2000-2025 extracted from Landsat 5, Landsat 7, Landsat 8 and Sentinel-2 images, as well as climate data from ERA5-Land via the Google Earth Engine platform.

The XGBoost Regression algorithm was used to build a model that predicts the NDVI_mean value based on temperature, precipitation, humidity, dew point, and statistical NDVI characteristics, with performance assessed using MAE, RMSE, R^2, and correlation coefficient indicators. SHAP technology was employed to interpret model decisions and determine the impact of each variable on results, enhancing system transparency and reliability.

The results demonstrated the model's ability to perform well in predicting the mean values of the NDVI index, achieving a coefficient of determination (R^2 = 0.6629), a root mean square error (RMSE = 0.0791), a mean absolute error (MAE = 0.0621), and a correlation coefficient of 0.8259, confirming its potential for use in environmental decision-support systems, with the possibility of further developing the system in the future by updating the data and expanding its scope of application to include other regions within Libya.

الملخص .

تقدم هذه الدراسة نظام دعم قرار بيئي ذكي لمراقبة الغطاء النباتي ورصد علامات التصحر في منطقة الجبل الأخضر بليبيا من خلال دمج الاستشعار عن بعد ونظم المعلومات الجغرافية والتعلم الآلي في منصة موحدة. استخدمت الدراسة البيانات الموسمية للفترة 2000-2025 المستخلصة من صور الأقمار الصناعية Landsat 5 و Landsat 7 و Landsat 8 و Sentinel-2 ، بالإضافة إلى بيانات المناخ من ERA5-Land التي تم الحصول عليها من خلال محرك Google Earth .

تم استخدام XGBoost Regression لتطوير نموذج للتنبؤ NDVI_mean من درجة الحرارة والأمطار والرطوبة ونقطة الندى وميزات NDVI الإحصائية. قيم أداء النموذج باستخدام MAE و RMSE و R^2 ومعامل الارتباط. كما تم استخدام SHAP لتفسير تنبؤات النماذج وتحديد مساهمة كل متغير ، وبالتالي تعزيز شفافية النظام وموثوقيته.

أشارت النتائج الى تحقيق النموذج أداءاً تنبؤياً مقبولاً لـ NDVI_mean بمعامل تحديد R^2=0.6629

وخطأ مربع متوسط الجذر (RMSE) من 0.0791 ، وخطأ معياري متوسط (MAE) من 0.0621 ، ومعامل ارتباط من 0.8259. وتدعم هذه النتائج الاستخدام المحتمل للنموذج في نظم دعم القرارات البيئية. ويمكن تطوير النظام بشكل افضل عن طريق تحديث البيانات وتوسيع نطاقه ليشمل مناطق أخرى من ليبيا

Received: 26-07-2026 **Accepted: 18-08-2026** **Published: 02-09-2026**

Keywords: desertification, decision support, NDVI index, remote sensing, XGBoost, explainable artificial intelligence

الكلمات المفتاحية: التصحر، دعم القرار، NDVI، الاستشعار عن بعد، XGBoost، الذكاء الصناعي القابل للتفسير

Introduction

Desertification is one of the most significant natural phenomena threatening the sustainability of natural resources in arid and semi-arid regions, as it leads to land degradation, loss of soil fertility, and a decline in agricultural productivity, which negatively impacts food security and the stability of ecosystems. Given its location within arid and semi-arid climatic zones, Libya is among the countries most vulnerable to these risks

due to water scarcity, rising temperatures, sparse vegetation, and increasing human pressure on the land; this makes monitoring natural changes and analyzing the factors driving them a scientific necessity to support sustainable management strategies.

Rapid advances in remote sensing technologies and geographic information systems (GIS) have facilitated large-scale analyses of satellite imagery, making it possible to derive accurate environmental indicators for monitoring changes to the Earth's surface. In recent years, machine learning algorithms—most notably "XGBoost"—have emerged as effective tools for analyzing complex environmental datasets and uncovering nonlinear relationships between climatic and spectral variables, thereby enhancing the accuracy of predictive models. However, reliance on such models poses a challenge due to a lack of transparency regarding their decision-making mechanisms. This has led to the adoption of "explainable artificial intelligence" (XAI) techniques—most notably "SHAP" (SHapley Additive exPlanations)—which have proven effective in determining the quantitative contribution of each variable to the model's output, ultimately enhancing its overall reliability.

Although an increasing number of studies have examined the application of remote sensing and machine learning for vegetation cover analysis and land degradation monitoring, there remains a need to develop integrative frameworks that combine remote sensing data, climatic variables, advanced predictive models, and explainable artificial intelligence (XAI) methods. Moreover, deploying such an integrative approach in the Al Jabal Al Akhdar region of Libya can yield deeper insights and contribute to the development of scientifically robust tools applicable to monitoring land degradation indicators and supporting environmental decision-making.

Building on the foregoing, this study aims to develop an environmental decision support system for analyzing vegetation cover status and predicting mean NDVI (NDVI_mean) values in the Al Jabal Al Akhdar region of Libya. The proposed system relies on extracting environmental indicators from satellite data via the Google Earth Engine platform , employing an XGBoost Regression model to predict vegetation cover values, and utilizing the SHAP technique to interpret model outputs, thereby providing decision-makers with reliable scientific information to support sustainable natural resource planning and management. This integrative approach is

expected to enhance understanding of the drivers influencing vegetation cover dynamics and to strengthen the transparency of predictive models

2. Theoretical Framework and Literature Review

The field of desertification studies and land degradation assessment has undergone a fundamental methodological shift, driven by the integration of remote sensing and geographic information systems (GIS) on one hand, and machine learning algorithms alongside explainable artificial intelligence (XAI) on the other. This transformation has been particularly enabled by the Google Earth Engine platform, which allows researchers to process vast volumes of satellite imagery at a planetary scale without requiring heavy computational infrastructure . Two comprehensive reference reviews have documented this evolution: the first examined the temporal trajectory and global geographic distribution of remote sensing-based desertification studies since their inception in 1991 , while the second focused specifically on techniques for monitoring and tracking vegetation cover changes in arid and semi-arid environments . Meanwhile, AbdelRahman provided a more integrative review encompassing land degradation, desertification, and sustainable management through GIS and remote sensing applications.

2.1 Prior Studies in Libya

At the local Libyan level, the study of Alghariani et al. aimed to monitor and track vegetation cover changes in northwestern Libya, leveraging cloud computing via Google Earth Engine (GEE) to analyze time series of NDVI, NDBI, and land surface temperature indices. The results revealed sharp fluctuations in vegetation density associated with climatic variability and rapid urban sprawl; however, the study was limited to retrospective historical analysis. In the same local context, the doctoral thesis by Almesmari at Cranfield University examined changes in natural vegetation cover in the Al Jabal Al Akhdar region of northeastern Libya, employing remote sensing techniques supported by machine learning classification algorithms that demonstrated accurate discrimination among vegetation patterns. Nevertheless, this study remained confined to a static, descriptive-classificatory framework.

2.2 Regional Studies in North Africa

At the Maghreb regional level, a research team from Newcastle University developed an open-source machine learning methodology based on the Random Forest algorithm to detect historically desertified lands in the oases of southern Morocco using Sentinel-1/2 data, achieving a classification accuracy of 74–76% specifically for the desertification class. Nevertheless, the study was limited to immediate diagnosis (detection) and did not extend to predicting the future (forecasting). A subsequent study combined GEE analysis with a "Gradient Tree Boosting" (Gradient Tree Boosting), hydrogeological fieldwork, and socioeconomic surveys to track the dynamics of desertification in the Ternata Oasis in southeastern Morocco over a full four-decade period from 1984 to 2024. This represents a notable advance in integrative analytical frameworks, although it has remained limited to the Moroccan environment without calibration to other arid contexts in the Mediterranean region.

