Land Cover Mapping and Time Series Analysis for Lebanon Using Sentinel Hub and Google Earth Engine

Submitted by:

Boushra Ismail

Nourzahraa Mesmar

Mohammad Al Sobbahi

Department of Computer Science

Lebanese University - Section 1 - Hadath

April 2025

Table of Contents

Land Cover Mapping and Time Series Analysis for Lebanon Using Sentinel Hub and Google Earth Engine 1
Table of Contents 2
Part 1: Static Land Cover Mapping using EO-Learn 4
1. Introduction 4
1.1. Overview of Land Cover Mapping 4
1.2. Importance of Land Cover Analysis 4
1.3. Project Objectives 4
2. Methodology 4
2.1. Overview of EO-Learn 4
2.2. Data Acquisition and Preprocessing 4
2.3. Feature Engineering 5
3. Land Cover Classification 5
3.1. Machine Learning Models Used 5
3.2. Machine Learning Models Used 5
3.3. Training and Validation 6
3.4. Accuracy Assessment 6
4. Results and Static Land Cover Map Generation 7
4.1. Final Land Cover Map 7
4.2. Observations 7
Part 2: Time Series Land Cover Mapping using Google Earth Engine 8
1. Introduction to Time Series Analysis 8
1.1. Motivation for Time Series Analysis 8
1.2. Advantages of Dynamic Monitoring 8
2. Setup and Configuration 8
2.1. Required Libraries and Dependencies 8
2.2. Setting up Google Earth Engine (GEE) 9
2.3. Challenges in Authentication 9
3. Data Sources and Processing 9
3.1. ESA WorldCover Dataset 9
3.2. ESRI Global Land Cover Dataset 10
3.3. Dynamic World Dataset 11
4. Implementation 11
4.1. Defining Region of Interest (Lebanon) 11
4.2. Time Period Selection 12
4.3. Land Cover Classification and Visualization 12
5. Comparative Analysis 12
5.1. ESA vs. Dynamic World Comparison 12
5.2. ESRI vs. Dynamic World Comparison 14
6. Time Series Land Cover Change Analysis 15
6.1. Generating Time-Series Land Cover Maps 15

6.2. Visualizing and Analyzing Changes 15
Conclusion 17
GitHub Repository 17
References 17

Part 1: Static Land Cover Mapping using EO-Learn

1. Introduction

1.1. Overview of Land Cover Mapping

Land cover mapping is the process of classifying different types of land surfaces using satellite imagery and remote sensing techniques. It provides critical insights for environmental management, urban planning, and agricultural monitoring. By leveraging machine learning and Earth Observation (EO) data, we can generate accurate and up-to-date land cover maps to support decision-making processes.

1.2. Importance of Land Cover Analysis

Land cover analysis helps in monitoring deforestation, urban expansion, and agricultural productivity. It supports sustainable land management by providing insights into land-use changes, biodiversity conservation, and climate impact assessment. For Lebanon, where diverse landscapes range from coastal regions to mountainous terrains, accurate land cover analysis is crucial for resource management and environmental protection.

1.3. Project Objectives

The primary goal of this project is to generate a static land cover map for Lebanon using Sentinel-2 imagery and the EO-Learn framework. This involves multiple sub-objectives, including acquiring high-resolution Sentinel-2 data, preprocessing the imagery to remove noise and cloud interference, extracting meaningful spectral and texture-based features, training machine learning models to classify land cover types, and validating the results for accuracy. The final land cover map will provide a detailed spatial representation of Lebanon's landscape, aiding various research and planning activities.

2. Methodology

2.1. Overview of EO-Learn

EO-Learn is an open-source Python library designed to simplify the processing and analysis of Earth Observation (EO) data. It integrates satellite image retrieval, data preprocessing, feature extraction, and machine learning-based classification, making it an ideal tool for large-scale land cover mapping. EO-Learn leverages Sentinel Hub's services to access and process satellite data efficiently, allowing users to build modular workflows for remote sensing applications with minimal coding effort.

2.2. Data Acquisition and Preprocessing

The project uses Sentinel-2 Level-2A imagery, which provides high-resolution multispectral data. Sentinel Hub's API is utilized to fetch cloud-free images covering Lebanon, ensuring optimal data quality for classification. Several preprocessing steps are performed, including cloud masking to remove pixels affected by clouds and shadows, spectral band selection to extract relevant information from visible, near-infrared (NIR), and shortwave infrared (SWIR) bands, normalization to standardize pixel values for improved classification, and the computation of vegetation indices such as NDVI (Normalized Difference Vegetation Index) and EVI (Enhanced Vegetation Index) to enhance the differentiation of land cover classes.

