Monitoring the Status of Vegetation in the Agricultural Fields in the Area of Parco Lambro, Milan (Spring-Summer 2022/2023)

Author: Ozgur Gumus, Ghazaleh Monsef Shemordeh

Date: 17/06/2024

1. Introduction

Objective

The primary goal of this project is to monitor the status of vegetation in the agricultural fields of Parco Lambro, Milan, during Spring-Summer 2022/2023. This involves analyzing vegetation health using the Normalized Difference Vegetation Index (NDVI) and classifying different types of vegetation, specifically trees, meadows, and bushes.

Study Area

Parco Lambro is a significant green space in Milan, known for its ecological diversity and agricultural fields. Monitoring vegetation in this area is crucial for understanding its ecological health and aiding in effective land management.

Figure 1: Parco Lambro Area

Figure 2: Locations of the training points

2. Data Collection and Preprocessing

Data Sources

The analysis utilizes Sentinel-2 satellite imagery, which provides high-resolution optical images of the Earth's surface. Sentinel-2 data is particularly suitable for vegetation monitoring due to its multispectral capabilities.

Data Filtering

It takes five days for the Sentinel-2 satellite to complete a pass around the equator, so under normal circumstances, a new image is available every five days. However, clear images are essential for accurately identifying NDVI data, which is why cloudy days are eliminated from the dataset. To ensure the best results, cloud coverage is limited to a maximum of 15%.

3. NDVI Analysis

NDVI Calculation

The Normalized Difference Vegetation Index (NDVI) is a widely used indicator of vegetation health. It is calculated using the near-infrared (NIR) and red bands of the spectrum:

$$\text{NDVI} = (\frac{(NIR - Red)}{(NIR + Red)}$$

Higher NDVI values indicate healthier, denser vegetation, while lower values suggest sparse or stressed vegetation.

Temporal Analysis

NDVI values were extracted for the entire study period, and the median NDVI value was calculated to provide a comprehensive view of vegetation health over time.

Visualization

The median NDVI for Parco Lambro was visualized, revealing areas of healthy vegetation as well as regions that may require attention.

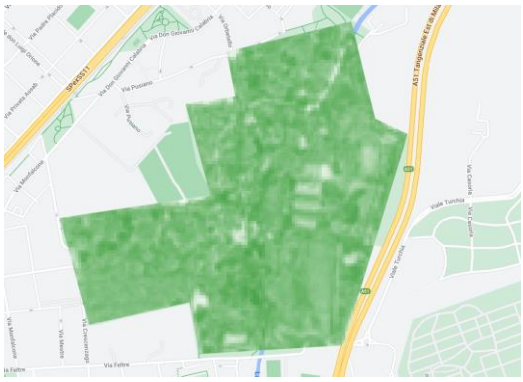

Figure 3: Median NDVI for Parco Lambro (2022-2023)

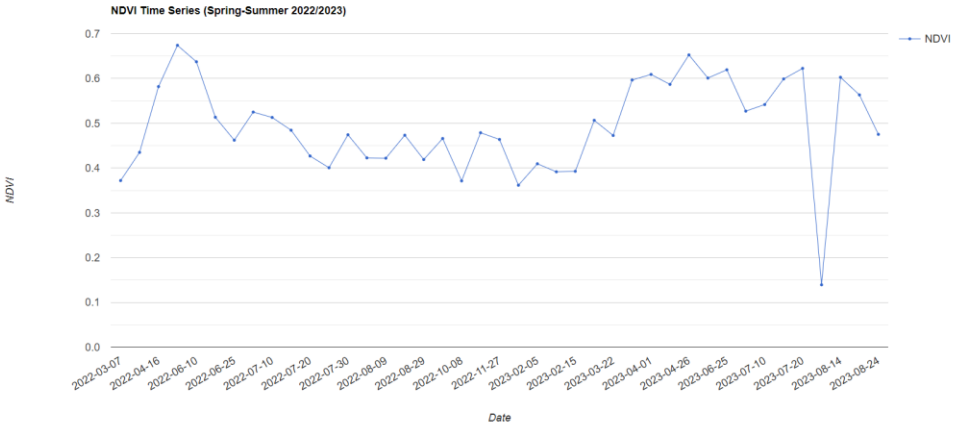

Figure 4: NDVI Time Series (Spring-Summer 2022/2023)

4. Vegetation Classification

Training Data

Training points for different vegetation types (trees, meadows, and bushes) were manually selected based on visual inspection and expert knowledge. These training points are crucial for training the classifiers.

Classification Methods

Two classification methods were employed:

- **Random Forest (RF):** An ensemble learning method that constructs multiple decision trees during training and outputs the mode of the classes.
- **Support Vector Machine (SVM):** A supervised learning model that classifies data by finding the hyperplane that best separates the classes in the feature space.

Classification Results

Both classifiers were trained using the selected training points and applied to the median NDVI image to classify the vegetation types.