2.3 Quantitative Studies in Asia

At the level of broader Asian regional studies—which provide a rich methodological reference through advanced quantitative approaches, despite their geographical context differing from Libya—applications have been diverse: a study of the Tarim Basin employed LULC indices over three decades to classify five desertification severity levels ranging from light to very severe. A Central Asian study utilized a comprehensive Land Desertification Status Index (LDSI) via GEE to analyze trends and drivers for the period 2001–2020. Meanwhile, studies on the Mongolian Plateau relied on Intensity Analysis and the Geographical Detector to disentangle natural and anthropogenic drivers of desertification. This approach was extended to the Xilingol Basin to identify four drivers responsible for 99% of desertification dynamics, to the Qilian–Hami Basin via a C5.0 Decision Tree model, and to the Wuliangsu–Hai Basin using the Desertification Degree Index (DDI) based on Albedo–NDVI. The same DDI model was further applied at the scale of the entire Yellow River Basin Collectively, this extensive body of Asian studies shares a reliance on advanced quantitative tools (DDI, Geographical Detector, Intensity Analysis, XGBoost–SHAP) to disentangle desertification drivers with high spatiotemporal resolution. However, all such efforts have remained confined to Chinese, Mongolian, and Central Asian contexts, with no attempt to calibrate these advanced methodological frameworks for a Mediterranean-arid environment such as Libya's Al Jabal Al Akhdar, which exhibits radically different topographic and climatic characteristics.

2.4 Applications of Machine Learning and Explainable Artificial Intelligence

In the direction of deploying Explainable Artificial Intelligence (XAI), Chen and Guestrin established the algorithmic foundation for the XGBoost model, while Lundberg and Lee formulated the theoretical framework for the SHAP (SHapley Additive exPlanations) technique, which is grounded in game theory to interpret each variable's contribution to machine learning model predictions. XGBoost has demonstrated superior predictive accuracy compared to traditional logistic regression, and SHAP has proven capable of uncovering non-linear relationships between influencing factors and rangeland degradation, producing spatial maps that highlight risk zones. However, these applications have been confined to the rangeland environment of Inner Mongolia and have not addressed Mediterranean or arid environments such as Al Jabal Al Akhdar.

A similar ensemble learning approach was employed in another study for remote modeling of desertification risk, with results consistently identifying precipitation as the most influential factor . Specifically in the Iraqi context, Al-Tameemi et al. presented a study that integrated an enhanced MEDALUS model with entropy weighting and the XGBoost algorithm, alongside SHAP interpretation to determine the relative importance of factors influencing desertification, achieving a high classification accuracy of 94.46% and a Kappa coefficient of 0.926, and revealing that the climatic quality and vegetation quality indices were the most influential. Nevertheless, this work remained limited to the Iraqi environment within the MEDALUS framework, necessitating testing of the same methodology in distinct Mediterranean environments such as Libya's Al Jabal Al Akhdar region

The present study seeks to address the limitations of existing application frameworks that integrate remote sensing data, climatic variables, cloud computing, machine learning-based predictive models, and explainable AI techniques within a single model dedicated to studying vegetation cover dynamics in Mediterranean and arid environments in Libya. This framework is expected to facilitate a transition from descriptive monitoring of environmental changes to interpretable predictive analysis, thereby enhancing the scientific value of the results and their utility in supporting land degradation and desertification monitoring.

3. Methodology

3.1 Overall Methodological Approach

The study adopts an applied-analytical research approach that integrates remote sensing techniques, reanalyzed climatic data, and machine learning algorithms, with the aim of addressing the research questions and testing the hypotheses concerning the capability of the XGBoost Regression model to predict mean NDVI (NDVI_mean) values in the Al Jabal Al Akhdar region, and the capacity of the SHAP technique to interpret the model's outputs transparently.

3.2 Study Area

The Al Jabal Al Akhdar region in northeastern Libya was selected as the geographical scope of this study, owing to its notable vegetation diversity and its concurrent exposure to varied environmental and climatic pressures, rendering it a representative model for investigating vegetation cover changes and monitoring manifestations of desertification using remote sensing techniques . The boundaries of the study area were delineated within the Google Earth Engine platform using fixed spatial boundaries to ensure consistency across all processing and data extraction operations throughout the study period spanning from 2000 to 2025.

Location map of the Al-Jabal Al-Akhdar region defined in the study, as shown in Figure 3-1.



Code Snippet 3.1: Defining the study area and time period in GEE

```
var roi = ee.Geometry.Rectangle([20.9, 32.30, 22.5, 32.90]);
Map.centerObject(roi, 9);
Map.addLayer( roi,{color: 'red'},'Study Area');
var START_YEAR = 2000;
var END_YEAR = 2025;
var SCALE = 30;
```

Remote sensing provides a time series of Earth's surface spectral measurements, enabling the derivation of vegetation indices and the comparison of their conditions across different periods. The study utilized Landsat 5, Landsat 7, Landsat 8, and Sentinel-2 within the GEE environment.

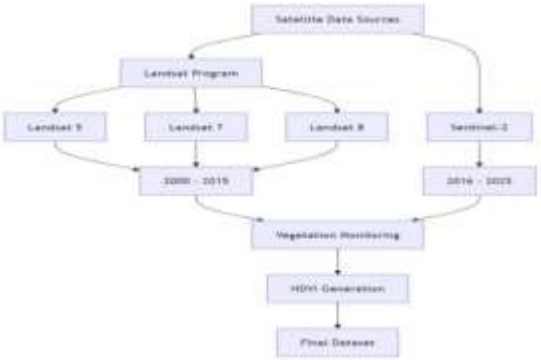
3.3 Data Sources and Collection

The study relied on two primary data sources, both automatically acquired via the Google Earth Engine cloud computing platform :

Table 3.1: Data sources used in the study and the function of each source

Data Type	Source	Usage / Application
Satellite	Landsat 5 / 7 / 8 – Collection 2 Level 2	Deriving NDVI and constructing the time series
Satellite	COPERNICUS/S2_SR_HARMONIZED	Deriving NDVI
Climatic	ECMWF/ERA5_LAND/MONTHLY_AGGR	Extracting Temperature, Dew Point, and Rainfall, and calculating Humidity
Auxiliary	JRC/GSW1_4/GlobalSurfaceWater	Excluding / Masking out water bodies

A diagram illustrating the satellite imagery sources used in the study, as shown in Figure 3-2.



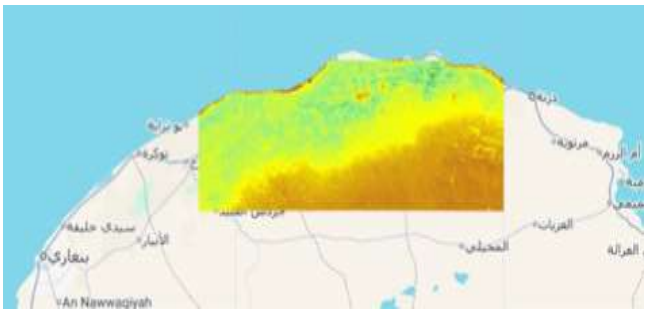
a.Satellite (Spectral) Data: Landsat satellite imagery (Landsat 5, 7, and 8) was utilized for the period 2000–2015, and Sentinel-2 imagery for the period 2015–2025, drawn from the Landsat Collection 2 and Sentinel-2 Harmonized datasets available through the GEE Data Catalog

Code Snippet 3.2: Importing image collections and the water database

```
var landsat5 = ee.ImageCollection( 'LANDSAT/LT05/C02/T1_L2');
var landsat7 = ee.ImageCollection( 'LANDSAT/LE07/C02/T1_L2');
var landsat8 = ee.ImageCollection( 'LANDSAT/LC08/C02/T1_L2');
var sentinel2 = ee.ImageCollection('COPERNICUS/S2_SR_HARMONIZED');
var waterDataset = ee.Image('JRC/GSW1_4/GlobalSurfaceWater');
var waterMask = waterDataset .select('occurrence').lt(50);
```

b. Climatic Data: Climatic variables were extracted from the ERA5-Land Monthly Aggregated global reanalysis dataset, produced by the European Centre for Medium-Range Weather Forecasts (ECMWF) under the Copernicus Climate Change Service , which provides an enhanced representation of the land surface component at a spatial resolution of up to 9 km. The dataset included temperature, dew point, and precipitation variables, while relative humidity was mathematically derived based on temperature and dew point values.