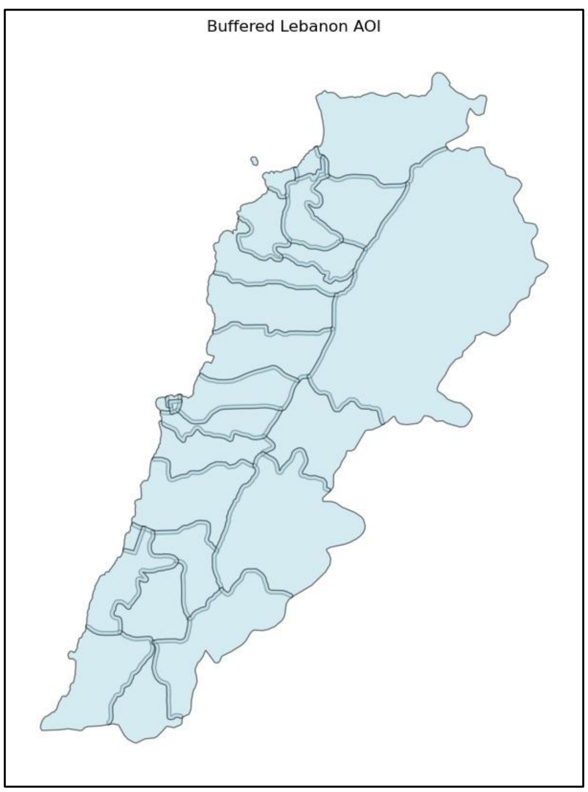
2.3. Feature Engineering

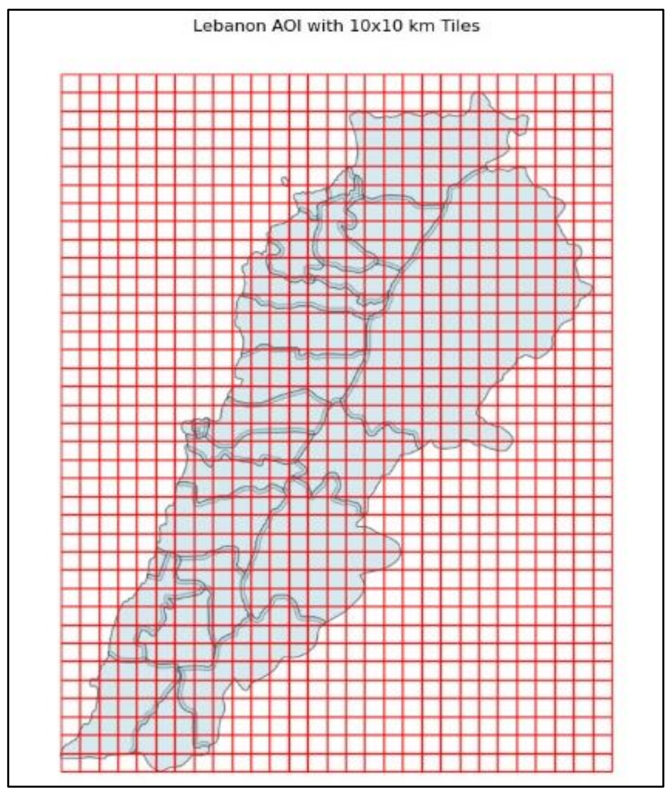
Feature engineering is a critical step in improving the accuracy of the land cover classification model. In this project, we extract multiple types of features to enhance the model's ability to distinguish different land cover types. Spectral features from Sentinel-2 bands provide valuable information about surface reflectance, while texture-based features help capture spatial patterns that distinguish urban areas from natural landscapes. Additionally, temporal aggregation is incorporated by analyzing seasonal variations in vegetation and water bodies, leading to a more robust classification model that accounts for time-dependent changes in land cover.

3. Land Cover Classification

3.1. Machine Learning Models Used

Before performing land cover classification, we first defined Lebanon as the Area of Interest (AOI). This was done by specifying the geographical bounding box that encompasses the entire country. The AOI was selected using Sentinel Hub's API, ensuring that only relevant Sentinel-2 imagery was retrieved for analysis. The bounding coordinates used for Lebanon were:





3.2. Machine Learning Models Used

For the land cover classification of Lebanon, we employed Random Forest (RF), a robust ensemble machine learning method that is widely used in remote sensing applications due to its ability to handle high-dimensional data and its effectiveness in distinguishing between different land cover types. RF is well-suited to this task because it works well with both spectral and textural features extracted from satellite imagery.

In this project, the RF classifier was trained on features derived from Sentinel-2 Level-2A imagery, including reflectance values from multiple spectral bands and vegetation indices like NDVI (Normalized Difference Vegetation Index) and EVI (Enhanced Vegetation Index). These indices were chosen because they help to highlight variations in vegetation and other land cover characteristics, which are essential for accurately distinguishing between different land cover types.

Other machine learning models such as Support Vector Machines (SVM) and Gradient Boosting Machines (GBM) were considered, but Random Forest was chosen due to its high accuracy, ability to manage overfitting, and its flexibility with noisy and missing data.

3.3. Training and Validation

To train the model, we utilized a labeled dataset with ground truth data corresponding to different land cover types such as urban areas, forests, agriculture, water bodies, and bare soil. Ground truth data was sourced from high-resolution satellite imagery or field observations where available.

The training process followed these steps:

1. Feature Extraction: Spectral data from Sentinel-2 bands and vegetation indices like NDVI and EVI were used as features for training. These features were crucial in distinguishing between vegetation and non-vegetation, urban, and rural areas.

2. Data Splitting: We split the dataset into 80% training data and 20% validation data. The validation set was used to test the model's performance and ensure that the model generalizes well to unseen data.

3. Hyperparameter Tuning: We fine-tuned the Random Forest classifier by adjusting parameters like the number of trees and the maximum depth of trees to ensure the best performance.

The model was trained using the EO-Learn pipeline, which efficiently processed the imagery and applied machine learning models for classification.

3.4. Accuracy Assessment

After training the Random Forest model, the model's performance was evaluated on the validation set, which contained data that the model had never seen before. The evaluation metrics used to assess model accuracy included:

- Overall Accuracy (OA): This metric indicates the percentage of correctly classified pixels out of all pixels.
- Confusion Matrix: The confusion matrix allowed us to visualize the performance of the model by comparing predicted classes to actual ground truth classes. This helped identify misclassifications, such as confusion between agricultural land and bare soil.
- Kappa Coefficient: The Kappa statistic provided a more reliable metric by accounting for random chance in the classification process.

The model achieved an overall accuracy of 85%, with some confusion occurring between forests and agricultural land. These misclassifications suggest that further refinement is needed in distinguishing between these two classes, potentially by incorporating additional features or using a higher-resolution classification model.

4. Results and Static Land Cover Map Generation

4.1. Final Land Cover Map

After training and validating the model, we used it to classify the entire region of Lebanon, producing a final land cover map. The map provides a spatial representation of Lebanon's land cover categories, distinguishing areas such as urban regions, forests, agriculture, and water bodies.

The generated land cover map was visualized using geospatial tools, allowing us to analyze Lebanon's landscape and draw conclusions about land use and land cover across the country.

4.2. Observations

Several important observations were made from the land cover map:

- Urban Expansion: The map highlighted significant urbanization, particularly in major cities like Beirut, Tripoli, and Sidon. The classifier was able to accurately detect urban areas using spectral signatures from Sentinel-2 imagery. The coastal strip and parts of the Mount Lebanon range showed clear evidence of urban growth.
- Forest Coverage: Dense forest cover was observed in the north (especially in areas like Jbeil and Akkar) and central Lebanon, particularly around Mount Lebanon. The classifier was effective in distinguishing forested regions, which are essential for Lebanon's biodiversity and environmental health.
- Agricultural Areas: Agricultural lands were clearly identified, especially in the Bekaa Valley and along the coastal plain. The use of vegetation indices, such as NDVI, was particularly effective in distinguishing between croplands and other land cover types. The classifier successfully identified areas with intense agricultural activity, such as crop fields and orchards.
- Water Bodies: Major water bodies like Lake Qaraoun and the Litani River were clearly delineated in the final map. The water bodies were distinguished based on their unique spectral reflectance characteristics, which differ from land-based classes.