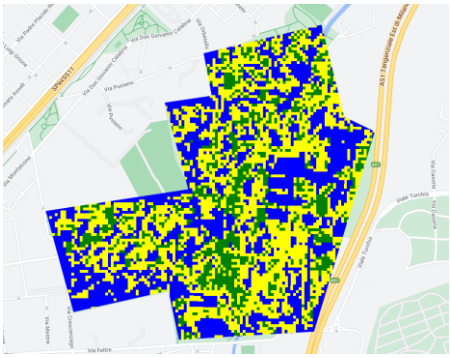

Figure 5: Random Forest Classification

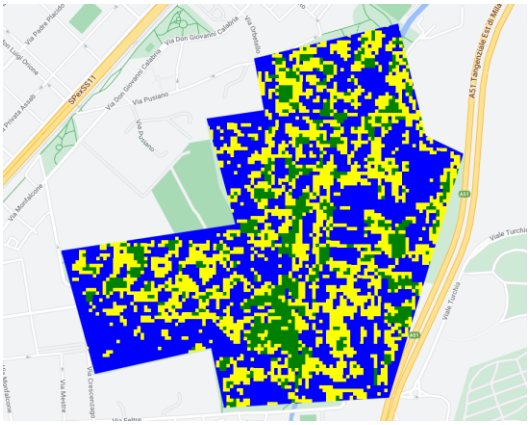

Figure 6: SVM Classification

Performance Metrics
The performance of both classifiers was evaluated using accuracy and kappa metrics. The Random Forest classifier demonstrated higher accuracy and kappa values compared to the SVM classifier.

Random Forest Performance:

- Accuracy: 0.967
- Kappa: 0.950

SVM Performance:

- Accuracy: 0.767
- Kappa: 0.650

5. Results and Discussion

NDVI Analysis Results

The NDVI analysis revealed distinct spatial patterns of vegetation health within Parco Lambro. Areas with high NDVI values corresponded to healthy, dense vegetation, primarily in regions dominated by trees and well-maintained meadows. Lower NDVI values were observed in areas with less dense vegetation or where vegetation might be under stress.

Classification Results

Both the Random Forest and SVM classifiers successfully differentiated between trees, meadows, and bushes. The classification maps provided valuable insights into the spatial distribution of these vegetation types.

- **Random Forest Classification:** This method showed a high accuracy in distinguishing between different vegetation types, leveraging its ensemble approach to minimize overfitting and improve generalization.

- **SVM Classification:** The SVM classifier also performed well, particularly in separating distinct vegetation classes. However, it was slightly less effective than Random Forest in dealing with overlapping vegetation classes.

Comparison of Classifiers

The Random Forest classifier demonstrated a slight edge over the SVM classifier in terms of accuracy and robustness. This can be attributed to its ability to handle a larger feature space and its ensemble nature, which mitigates the impact of noisy data.

Temporal Analysis of Vegetation Classes

The temporal analysis of the classified vegetation types was conducted to understand the changes in vegetation cover over time. The following chart illustrates the area of each vegetation class (trees, meadows, and bushes) over the study period.

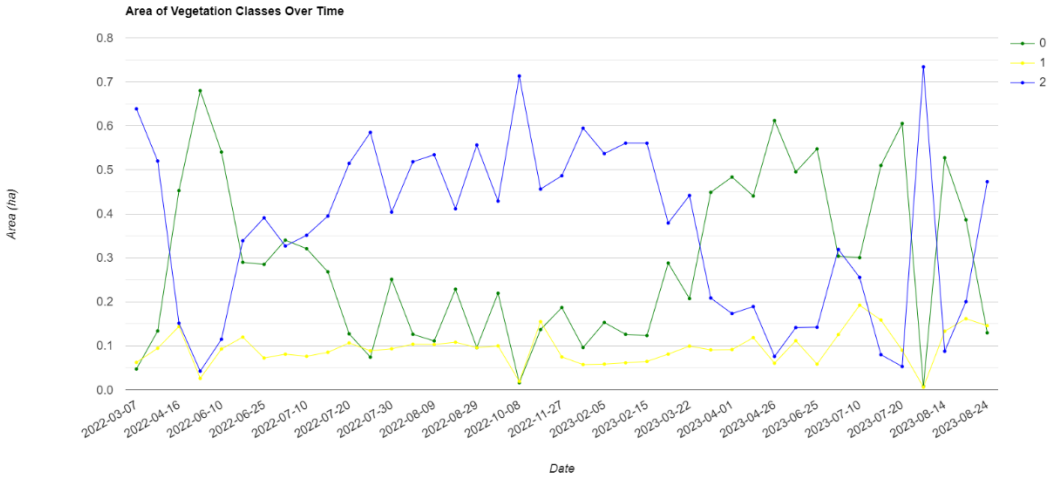

Figure 7: Area of Vegetation Classes Over Time

This chart indicates the fluctuation in the area covered by different vegetation types throughout the study period. Notably, there are significant variations in the area covered by trees and bushes, reflecting the dynamic nature of vegetation growth and health in Parco Lambro.

- **Trees (Class 0):** The area covered by trees showed considerable variation, with peaks and troughs reflecting periods of growth and potential stress or loss.
- **Meadows (Class 1):** Meadows maintained relatively stable coverage with minor fluctuations, indicating consistent management or natural stability.
- **Bushes (Class 2):** The area covered by bushes also exhibited fluctuations, possibly due to seasonal growth patterns or management practices.

Implications of Findings

The classification maps and NDVI analysis provide a comprehensive understanding of the vegetation dynamics in Parco Lambro. These findings can inform local authorities and land

managers about areas that require intervention or further monitoring. The methodology demonstrated here can be applied to other regions and time periods to support ongoing ecological monitoring and land management efforts.

6. Conclusion

This study successfully monitored the vegetation status in Parco Lambro using NDVI and classified vegetation types with two different classifiers. The NDVI analysis highlighted the areas of healthy vegetation as well as those under potential stress. The classification results provided detailed insights into the distribution of trees, meadows, and bushes within the park.

7. References

- Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5-32.
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine Learning, 20(3), 273-297.