A screenshot from the GEE platform showing the ERA5-Land data used in the study, as shown in Figure 3-3

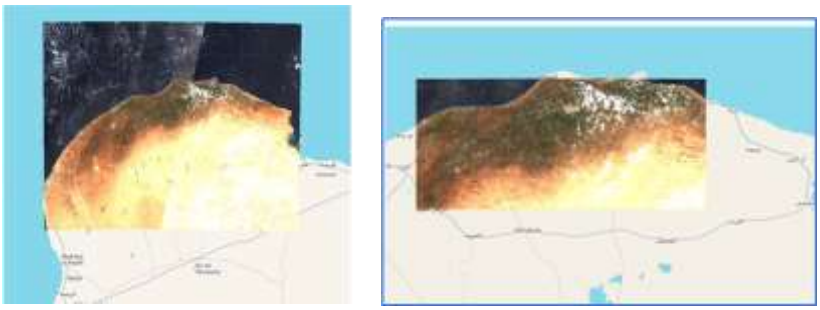


3.4 Data Preprocessing

The satellite imagery underwent a series of preprocessing stages to ensure the quality and consistency of spectral data:

1. **Geographic Clipping:** All images were clipped according to the study area boundaries, with the aim of constraining the analysis to the targeted spatial extent and minimizing the processing of data located outside the region.

Figure 3-4: An example of an image before and after the cropping process

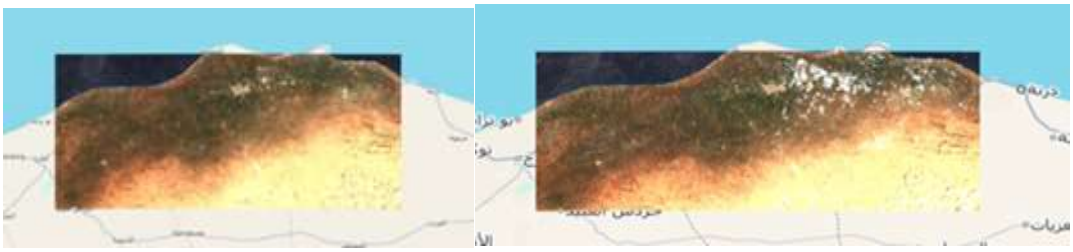


Code Snippet 3.3: Filtering images and clipping the final composite to the ROI

```
var landsat5 = ee.ImageCollection(
 'LANDSAT/LT05/C02/T1_L2'
);
var landsat7 = ee.ImageCollection(
 'LANDSAT/LE07/C02/T1_L2'
);
var landsat8 = ee.ImageCollection(
 'LANDSAT/LC08/C02/T1_L2'
);
var sentinel2 = ee.ImageCollection(
 'COPERNICUS/S2_SR_HARMONIZED'
);
var waterDataset = ee.Image(
 'JRC/GSW1_4/GlobalSurfaceWater'
);
var waterMask = waterDataset
 .select('occurrence')
 .lt(50);
```

2. **Cloud Removal:** Clouds were removed using the Scene Classification (SCL) layer for Sentinel-2 imagery and the Pixel Quality (QA_PIXEL) layer for Landsat imagery.

Figure 3-5: An example of an image before and after cloud removal



Code Snippet 3.4: Applying quality masks to Landsat and Sentinel-2

```
function maskLandsat(image) {
 var qa = image.select('QA_PIXEL');

 var mask = qa.bitwiseAnd(1 << 1).eq(0)
  .and(qa.bitwiseAnd(1 << 2).eq(0))
  .and(qa.bitwiseAnd(1 << 3).eq(0))
  .and(qa.bitwiseAnd(1 << 4).eq(0))
  .and(qa.bitwiseAnd(1 << 5).eq(0));
 return image .updateMask(mask).updateMask(waterMask).copyProperties( image,['system:time_start']
);
}

function maskSentinel2(image) {
 var scl = image.select('SCL');
 var mask = scl.neq(3)
  .and(scl.neq(8))
  .and(scl.neq(9))
  .and(scl.neq(10))
  .and(scl.neq(11));
 return image
  .updateMask(mask)
  .updateMask(waterMask)
  .copyProperties(
   image,
   ['system:time_start']
  );
}
```

Figure 3.6: A practical example of applying quality masks to satellite imagery.



3. **Water Body Masking:** Water bodies were excluded using the Global Surface Water database

Code Snippet 3.5: Creating and applying the water mask

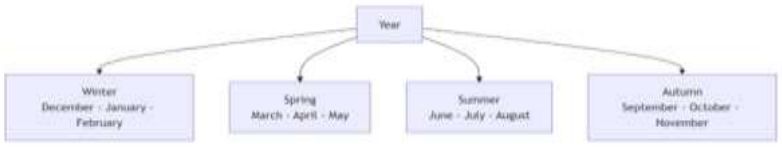
```
var waterOccurrence =
 waterDataset.select('occurrence');
var waterMask =
 waterOccurrence.lt(50);
var maskedImage =
 image.updateMask(waterMask);
```

Figure 3-7:A map illustrating the water mask used, as shown in



4. **Seasonal Compositing:** Each year was divided into four seasons, with a single composite image generated per season by computing the statistical median (Median Composite).

Figure 3.8: Dividing the year into four seasons and generating the seasonal median composite.



Code Snippet 3.6: Generating the seasonal composite

```
function getNDVISeasonalComposite(startDate,endDate) {
 var collection = satelliteNDVI.filterBounds(roi).filterDate( startDate, endDate );
 return collection .median()
  .clip(roi)
  .resample('bilinear');
}
```

5. **Spatial Resolution Harmonization:** All images were resampled to a uniform spatial resolution of 30 meters per pixel.

3.5 Extraction of NDVI Index and Statistical Properties

The Normalized Difference Vegetation Index (NDVI) was computed according to the standard mathematical formula:

$$\mathrm{NDVI} = \frac{\mathrm{NIR} - \mathrm{Red}}{\mathrm{NIR} + \mathrm{Red}}$$

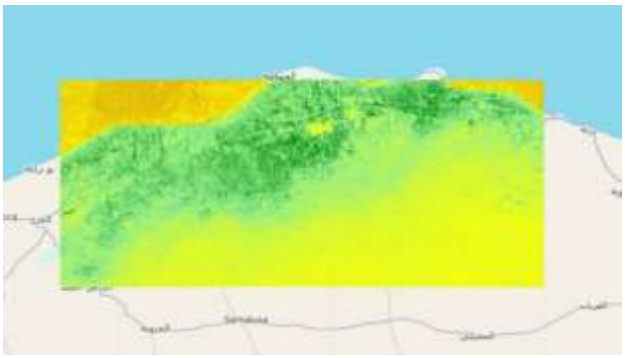
where NIR denotes spectral reflectance in the near-infrared band and Red denotes reflectance in the red band, with index values ranging from −1 to +1. For each seasonal composite image, four spatial statistical properties value (NDVI_min), were extracted: mean (NDVI_mean), standard deviation (NDVI_std), minimum and maximum value (NDVI_max).

Code Snippet 3.7: Calculating NDVI based on the spectral bands for each sensor

```
function addNDVI_L5(image) {
 var ndvi = image
  .normalizedDifference([
   'SR_B4', 'SR_B3'
  ])
  .rename('NDVI');
 return image.addBands(ndvi);
}
function addNDVI_L8(image) {
 var ndvi = image
  .normalizedDifference([
   'SR_B5', 'SR_B4'
  ])
  .rename('NDVI');
 return image.addBands(ndvi);
}
function addNDVI_S2(image) {
 var ndvi = image
  .normalizedDifference([
   'B8', 'B4'
```

```
  ])
  .rename('NDVI');
 return image.addBands(ndvi);
}
```

Figure 3.9: An actual NDVI map for one of the seasonal composites within the study area.



Code Snippet 3.8: Extracting spatial statistics

```
function getNDVIStatistics(composite) {
 return composite.reduceRegion({
  reducer: ee.Reducer.mean()
   .combine({
    reducer2:
     ee.Reducer.stdDev(),
    sharedInputs: true
   })
   .combine({
    reducer2:
     ee.Reducer.minMax(),
    sharedInputs: true
   }),
  geometry: roi,
  scale: 30,
  maxPixels: 1e13,
  bestEffort: true
 });
}
```

3.6 Final Dataset Construction

The spectral properties were integrated with climatic variables matched temporally and spatially, yielding a unified dataset comprising 104 seasonal statistical records. This dataset resulted from a temporal span covering 26 years (2000–2025), at a rate of four seasonal records per year.