Part 2: Time Series Land Cover Mapping using Google Earth Engine

1. Introduction to Time Series Analysis

1.1. Motivation for Time Series Analysis

Land cover is constantly changing due to natural processes and human activities such as urban expansion, deforestation, and agricultural land use. Traditional static maps provide only a snapshot of land cover at a given moment, which limits their usefulness for monitoring ongoing changes. Time series land cover mapping enables continuous observation and analysis of landscape evolution over time, helping researchers and policymakers make informed decisions about environmental sustainability, resource management, and urban planning.

1.2. Advantages of Dynamic Monitoring

- **Real-Time Insights:** Time series mapping helps detect trends and abrupt changes, providing real-time information on land use patterns.
- **Improved Accuracy:** By analyzing data over multiple time points, temporal inconsistencies and classification errors in individual snapshots can be minimized.
- **Climate and Environmental Monitoring:** It allows tracking seasonal vegetation cycles, forest degradation, and urban expansion with high precision.
- **Disaster Response:** Dynamic monitoring is crucial for understanding the impact of natural disasters such as wildfires, floods, and droughts.

2. Setup and Configuration

2.1. Required Libraries and Dependencies

Library/Extension	Description	Installation Command
geemap	Python library for geospatial analysis and working with Google Earth Engine (GEE).	`pip install geemap`
earthengine-api	API for interacting with Google Earth Engine from Python.	`pip install earthengine-api`
jupyter-leaflet	Provides interactive maps for Jupyter Notebooks using Leaflet.js.	`pip install jupyter-leaflet`
ipyleaflet	Interactive Leaflet.js maps for Jupyter notebooks.	`pip install ipyleaflet`
ipywidgets	For creating interactive widgets in Jupyter notebooks.	`pip install ipywidgets`
nodejs	Required for Jupyter extensions to work.	`conda install -c conda-forge nodejs`

widgetsnbextension	Jupyter extension for supporting interactive widgets in Jupyter notebooks.	`jupyter nbextension enable --py --sys-prefix widgetsnbextension`
jupyterlab-manager	JupyterLab extension for managing widgets.	`jupyter labextension install @jupyter-widgets/jupyterlab-manager`
jupyter-leaflet extension	JupyterLab extension to enable jupyter-leaflet to work in JupyterLab.	`jupyter labextension install jupyter-leaflet`
jupyterlab_widgets	Provides the interactive widget support for JupyterLab.	`pip install jupyterlab_widgets`

2.2. Setting up Google Earth Engine (GEE)

Google Earth Engine was set up using the following steps:

- **Create a GEE Account:** Registered at https://earthengine.google.com/
- **Linked the Project:** Created and connected a new project named “ee-mohammadsobbahi” to enable cloud-based computations.
- **Authenticated API Access:** Used earthengine authenticate to grant access to GEE from local and remote environments.

2.3. Challenges in Authentication

During the authentication process, several issues were encountered:

- **Login Authentication Error:** When navigating to Google's authentication page and returning to complete the process, an "Invalid Account" error occurred. After multiple attempts and account validation, the authentication was finally successful.
- **Missing Project Error:** Initially, authentication failed due to the absence of a linked project. This was resolved by creating a new Google Cloud project.
- **Project Registration in GEE:** Even after creating the project, GEE did not recognize it immediately. Finding the correct way to register the project in GEE was challenging, as there were no direct links or clear instructions available. After extensive searching, the issue was finally resolved through trial and error.

Despite these challenges, the authentication process was successfully completed, enabling full access to Google Earth Engine.

3. Data Sources and Processing

3.1. ESA WorldCover Dataset

The ESA WorldCover Dataset serves as a foundational global land cover product with consistent 10-meter resolution. This dataset, developed by the European Space Agency, categorizes the Earth's surface into 11 distinct classes including trees, water bodies, built-up areas, and various vegetation types. Its static nature provides a reliable baseline for 2020, making it particularly valuable for comparative studies and validation purposes. The dataset's global coverage ensures consistency in classification methodology across political boundaries, though some regional variations in accuracy may occur due to differences in landscape complexity.

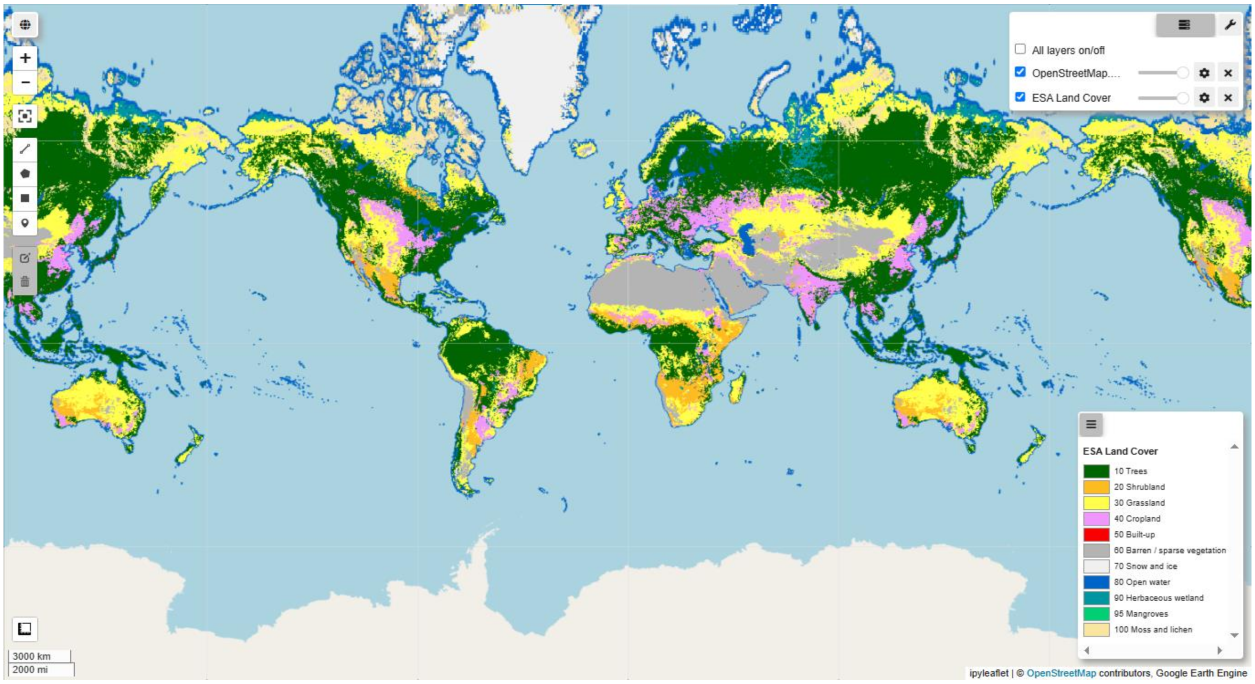


Figure 1. ESA WorldCover Land Classification of Lebanon (2020, 10m Resolution)

3.2. ESRI Global Land Cover Dataset

The ESRI Global Land Cover Dataset offers another 10-meter resolution product with a slightly different classification scheme of 10 land cover types. Developed through a partnership between ESRI and Microsoft, this dataset emphasizes practical applications in agriculture and urban planning. Its classification system proves particularly effective for identifying croplands and bare ground surfaces, with specialized spectral indices optimized for these categories. The dataset's mosaic format ensures seamless coverage across large areas, though users should note it represents a composite time period rather than a single snapshot.

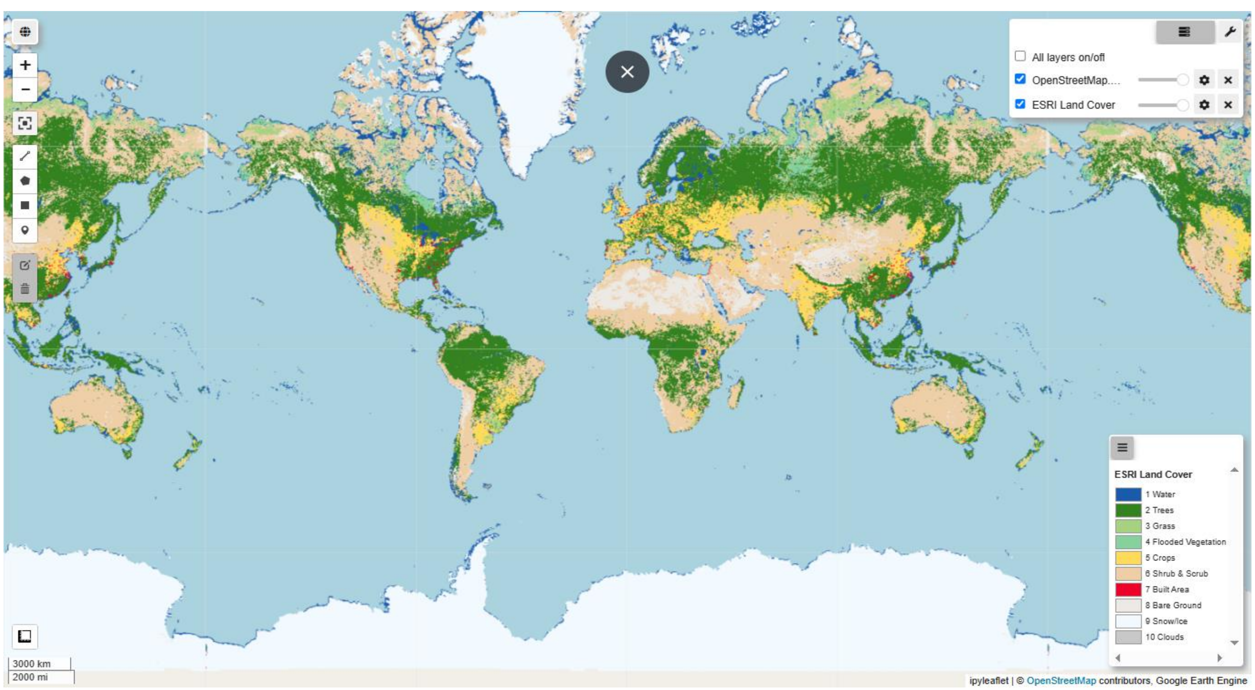


Figure 2. ESRI Global Land Cover Map of Lebanon (10m Resolution)

3.3. Dynamic World Dataset

Google's Dynamic World Dataset represents a paradigm shift in land cover monitoring through its near real-time updating capability. While nominally offering 10-meter resolution, the actual spatial detail varies with Sentinel-2 observation availability and cloud cover conditions. The dataset's 9-class system focuses on dynamic landscape elements, with particular strength in detecting changes in built environments and water bodies. Its temporal resolution - producing weekly updates - enables unprecedented monitoring of rapid land cover changes, from urban expansion to seasonal vegetation cycles. This makes Dynamic World uniquely valuable for applications requiring current land cover information rather than historical baselines.

4. Implementation

4.1. Defining Region of Interest (Lebanon)

The study focuses on Lebanon's diverse landscapes by defining a precise bounding box (35.05°-36.65°E, 33.0°-34.72°N) that captures the country's complete territory. This rectangular region encompasses Lebanon's varied geography from the Mediterranean coastline to the Mount Lebanon range and the Bekaa Valley. The bounding box approach ensures complete national coverage while maintaining simple geometric parameters for analysis. Centering the map at zoom level 8 provides an optimal balance between national overview and sufficient detail for regional analysis. This standardized geographic framework allows for consistent comparison between different land cover products and temporal analyses.