Table 3.2 . Variables in the Final Dataset

S_no	Variable	Symbol	Unit	Overall Mean (Approx.)
1	Year	Year	—	2025–2000
2	Season	Season	—	Four seasons
3	Temperature	Temperature	°C	17.81
4	Dew Point	DewPoint	°C	8.42
5	Precipitation	Rainfall	mm	17.31
6	Relative Humidity	Humidity	%	59.6
7	Mean NDVI	NDVI_mean	—	0.0818
8	NDVI Standard Deviation	NDVI_std	—	0.0421
9	NDVI Minimum Value	NDVI_min	—	−0.112
10	NDVI Maximum Value	NDVI_max	—	0.387

Figure 3.10: An actual snippet from the final dataset

	A	B	C	D	E	F	G	H	I	J
1	DewPoint	Rainfall	Season	Temperature	Year	NDVI_mean	NDVI_std	NDVI_min	NDVI_max	Humidity
2	5.375888828	19.88952887	Winter	11.14849302	2000	-0.05187807714	0.208672472	-0.9106823802	0.2820314765	87.59597915
3	8.187470588	13.41572189	Spring	16.23280334	2000	-0.06580423191	0.1591021217	-0.732811537	0.5013244748	51.39100943
4	13.58113171	0.283893819	Summer	23.35866556	2000	0.100033689	0.08891676338	-0.9512330174	0.5473712887	54.19158354
5	10.21873189	10.92588543	Autumn	18.23081369	2000	0.1127343024	0.09218688533	-0.8307687163	0.247206375	55.8861991
6	5.788817415	58.69393242	Winter	9.734968021	2001	-0.1530123484	0.2012641558	-0.8268254995	0.2102867812	78.4822211
7	8.230244803	8.488054151	Spring	17.28725821	2001	0.02988852284	0.1167143386	-0.8317956229	0.2283094525	49.19644994
8	13.76880512	0.3277277945	Summer	23.81111868	2001	0.1188147574	0.09979573287	-0.7038726871	0.5634258389	53.37414306
9	11.6863402	15.66436984	Autumn	18.27670819	2001	0.1277511653	0.0848376581	-0.6813516168	0.2743191719	61.46741099
10	5.25830283	39.30537484	Winter	10.4416815	2002	-0.1010062282	0.2011692564	-0.3611646805	0.2516388543	70.27775372

3.6.1 Extraction of climate variables and calculation of Humidity

code 3.9: Extract Temperature, DewPoint and Rainfall from ERA5-Land.

```
var era5 = ee.ImageCollection(
 'ECMWF/ERA5_LAND/MONTHLY_AGGR'
).filterBounds(roi);

function getClimateData(
 startDate,
 endDate
) {
 var climate =
  era5.filterDate(
   startDate,
   endDate
  );
 var temperature =
  climate
   .select('temperature_2m')
   .mean()
   .reduceRegion({
    reducer: ee.Reducer.mean(),
    geometry: roi,
    scale: 11132,
    maxPixels: 1e13
   })
```

```
    .get('temperature_2m');
  var dewPoint =
   climate
    .select('dewpoint_temperature_2m')
    .mean()
    .reduceRegion({
     reducer: ee.Reducer.mean(),
     geometry: roi,
     scale: 11132,
     maxPixels: 1e13
    })
    .get('dewpoint_temperature_2m');
  var rainfall =
   climate
    .select('total_precipitation_sum')
    .sum()
    .reduceRegion({
     reducer: ee.Reducer.mean(),
     geometry: roi,
     scale: 11132,
     maxPixels: 1e13
    })
    .get('total_precipitation_sum');
  return {
   temperature: temperature,
   dewPoint: dewPoint,
   rainfall: rainfall
  };
}
```

3.7 Correlation Analysis Between Variables

Pearson Correlation analysis was conducted among the ten variables to test the third research hypothesis, which posits that the statistical properties of the NDVI index and the climatic variables contribute to improving prediction quality.

Heatmap of correlation coefficients, as shown in Figure 3-11

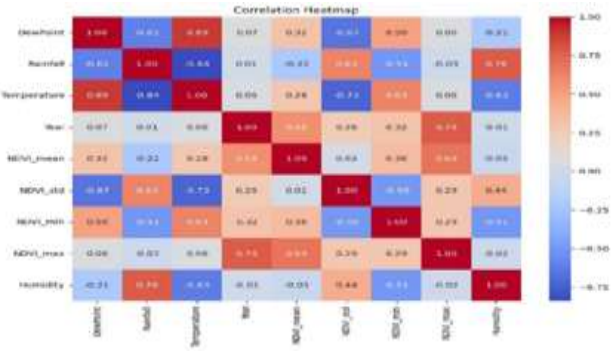


Table 3.3: Summary of Pearson correlation analysis results with NDVI_mean

Relationship	Correlation Coefficient (r)	Direction	Environmental Interpretation
NDVI_mean ↔ NDVI_max	0.689	Strong Positive	Dynamic consistency between mean vegetation density and its seasonal peak
NDVI_mean ↔ Year (Trend)	0.576	Moderate Positive	General environmental recovery trend of vegetation over 26 years
NDVI_mean ↔ Rainfall	Negative	Inverse (Negative)	Time-lag effect between winter rainfall and spring peak growth
NDVI_mean ↔ Temperature / DewPoint	Positive	Direct (Positive)	Stimulation of photosynthetic activity in the temperate Mediterranean climate

Code Snippet 3.10: Calculating Humidity based on Temperature and DewPoint

```
function calculateHumidity(
 temperature,
```

```
 dewPoint
) {
 var a = 17.625;
 var b = 243.04;
 var tempC =
  ee.Number(temperature)
    .subtract(273.15);
 var dewC =
  ee.Number(dewPoint)
    .subtract(273.15);
 var numerator =
  ee.Number(a)
    .multiply(dewC)
    .divide(
      ee.Number(b).add(dewC)
    );

 var denominator =
  ee.Number(a)
    .multiply(tempC)
    .divide(
      ee.Number(b).add(tempC)
    );
 return ee.Number(100)
  .multiply(
    numerator
      .subtract(denominator)
      .exp()
  );
}

}
```

3.8 Feature Engineering and Target Variable

The overall mean temperature was 17.81 °C, the mean precipitation was 17.31 mm, and the mean humidity was 59.09%, while the overall mean of NDVI_mean was approximately 0.0819. These values represent a statistical description of the final dataset.Heatmap of correlation coefficients, as shown in Figure 3-9

Table 3.4: Overall means of the key dataset variables

Variable	Overall Mean	Unit
Temperature	17.81	°C
Rainfall	17.31	mm
Humidity	52.6	%
NDVI_mean	0.0818	Dimensionless

Figure 3.12: Descriptive statistics extracted from the dataset

	A	B	C	D	E	F
1		DewPoint	Rainfall	Temperature	Year	Humidity
2	count	104	104	104	104	104
3	mean	9.519777686	17.30777585	17.80863083	2012.5	59.08523313
4	std	3.813729729	15.90440992	4.988173198	7.536319825	8.606543035
5	min	4.197225486	0.258699C896	9.371579975	2000	38.95290968
6	25%	5.976516441	4.479471137	14.2729	2006	52.66746502
7	50%	8.884319769	13.51426496	18.38095073	2012.5	57.49207178
8	75%	13.13396436	24.62955775	21.04446921	2019	65.41336
9	max	15.47759432	68.18400593	26.87406435	2025	76.70793781

3.9 Modeling Algorithm Selection

The XGBoost Regression algorithm—a tree-based boosted regression method —was adopted for its proven capability to handle non-linear relationships, its efficiency with medium-sized datasets, and its inclusion of regularization parameters that mitigate overfitting.

The trained model was saved in binary format (.pkl) using the Joblib library, to ensure it can be directly invoked within the proposed system without requiring retraining.

Table 3.5: Dataset Splitting (Hold-Out Validation)

Dataset / Subset	Percentage	Number of Records
Training Set	80%	83 records
Test Set (Testing Set)	20%	21 records
Total	100%	104 records

3.11 Hyperparameter Tuning

Table 3.6: Final Hyperparameters for the XGBoost Model

Hyperparameter	Value	Description / Function
n_estimators	300	Number of sequential trees
learning_rate	0.05	Boosting step size / Shrinkage rate
max_depth	4	Maximum depth of a tree

The trained model was saved in a binary format (.pkl) using the Joblib library to ensure it can be loaded directly into the proposed system without the need for retraining.

4: Discussion of results

The results indicate that the model achieved excellent performance on the training data, while maintaining good and reliable performance on independent test data ($R^2 \approx 0.66$, Correlation ≈ 0.83), thereby confirming its generalization capability and supporting the validity of the second research hypothesis.

The SHAP technique was employed to interpret each variable's contribution with greater transparency. The results demonstrated that features derived from the NDVI index—particularly NDVI_max—constituted the dominant factor influencing prediction accuracy, followed by temperature, whereas the effects of precipitation and humidity were relatively limited. This outcome is attributed to the Time-Lag Effect, as elucidated in Section 3.7, thereby supporting the validity of the fourth research hypothesis

4.1 Linking Results to Research Questions and Hypotheses

The derived results collectively support all five research hypotheses: the integration of satellite and climatic data enhanced the model's predictive capability (Hypothesis 1); XGBoost demonstrated efficiency in learning non-linear relationships (Hypothesis 2); the statistical properties of the NDVI index confirmed their pivotal role (Hypothesis 3); the SHAP technique successfully interpreted the model and increased its transparency(Hypothesis 4),and these findings pave the way for deploying the results within the proposed environmental decision support system (Hypothesis 5).

5. Results

This section presents the findings obtained through the successive stages of the research, beginning with the exploratory statistical analysis of interrelationships among environmental and climatic variables and the vegetation cover index (NDVI), followed by the seasonal analysis results for these variables, and concluding with the training and evaluation outcomes of the proposed intelligent model based on the XGBoost algorithm, alongside the results of interpreting its decisions using the SHAP technique. Collectively, these findings constitute the empirical foundation upon which the proposed environmental decision support system was subsequently built

5.1 Exploratory Statistical Analysis and Correlation Results

Computation of the Pearson Correlation Matrix among hydroclimatic variables and the NDVI index revealed several statistically significant patterns with clear environmental implications. The correlation between the arithmetic mean of the vegetation cover index (NDVI_mean) and the maximum value of the same index (NDVI_max) recorded a high value of 0.689, reflecting a dynamic consistency between the improvement of general environmental conditions and the peak vegetation density observed within the same season.

A positive correlation was also observed between mean NDVI and both near-surface air temperature and dew point, a relationship that can be interpreted in light of the Mediterranean climatic nature of the Al Jabal Al Akhdar region. Conversely, the analysis revealed a negative correlation between cumulative seasonal

precipitation and mean NDVI—a result that appears counterintuitive from an ecological standpoint but is statistically attributed to the presence of a Time-Lag Effect in the environmental response

At the extended temporal scale, the analysis revealed a positive and notable trend between the progression of years and mean NDVI, with a correlation coefficient of 0.576 over the 26-year period (2000–2025). This suggests that the region has experienced periods of ecological recovery and resilience in its vegetation cover, or that growing seasons have shifted toward an overall positive trend as a result of climatic changes. This observation will be revisited and dissected with greater precision later through the analysis of feature weights using the SHAP technique.

5.2 Seasonal Analysis Results of Environmental Variables

Analysis of the seasonal means of climatic and biotic variables over the 26-year time series revealed clear seasonal variability reflective of the prevailing climatic regime in the Al Jabal Al Akhdar region. The highest cumulative precipitation rates coincided with the highest relative humidity levels during winter, consistent with the traditional climatic pattern of the eastern Mediterranean, where rain-bearing low-pressure systems are concentrated in this season. Conversely, the highest surface temperature rates were concentrated in summer, accompanied by a sharp decline in precipitation and atmospheric humidity, representing the peak of thermal climatic stress on the region's ecosystem.

The most significant finding, however, was the recording of the highest annual mean NDVI during summer—not during the rainy winter season, as might be expected.

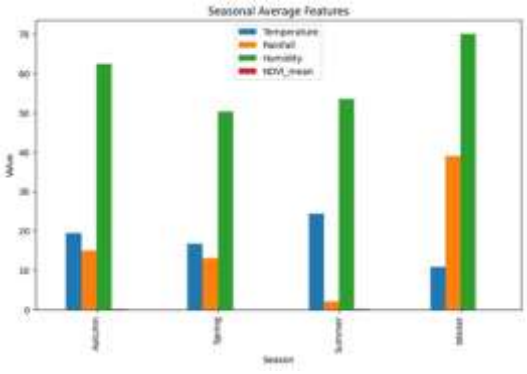


Figure 5.1: Chart of mean environmental-variable values by season.

6. Model Development

Following the completion of dataset preparation, the feature set used in the training process was constructed, comprising the following variables:

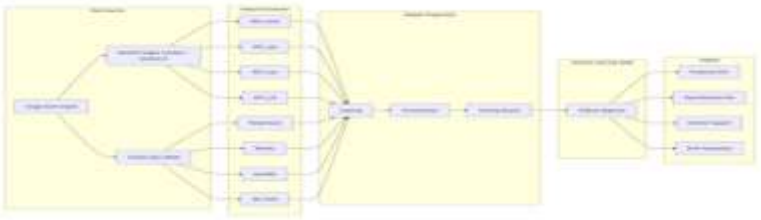
1. Temperature
2. DewPoint
3. Humidity
4. Rainfall
5. NDVI_std
6. NDVI_min
7. NDVI_max

The mean vegetation index (NDVI_mean) was adopted as the target variable that the model seeks to predict. The XGBoost Regression algorithm was selected for several reasons:

1. Its high capability in handling complex relationships and non-linear data.
2. Its efficiency and high speed in processing small to medium-sized datasets.
3. Its effective resistance to the overfitting problem.
4. Its distinguished capacity to extract interdependent relationships among diverse environmental variables.

The model's learning process relies on the direct relationship between environmental variables and vegetation index statistics. It does not use "Year" or "Season" as direct inputs in the training process; rather, their role within the system is limited to providing the temporal context of the data, not as features for training the NDVI_mean prediction model.**Outline of model structure and data as shown in Figure 6-1**

Figure 6-1:Outline of model structure and data



6.1 Model Training

The data were partitioned using the Hold-Out Validation approach, with the following allocation:

- 80% of the data for the training process (Training Set).
- 20% of the data for the testing process (Testing Set).

Accordingly, the training set comprised 83 records, while the test set consisted of 21 records.

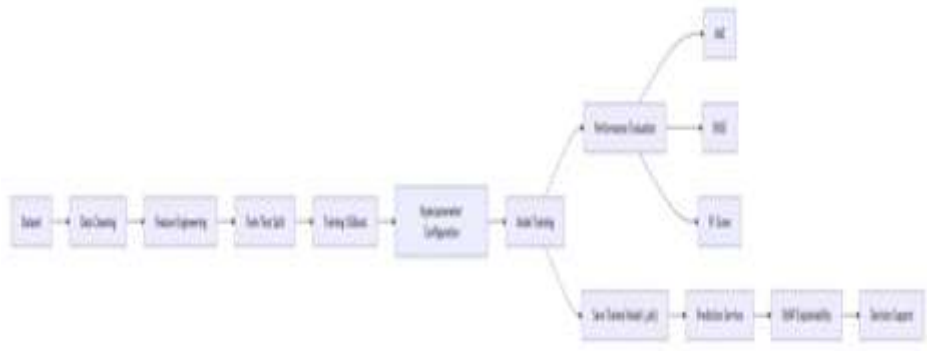
The model was trained using a set of experimental hyperparameters that demonstrated high efficiency and suitability for the available data, the most important of which include:

- Number of trees (n_estimators): 300 trees.
- Learning rate (learning_rate): 0.05.
- Maximum tree depth (max_depth): 4 levels.

Training data were used to construct the XGBoost model, while test data were reserved for evaluating model performance on unseen data, with the aim of accurately measuring its generalization capability.

Upon completion of the training phase, the model was saved in .pkl format using the Joblib library, enabling it to be directly reloaded within the system whenever NDVI_mean prediction operations are required, without the need to retrain the model.