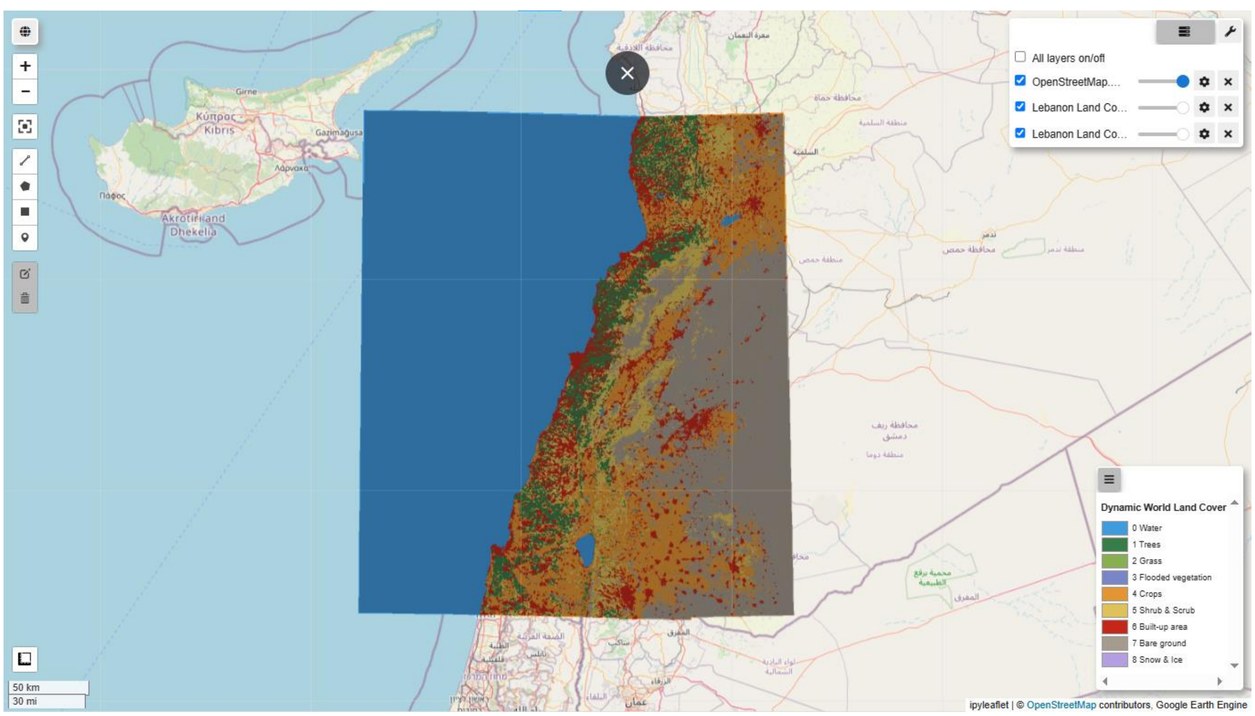


Figure 3. Lebanon's Bounding Box (35.05°E–36.65°E, 33.0°N–34.72°N)

4.2. Time Period Selection

The analysis period from March 2020 to March 2023 was strategically chosen to capture three complete annual cycles of Lebanon's distinct seasons. Beginning in late March ensures inclusion of the spring growing season while avoiding winter cloud cover that might obscure baseline observations. The three-year duration enables identification of both short-term changes and longer-term trends, while remaining manageable in terms of data volume and processing requirements. This temporal window also coincides with significant recent developments in Lebanon's landscape, including urban expansion and vegetation changes related to economic fluctuations. The consistent annual timeframe (March-March) controls for seasonal variations when comparing interannual changes.

4.3. Land Cover Classification and Visualization

Effective visualization of land cover data requires careful parameter selection to maximize interpretability. For ESA WorldCover, the straightforward single-band visualization highlights the discrete classification scheme. Dynamic World employs a more sophisticated color palette that intuitively represents different land cover types - blues for water, greens for vegetation, and reds for urban areas. These visualization schemes were chosen to facilitate quick interpretation while maintaining scientific accuracy. The map layers are designed with appropriate transparency and overlay options to enable comparison with underlying base maps, providing geographic context to the classification data.

5. Comparative Analysis

5.1. ESA vs. Dynamic World Comparison

The comparative analysis between ESA WorldCover and Dynamic World reveals fundamental differences in their design philosophies and applications. ESA's static 2020 snapshot provides a consistent baseline unaffected by seasonal variations, making it ideal for long-term change detection when compared with historical or future data. Dynamic World's constantly updated classifications excel at capturing recent changes and seasonal patterns, though with potentially less interannual consistency. The split-map visualization clearly shows how urban areas appear more extensive in Dynamic World due to its temporal sensitivity, while ESA provides a more conservative estimate of persistent built-up areas. This comparison highlights the importance of selecting datasets appropriate for specific monitoring objectives.