Figure 6-2 An Environmental Decision Support System Based on Remote Sensing and Explainable Machine Learning



6.2 Model Training and Performance Evaluation Results

The XGBoost Regression model was trained using a set of empirically optimized hyperparameters, consisting of 300 trees, a learning rate of 0.05, and a maximum tree depth of 4 levels. The training and evaluation process yielded the results presented in the following table:

Table 6.1. Statistical Performance Metrics of the Model on Training and Test Data

Statistical Scale	Training Data (Training)	Test Data (Testing)
MAE	0.001003	0.062090
MSE	0.000002	0.006259
RMSE	0.001318	0.079114
R^2	0.999880	0.662871
Correlation coefficient (Correlation)	0.999946	0.825874

These results indicate that the model achieved near-perfect fit with the training data, whereas the test results—representing the true measure of the model's generalization capability—demonstrated good and reliable performance. Specifically, the model attained a coefficient of determination (R^2) of 0.6629 on the test data, meaning it can explain approximately 66% of the variance in previously unseen NDVI_mean values. Furthermore, a correlation coefficient of 0.8259 was observed between the actual and predicted values, reflecting a strong relationship between the model's outputs and the ground-truth values.

6.3 Feature Importance Analysis Results

The analysis revealed that the maximum value of the vegetation index (NDVI_max) exerted the greatest influence on predicting the mean index value, followed by temperature (Temperature) in second place, then the minimum index value (NDVI_min) and its standard deviation (NDVI_std). In contrast, dew point (DewPoint), precipitation (Rainfall), and relative humidity (Humidity) ranked lower in terms of relative influence.

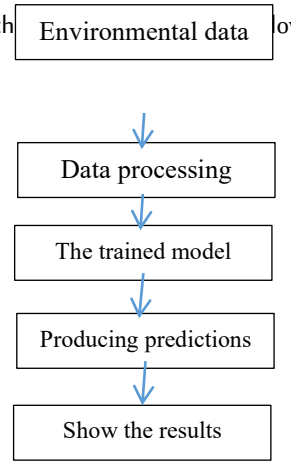
6.4 Model Interpretation Results Using SHAP

The SHAP technique was applied, and the results confirmed that NDVI_max remained the most influential factor in prediction, followed by temperature in second place. Meanwhile, features derived from NDVI statistics—such as the maximum value, minimum value, and standard deviation—played the predominant role in enhancing model accuracy. In contrast, the influence of humidity and precipitation remained relatively limited compared to the other variables.

These findings are consistent with the feature importance outputs directly derived from the XGBoost algorithm, thereby reinforcing confidence in the model's internal consistency.

6.5 System Development

Upon completion of model training, the model was integrated into the overall architecture of the proposed system, constituting a core component of the environmental decision support framework. The system's general arch



lowing stages

6.5.1. Model Persistence

The model was saved in PKL format using the Joblib library, enabling it to be reloaded within the system and execute predictions directly.

6.5.2 Future Prediction Support

The system was designed to enable the trained model to estimate NDVI_mean values upon input of new environmental data. The predicted value is then utilized to assess the current environmental status and display the vegetation cover level in a manner that assists the user in comprehending the condition of the study area.

6.5.3: Future Prediction

Extract recent environmental data.

1. Compute vegetation cover indices.
2. Generate the required features.
3. Load the trained model.
4. Execute the prediction process.
5. Display the results within the system.

Upon obtaining the NDVI_mean value, the system analyzes the result and classifies the environmental status, then presents decision-support recommendations. These include suggesting appropriate afforestation measures and vegetation cover improvement actions tailored to the estimated condition, along with the capability to display the model's decision interpretation using SHAP results to enhance result transparency for the user.

7. Challenges Encountered in System Development

Despite the effectiveness of the proposed methodology in integrating remote sensing data, climatic variables, and machine learning algorithms within an environmental decision support system, the development and implementation of the system encountered several methodological and technical challenges, summarized as follows:

7.1 Satellite Data Quality and Homogeneity

One of the most prominent challenges lay in managing the impact of cloud cover, alongside the spatial and temporal resolution discrepancies among the satellite data sources employed.

7.2 Selection and Updating of Climatic Variables

Identifying the climatic variables most strongly associated with vegetation cover change posed a methodological challenge, owing to the multiplicity of influencing variables and the variability in their relative importance.

7.3 Machine Learning Model Construction and Tuning

Constructing the XGBoost model required appropriate tuning of hyperparameters to enhance its generalization capability and mitigate the risk of overfitting. Furthermore, the limited available computational resources constituted a practical constraint on conducting extensive hyperparameter space searches or testing more complex models.

7.4 Interpretation of AI Model Outputs

A significant challenge involved translating model outputs into information comprehensible to non-specialist users. Integrating interpretability mechanisms within the system introduced additional processing and runtime requirements, necessitating a balance between the level of interpretability and response speed.

7.5 Integration of Decision Support System Components

System development required functional integration among multiple software and technical components, including the database, Google Earth Engine (GEE) platform, XGBoost model, SHAP tool, and interactive user interfaces. Additionally, designing a mechanism to convert the model's digital outputs into actionable environmental information and recommendations usable by decision-makers was essential.

7.6 Scalability and Sustainability

The system's scalability to cover new geographical areas was contingent upon re-evaluating model performance and retraining it when environmental characteristics and data sources differ from those of the original study area. Moreover, operating the system at a broader scale requires establishing a sustainable mechanism for data updates and periodic model performance monitoring.

8. Discussion

This section provides an in-depth analysis and discussion of the results of the proposed environmental decision support system, which integrates remote sensing techniques, geographic information systems, and machine learning algorithms to monitor vegetation cover and predict desertification risk in the Al Jabal Al Akhdar region of Libya.

At its core, the system relies on integrating satellite data extracted from the Google Earth Engine (GEE) platform with multi-source climatic data, employing the Normalized Difference Vegetation Index (NDVI) as the primary dependent variable, alongside the XGBoost Regression model for predicting its values, and the

SHAP technique for interpreting each variable's contribution to the model's decision and enhancing its transparency.

8.1 Discussion of Data Quality and Sources

The accuracy of any environmental predictive model depends largely on the quality of the data upon which it is built. The GEE platform enabled direct extraction of the NDVI index and execution of necessary preprocessing steps, such as cloud masking, delineation of the study area boundaries, and derivation of statistical vegetation indices. These data were supplemented with temperature, precipitation, humidity, and dew point variables; their integration with remote sensing data yielded a marked improvement in the model's capacity to explain environmental changes compared to relying on NDVI alone.

The data preparation phase included cleaning, handling missing values, and temporal–spatial harmonization, which contributed to reducing noise levels and enhancing model performance stability during training.

8.2 Discussion of Environmental Indices Employed

The proposed system relied on integrating the NDVI index with a set of climatic variables. Four statistical properties of NDVI were used—mean, minimum, maximum, and standard deviation—to provide a more comprehensive description of vegetation cover status. The results revealed that relying on the mean alone is insufficient, as the remaining properties contributed to clarifying spatial variability among regions and identifying locations of degradation or high vegetation density.

Furthermore, the results indicated that temperature, precipitation, humidity, and dew point exert a direct influence on NDVI value changes: elevated temperatures coupled with reduced precipitation lead to diminished vegetation activity, whereas increased humidity and water availability enhance plant growth. Collectively, these findings confirm that the selection of vegetation and climatic indices was appropriate, and that their integration improved the predictive model's performance and delivered more comprehensive information to support environmental decision-making.

8.3 Discussion of XGBoost Model Performance

The success of environmental forecasting systems hinges on selecting a machine learning algorithm capable of handling complex non-linear relationships among variables. Accordingly, the XGBoost model was chosen for

its high accuracy, resistance to overfitting, and efficiency in processing multi-feature datasets. Training results demonstrated high accuracy in learning the relationships between environmental variables and the target variable, while test results indicated good generalization capability: the model achieved a coefficient of determination (R^2) of 0.6629 and a correlation coefficient of 0.8259, confirming its suitability for estimating mean NDVI values within the proposed system

8.4 Discussion of SHAP Technique Results

The SHAP technique was integrated to clarify each variable's contribution to the XGBoost model's outputs. This technique successfully identified the relative importance of variables, with clear variability in their influence according to prevailing environmental conditions, thereby aiding in interpreting the model's decision-making mechanism and enhancing result transparency for the end user.

SHAP-generated visualizations revealed that the influence of environmental factors varies from one state to another: precipitation may be the most influential factor during certain time periods, whereas temperature or humidity gains greater importance in others, reflecting the dynamic nature of ecosystems.

8.5 Discussion of the Environmental Decision Support System

The results demonstrated that the system enables visualization of the current vegetation cover status, analysis of relevant climatic indices, estimation of mean NDVI based on input environmental data, classification of environmental status, and provision of practical afforestation and vegetation improvement recommendations, alongside simplified SHAP-based interpretation of model outputs for the user.

Moreover, the system was found to be scalable for future enhancement through automatic data updates via GEE, the potential addition of new environmental indices, and expansion of its application scope to encompass other Libyan regions facing similar risks.

8.6 Comparison of Results with Prior Studies

To contextualize the current findings within the broader scientific landscape, Table 8.1 presents a comparison between the results of this study and three prior studies that employed machine learning and remote sensing techniques for vegetation cover and desertification analysis, with respect to study area, algorithm used, target variable, evaluation metrics, and key findings.

Table 8.1: Comparison of Results with Prior Studies

Study	Study area	Algorithm	Target variable	Evaluation metrics	Main findings
Present study	Al Jabal Al Akhdar, Libya	XGBoost Regression	NDVI_mean	R^2, RMSE, MAE, Correlation	R^2=0.6629, RMSE=0.0791, MAE=0.0621, Corr=0.8259
[1]Li et al.	China	XGBoost Regression	NDVI	R^2, error rate, accuracy	$R^2 > 0.80$; error < 5%; accuracy > 89%
[2]Hayeri Yazdi	Isfahan, Iran	XGBoost Regression	NDVI	MAE, RMSE, Correlation, NSE	MAE=0.0061, RMSE=0.0079, Corr≈0.692
[3]Al-Tameemi et al.	Iraq	XGBoost Classification + SHAP	Desertification-risk classification	Accuracy, Kappa	Accuracy=94.46%, Kappa=0.9261

The results of this comparison show that the current study is consistent with previous literature in demonstrating the efficiency of the XGBoost algorithm in addressing nonlinear relationships between climate variables and vegetation indicators. The proposed model achieved a determination factor (R2 = 0.6629) and a correlation factor (0.8259), demonstrating a good ability to interpret variation in vegetation data and predict seasonal trends in the Green Mountain region.

Comparing this finding to the Li et al study[1]. The study, which developed a machine learning model using the Extreme Gradient Boosting algorithm to predict vegetation growth throughout the growing season in China from 2001 to 2018, shows that the value of the determination coefficient was higher (R2> 0.80)

Compared to the Hayeri Yazdi study[2] , this study relied on drought indicators (SPI and VCI) combined with previous NDVI values to predict future NDVI values in the short term, and achieved relatively low values for both MAE and RMSE when predicting one month later.

For the Al-Tameemi et al study[3], which employed the XGBoost model along with SHAP technology to explain the contribution of each variable to the classification of desertification risk scores in Iraq, the comparison is limited to the methodological aspect of the sub-digital, due to the different nature of the problem The Iraqi study relied on classification, while the current study relied on regression. However, the two studies agree to emphasize the importance of vegetation indicators and climate variables as one of the most influential factors in model performance.

Overall, this comparison indicates that the results of the current study are consistent with recent trends in the use of machine learning and remote sensing for vegetation analysis, and confirms that the XGBoost algorithm is an effective option for building intelligent models for predicting vegetation status and supporting desertification monitoring systems

8.7 System Evaluation: Strengths and Limitations

Post-implementation system evaluation revealed that it achieved most of the objectives defined at the study's outset, successfully integrating, processing, and utilizing multiple data sources to construct a relatively accurate and interpretable predictive model simultaneously. Among the most prominent strengths identified were the system's capability to integrate satellite data with climatic data within a single platform, the provision of an interactive interface facilitating user data exploration and result analysis, and the speed of prediction execution suitable for practical applications. Furthermore, the use of SHAP enhanced system reliability by interpreting the reasons behind predictions, thereby increasing transparency and facilitating its adoption in decision support.

Conversely, the evaluation uncovered certain limitations, most notably that the model was trained exclusively on data from the Al Jabal Al Akhdar region, necessitating retraining with new data when attempting to generalize it to other areas. Additionally, the system's continued accuracy remains contingent upon periodic data updates. Compounding this is the complex nature of ecosystems, influenced by multiple factors that the model may not fully represent—such as natural disasters, abrupt changes, and human activity—which may introduce some deviation in predictions. Despite these limitations, the system remains highly amenable to future development, particularly given the potential for direct linkage to GEE data, geographical scope expansion, and subsequent integration of more advanced models.

9. Conclusions

This study concludes that the integration of remote sensing techniques, reanalyzed climatic data, and interpretable machine learning algorithms can provide a practical framework for environmental decision support, monitoring vegetation cover dynamics, and assessing desertification risk in semi-arid environments, as methodologically demonstrated in the Al Jabal Al Akhdar region of Libya. The results demonstrated that integrating NDVI time series extracted via the Google Earth Engine (GEE) platform with high-resolution climatic variables from the ERA5-Land reanalysis dataset furnished a rich data environment, enabling the XGBoost model to capture the complex, non-linear dynamics of environmental degradation with advanced accuracy ($R^2 = 0.6629$, $R = 0.8259$).

Furthermore, the integration of the SHAP technique proved methodologically successful in deconstructing the "black-box" nature of the machine learning model: spatial and seasonal interpretations provided precise quantitative explanations of the relative weights of temperature, precipitation, humidity, and dew point in guiding prediction decisions, thereby enhancing the scientific reliability of outputs for decision-makers. The system thus transforms into an operational tool that translates complex environmental data into actionable strategies, contributing to bridging the gap between academic analyses and field practice.

Despite the system's success in achieving its core objectives, certain challenges remain associated with the nature of local data, cloud cover effects, spatial and temporal resolution discrepancies among sources, and sample size limitations—necessitating continuous data updates and periodic model retraining, in addition to

the difficulty of representing all sudden human and natural factors within the current model. The outputs of this research affirm that digital solutions based on the integration of remote sensing and artificial intelligence constitute a fundamental pillar for developing proactive and sustainable environmental management.

10. Recommendations

Based on the results and challenges identified, this study recommends the following:

First: Expand environmental indicators by adding spectral indices tailored to arid regions (e.g., SAVI, BSI) and linking them to anthropogenic and economic factors; generalize the methodology to other regions in Libya; and integrate variables related to soil, topography, and land use.

Second: Automate the update and retraining loop; adopt hybrid models and provide confidence intervals for predictions; and leverage near-real-time data (Sentinel-2, MODIS) for detecting rapid changes. It is also advisable to test hybrid models and alternative machine learning algorithms, systematically comparing them with the XGBoost model used in this study, while employing appropriate methods to estimate uncertainty and provide predictive intervals or uncertainty bounds accompanying the predictions.

Third: Establish institutional partnerships to ensure data continuity and provide the necessary infrastructure; implement intensive programs for building local capacity in environmental data analysis and AI deployment to ensure system sustainability and future scalability. It is further proposed to execute specialized training programs to build local capacity in remote sensing, geographic information systems, environmental data analysis, and explainable AI applications.

References

[1] United Nations Convention to Combat Desertification (UNCCD), The Global Land Outlook, Second Edition: Land Restoration for Recovery and Resilience—Summary for Decision Makers. Bonn, Germany: UNCCD, 2022.

[2] J. R. Jensen, Introductory Digital Image Processing: A Remote Sensing Perspective, 4th ed. Pearson Education, 2015. ISBN: 978-0-13-405816-0.

[3] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore, "Google Earth Engine: Planetary-scale geospatial analysis for everyone," Remote Sensing of Environment, vol. 202, pp. 18–27, 2017, doi: 10.1016/j.rse.2017.06.031.

[4] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 2016, pp. 785–794, doi: 10.1145/2939672.2939785.

[5] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in Advances in Neural Information Processing Systems, vol. 30, 2017, pp. 4765–4774.

[6] D. Rivera-Marin, J. Dash, and B. O. Ogutu, "The use of remote sensing for desertification studies: A review," Journal of Arid Environments, vol. 206, Art. no. 104829, 2022, doi: 10.1016/j.jaridenv.2022.104829.