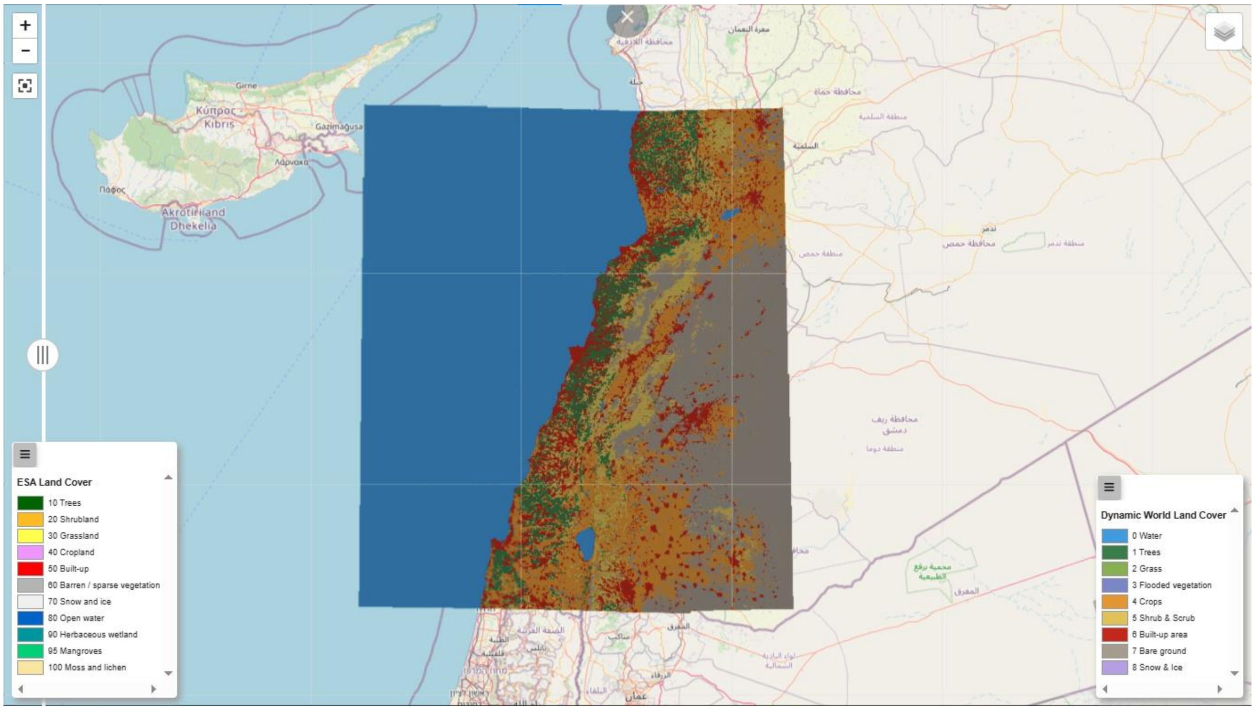


Figure 4. Comparative Land Cover Classification: ESA WorldCover

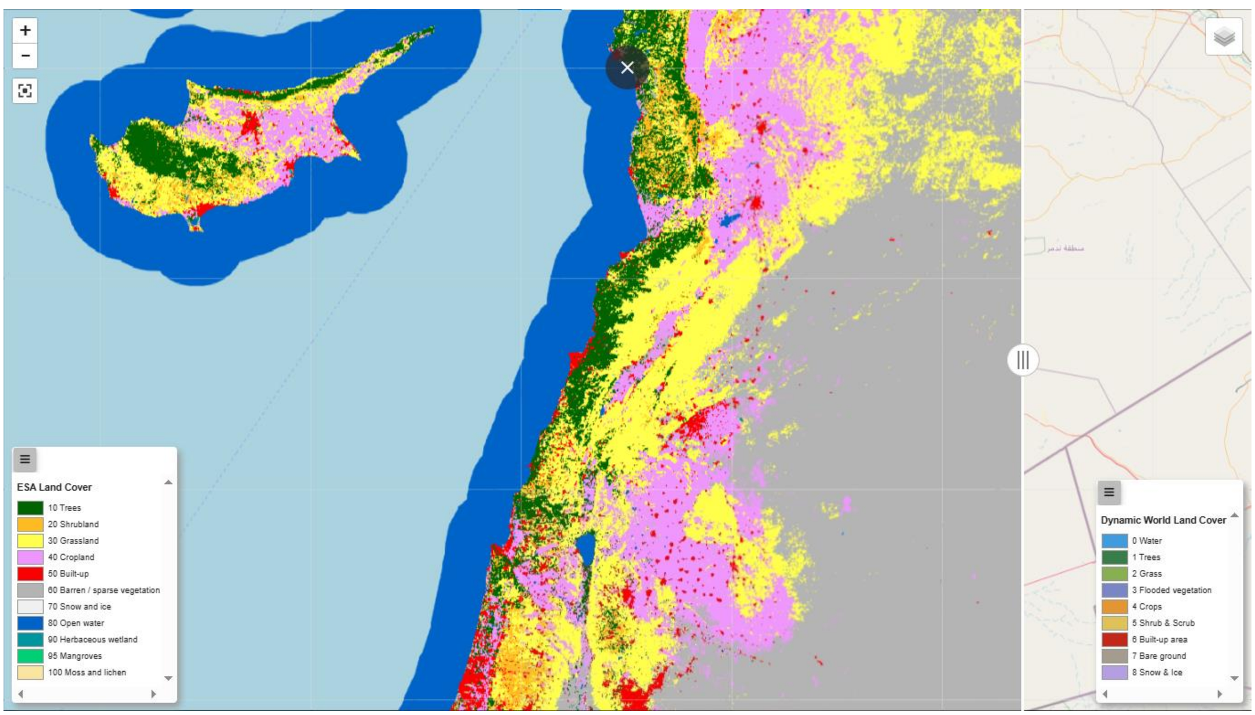


Figure 5. Comparative Land Cover Classification: Dynamic World

5.2. ESRI vs. Dynamic World Comparison

When comparing ESRI's global land cover product with Dynamic World, distinct strengths emerge for different applications. The ESRI dataset demonstrates superior performance in agricultural land identification, with clearer differentiation between crop types and natural vegetation. Dynamic World shows greater sensitivity to recent land use changes, particularly in peri-urban zones and water bodies. The side-by-side comparison reveals how ESRI's composite approach smooths out seasonal variations that Dynamic World explicitly captures. For Lebanon's diverse landscapes, this comparison proves particularly valuable in agricultural regions like the Bekaa Valley, where the datasets may classify the same fields differently based on observation timing and classification algorithms.