[7] R. Almalki, M. Khaki, P. M. Saco, and J. F. Rodriguez, "Monitoring and mapping vegetation cover changes in arid and semi-arid areas using remote sensing technology: A review," Remote Sensing, vol. 14, no. 20, Art. no. 5143, 2022, doi: 10.3390/rs14205143.

[8] M. A. E. AbdelRahman, "An overview of land degradation, desertification and sustainable land management using GIS and remote sensing applications," Rendiconti Lincei. Scienze Fisiche e Naturali, vol. 34, pp. 767–808, 2023, doi: 10.1007/s12210-023-01155-3.

[9] M. S. Alghariani, E. M. Sagar, H. Bedair, and A. M. F. Al-Quraishi, "Google Earth Engine (GEE) to detect vegetation cover changes in northwest of Libya," in Climate Change and Environmental Degradation in the MENA Region, Cham, Switzerland: Springer, 2024, pp. 1–18, doi: 10.1007/698_2024_1103.

[10] N. M. G. Almesmari, "Natural vegetation cover changes in north-east Libya," Ph.D. dissertation, School of Water, Energy, Environment and Agrifood, Cranfield University, Cranfield, U.K., 2019.

[11] L. Rayne, F. Brandolini, J. L. Makovics, E. Hayes-Rich, J. Levy, H. Irvine, L. Assi, and Y. Bokbot, "Detecting desertification in the ancient oases of southern Morocco," Scientific Reports, vol. 13, Art. no. 19424, 2023, doi: 10.1038/s41598-023-46319-1.

[12] A. Moumane, Y. Azougarh, A. Ait Enajar, W. S. Alkhuraiji, I. Bahdou, J. Al Karkouri, F. Nahas, N. Y. Rebouh, and Y. M. Youssef, "Desertification monitoring in arid oasis environment using Google Earth Engine, machine learning, and field-based hydrogeological assessment," Scientific Reports, vol. 16, Art. no. 9212, 2026, doi: 10.1038/s41598-026-41216-9.

[13] Batunacun, R. Wieland, T. Lakes, and C. Nendel, "Using Shapley additive explanations to interpret extreme gradient boosting predictions of grassland degradation in Xilingol, China," Geoscientific Model Development, vol. 14, no. 3, pp. 1493–1510, 2021, doi: 10.5194/gmd-14-1493-2021.

[14] A. Boali, H. R. Asgari, A. M. Behbahani, A. Salmanmahiny, and B. Naimi, "Remotely sensed desertification modeling using ensemble of machine learning algorithms," Remote Sensing Applications: Society and Environment, vol. 34, Art. no. 101149, 2024, doi: 10.1016/j.rsase.2024.101149.

[15] X. Meng, S. Li, K. Akhmadi, P. He, and G. Dong, "Trends, turning points, and driving forces of desertification in global arid land based on the segmental trend method and SHAP model," GIScience & Remote Sensing, vol. 61, no. 1, Art. no. 2367806, 2024, doi: 10.1080/15481603.2024.2367806.

[16] N. Al-Tameemi, Z. Xuexia, F. Shahzad, K. Mehmood, and J. Zhou, "Enhancing desertification risk mapping with entropy-based weighting and machine learning: Insights from Iraq," Journal of Environmental Management, vol. 396, Art. no. 127976, 2025, doi: 10.1016/j.jenvman.2025.127976.

[17] X. Cao, M. Cui, L. Xi, and Y. Feng, "Spatial-temporal process of land use/land cover and desertification in the Circum-Tarim Basin during 1990–2020," Land, vol. 13, no. 6, Art. no. 735, 2024, doi: 10.3390/land13060735.

[18] S. Zhao, J. Ding, J. Wang, X. Ge, L. Han, R. Wang, and S. Qin, "Central Asia's desertification challenge: Recent trends and drives explored with Google Earth Engine," Journal of Cleaner Production, vol. 460, Art. no. 142595, 2024, doi: 10.1016/j.jclepro.2024.142595.

[19] Y. Wang, E. Guo, Y. Kang, and H. Ma, "Assessment of land desertification and its drivers on the Mongolian Plateau using intensity analysis and the geographical detector technique," Remote Sensing, vol. 14, no. 24, Art. no. 6365, 2022, doi: 10.3390/rs14246365.

[20] J. Hao, Q. Lin, T. Wu, J. Chen, W. Li, X. Wu, G. Hu, and Y. La, "Spatial-temporal and driving factors of land use/cover change in Mongolia from 1990 to 2021," Remote Sensing, vol. 15, no. 7, Art. no. 1813, 2023, doi: 10.3390/rs15071813.

[21] J. Li, C. Cao, M. Xu, X. Yang, X. Gao, K. Wang, H. Guo, and Y. Yang, "A 20-year analysis of the dynamics and driving factors of grassland desertification in Xilingol, China," Remote Sensing, vol. 15, no. 24, Art. no. 5716, 2023, doi: 10.3390/rs15245716.

[22] W. Shao, L. Li, M. Yan, Z. Meng, L. Zhang, Q. Zhang, and Y. Chen, "Thirty years' spatio-temporal evolution of desertification degrees and driving factors in Turpan–Hami Basin, Xinjiang, China," Ecological Indicators, vol. 166, Art. no. 112484, 2024, doi: 10.1016/j.ecolind.2024.112484.

[23] W. Gao, Z. Huang, X. Li, B. Ji, N. Li, S. Li, X. Liu, Q. Zeng, G. Sun, and D. Zhao, "The spatial-temporal changes and driving factors of desertification in the Wuliangsuhai watershed based on remote sensing," Ecological Indicators, vol. 169, Art. no. 112851, 2024, doi: 10.1016/j.ecolind.2024.112851.

[24] L. Ran, L. Zhang, Y. He, S. Cao, Y. Ding, Y. Guo, X. Wei, and M. Filonchyk, "Spatiotemporal dynamic change and the driving mechanism of desertification in the Yellow River Basin," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 17, pp. 17134–17155, 2024, doi: 10.1109/JSTARS.2024.3453295.

[25] H. Li, K. Yang, Y. Cui, L. Ai, C. Wang, Z. Wang, and C. Zhang, "Desertification mitigation in Northern China was promoted by climate drivers after 2000," Remote Sensing, vol. 16, no. 19, Art. no. 3706, 2024, doi: 10.3390/rs16193706.

[26] Z. Jiang, X. Ni, and M. Xing, "A study on spatial and temporal dynamic changes of desertification in northern China from 2000 to 2020," Remote Sensing, vol. 15, no. 5, Art. no. 1368, 2023, doi: 10.3390/rs15051368.

[27] X. Chen, N. Wang, S. Peng, N. Meng, and H. Lv, "Analysis of spatiotemporal dynamics of land desertification in Qilian Mountain National Park based on Google Earth Engine," ISPRS International Journal of Geo-Information, vol. 13, no. 4, Art. no. 117, 2024, doi: 10.3390/ijgi13040117.

[28] M. Ngabire, T. Wang, J. Liao, and G. Sahbeni, "Quantitative analysis of desertification-driving mechanisms in the Shiyang River Basin," Remote Sensing, vol. 15, no. 12, Art. no. 2960, 2023.

[29] Q. Liu, G. Liu, and C. Huang, "Monitoring desertification processes in Mongolian Plateau using MODIS tasseled cap transformation and TGSI time series," Journal of Arid Land, vol. 10, no. 1, pp. 12–26, 2018.

[30] Z. Wang, Y. Shi, and Y. Zhang, "Review of desert mobility assessment and desertification monitoring based on remote sensing," Remote Sensing, vol. 15, no. 18, Art. no. 4412, 2023, doi: 10.3390/rs15184412.

[31] Sahara and Sahel Observatory (OSS), "GETCROP: Crop monitoring and early warning in North Africa," 2025. [Online]. Available: https://www.oss-online.org/

[32] J. Muñoz-Sabater, E. Dutra, A. Agustí-Panareda, C. Albergel, G. Arduini, G. Balsamo, S. Boussetta, M. Choulga, S. Harrigan, H. Hersbach, B. Martens, D. G. Miralles, M. Piles, N. J. Rodríguez-Fernández, E. Zsoter, C. Buontempo, and J.-N. Thépaut, "ERA5-Land: A state-of-the-art global reanalysis dataset for land applications," Earth System Science Data, vol. 13, pp. 4349–4383, 2021, doi: 10.5194/essd-13-4349-2021.