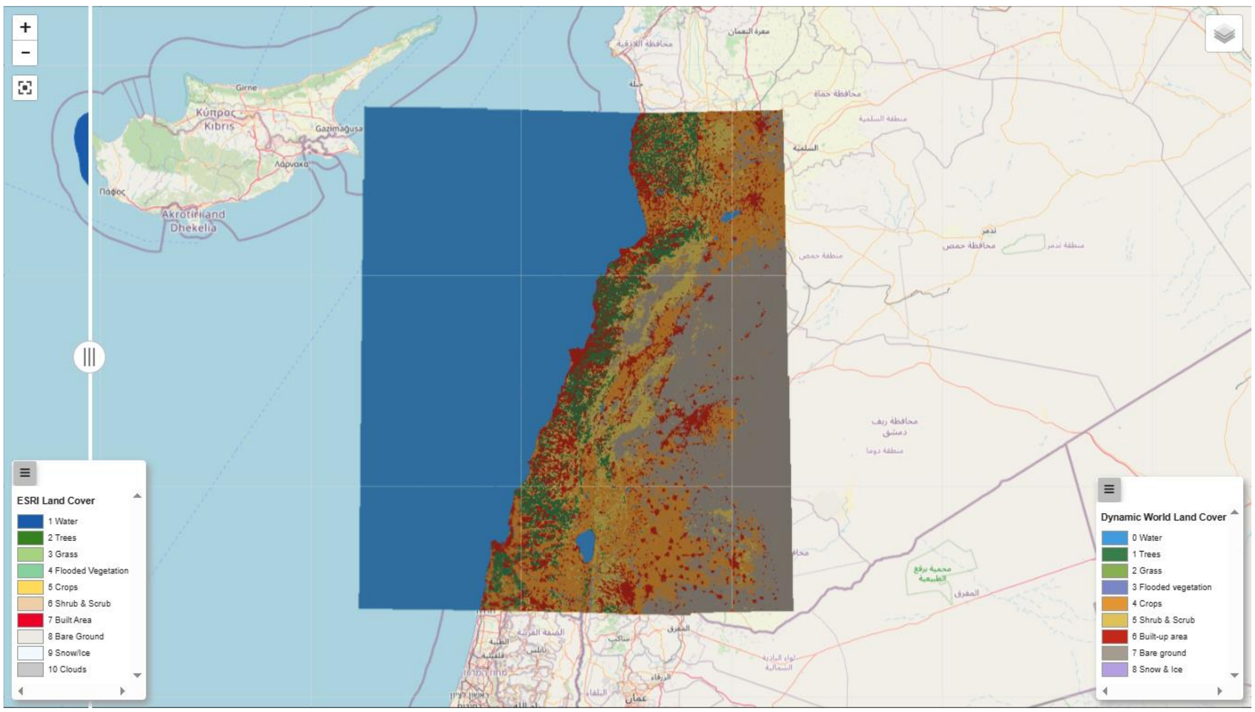


Figure 6. Comparative Land Cover Classification: ESR Land Cover

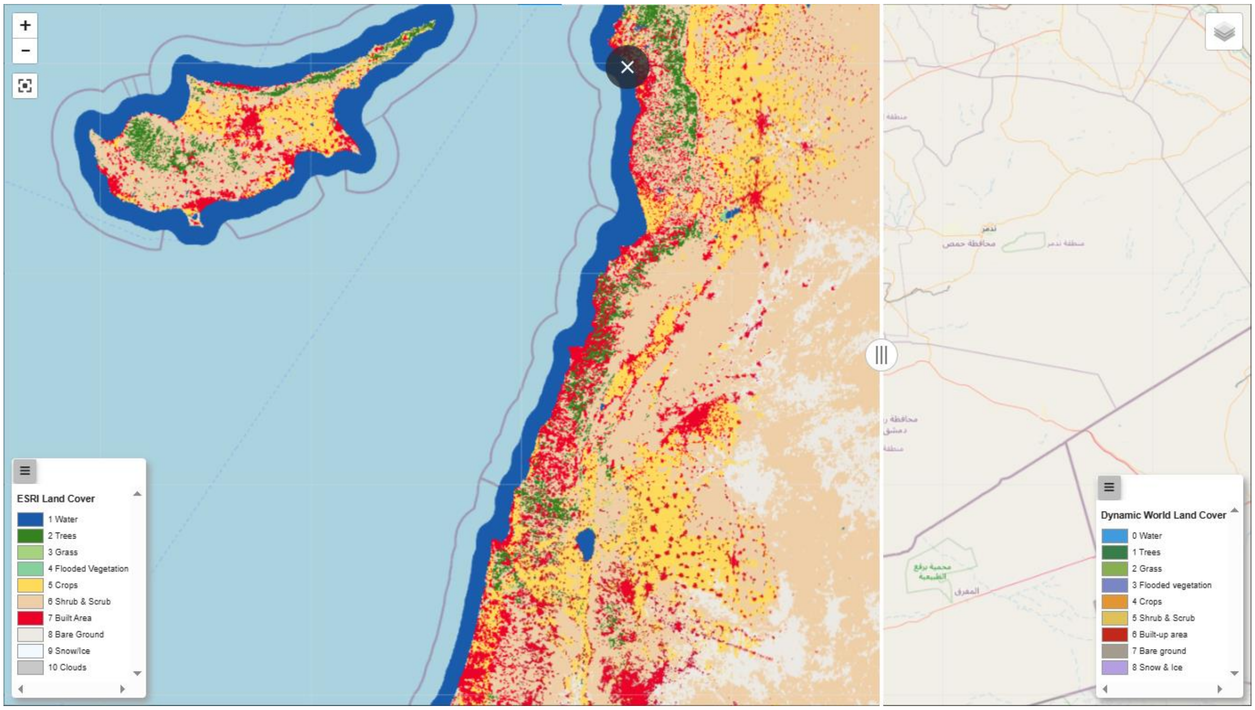


Figure 7. Comparative Land Cover Classification: Dynamic World Land Cover

6. Time Series Land Cover Change Analysis

6.1. Generating Time-Series Land Cover Maps

The time-series generation process leverages Dynamic World's temporal capabilities to create a sequence of land cover snapshots. This approach transforms static land cover analysis into a dynamic monitoring system, revealing progression of changes across seasons and years. The time-series properly handles missing data periods due to cloud cover by utilizing Sentinel-2's best available observations. For Lebanon's climate with distinct wet and dry seasons, this produces a clear phenological signal in vegetation classes while maintaining consistent classification of permanent features like urban areas. The resulting data cube supports both visual analysis and quantitative change detection metrics.

6.2. Visualizing and Analyzing Changes

Interactive visualization tools transform the time-series data into an exploratory analytical environment. The time slider interface allows users to animate land cover changes or focus on specific periods of interest. Pixel-level inspection capabilities enable precise identification of change timing and patterns, particularly useful for detecting gradual urban expansion or sudden land conversion events. The Beirut urban growth analysis demonstrates this capability, clearly showing expansion patterns through the red/green change visualization. These tools collectively support both broad-scale trend analysis and site-specific investigations, making the system valuable for researchers and policymakers alike.

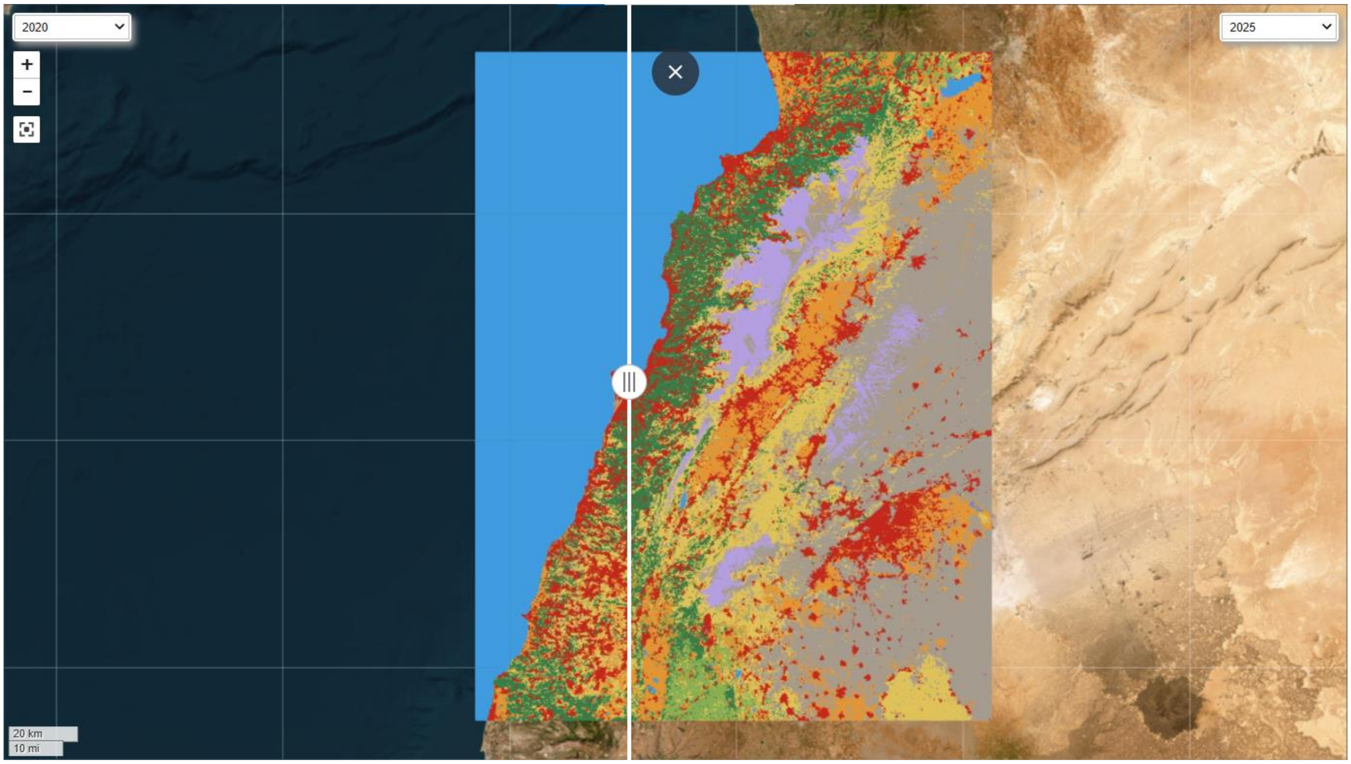


Figure 8. Lebanon Timeseries Land Cover 2020

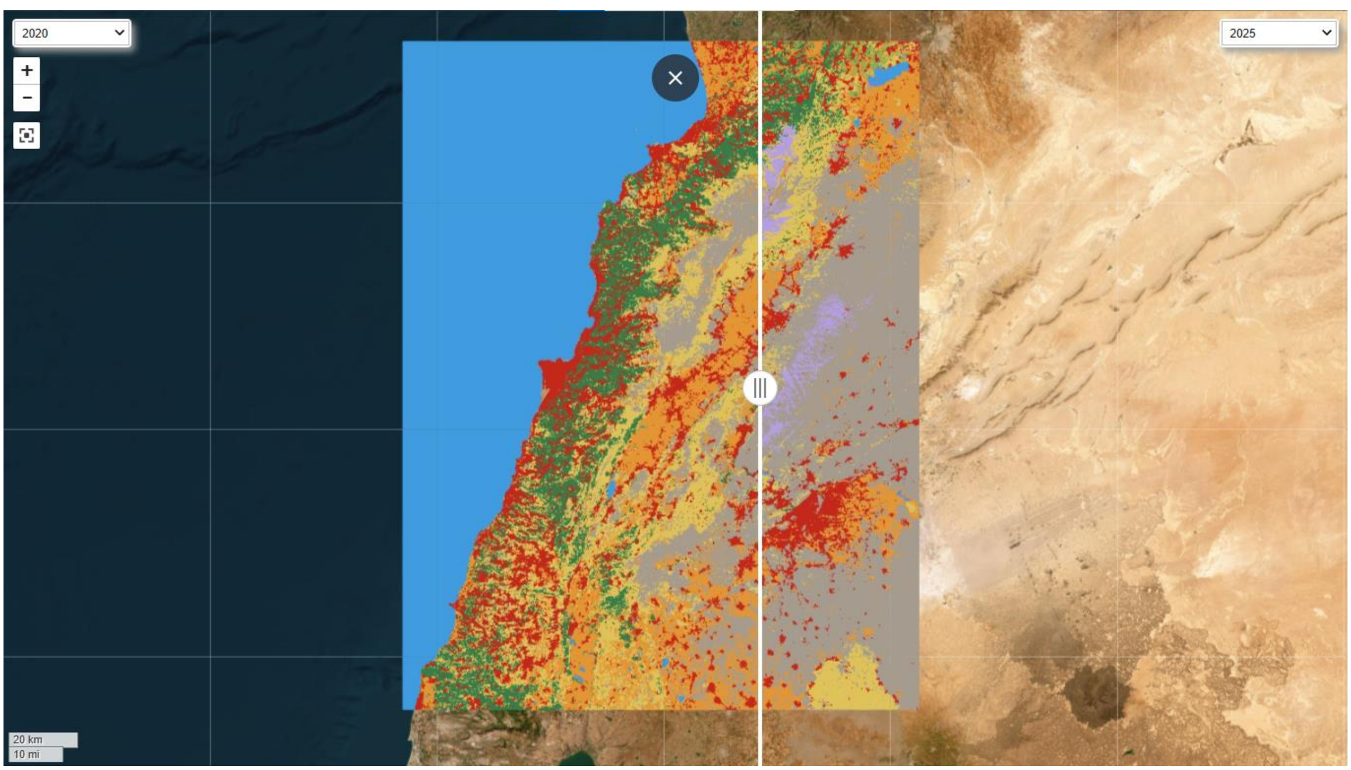


Figure 8. Lebanon Timeseries Land Cover 2025

Conclusion

This project developed a comprehensive land cover monitoring system for Lebanon, combining ESA and ESRI baseline maps with Dynamic World's near-real-time data. The analysis revealed key trends like Beirut's urban growth and seasonal vegetation changes, using interactive time-series tools and comparative visualizations. The integrated approach provides both historical reference points and dynamic change detection, creating a powerful tool for sustainable land management.

GitHub Repository

The complete source code and detailed documentation for the project are available on GitHub. You can access the repository at:

- Lebanon_LandCover_TimeSeries (main branch).

References

[1] Sentinel Hub, "SI_LULC_pipeline.ipynb," eo-learn, GitHub, 2025.

[2] "Land Cover Classification with Earth Engine," geemap.org, 2022.

[3] "Land Cover Time Series Analysis with Earth Engine," geemap.org, 2022.