Final Year Project Report

Full Unit – Final Report

The use of Data Science in “SDG-Goal 15”

Antony Sakr

Supervisor: Dr Ziad Abdallah

Co-Supervisor: Dr Mira Abboud

الجامعة اللبنانية
كلية الإعلام

Department of Data Science

Lebanese University

Declaration

This report has been prepared on the basis of my own work. Where other published and unpublished source materials have been used, these have been acknowledged.

Word Count: 8619

Student Name (ID): Antony Sakr (7517)

Date of Submission: 10 July 2022

Signature: Anthony

Contents

Abstract

We will try to calculate the SDG goal 15.3.1 which is Proportion of land that is degraded over the total land area by using the Google Earth Engine and remote sensing, we will use the JavaScript programming language, and we will finally visualize and analyze the results.

I. Introduction

The National Statistical Offices (NSO) are currently introducing the use of Artificial Intelligence to derive automated maps to compile SDG indicators from satellite images. Goal 15: Life and Land is an important goal for NSO. A prototype application will be developed within the scope of this project. It will be focused compilation of the relevant indicators by using data science knowledge.

The project's aim is to leverage modern methods and techniques, such as APIs and web applications, to undertake remote sensing work linked to cities and communities' indicators, rather than traditional work. These methods will assist us in reducing the number of resources required as well as the amount of time spent.

This project is constructed from many parts, in the first part, we will give a state-of-the-art review to introduce what remote sensing is, also we will talk about many important points in remote sensing, like what is its importance, its types (passive and active sensors) specifically about MODIS and Sentinel-2, in what we can use it, its data limitations, and finally a quick conclusion.

In the second part, we will talk about SDG goal 15 “life on land”, we have two indicators in this part to talk about are SDG 15.1.1 (Forest area as a proportion of total land area) and SDG 15.3.1 (Proportion of land that is degraded over the total land area). For each sub-indicator, we will talk about why each indicator is important for our world, and how to measure and monitor the indicator. Also, we will give an abstract about the metadata of our dataset, and finally a quick conclusion.

The third part is the Google Earth Engine part, we will start with a definition and what is it, then how can we use it, how we can access it, its cons and pros, and NDVI index calculation in GEE to calculate our SDG goal 15.3.1, and finally a conclusion.

finally, we will talk about the schema that we should follow to calculate the SDG 15.3.1 by using two kinds of satellites (MODIS and Sentinel-2), then we will show the resulting images and analyze them.

II. Remote sensing

A. Definition [1]

Remote sensing is a form of geospatial technology that samples emitted and reflected EM (electromagnetic) radiation from the Earth’s terrestrial, aquatic, and atmospheric ecosystems so you can discover and monitor the physical traits of a place without making physical contact. This technique of data collection normally consists of aircraft-based and satellite-based sensor technologies, which can be classified as passive sensors or active sensors.

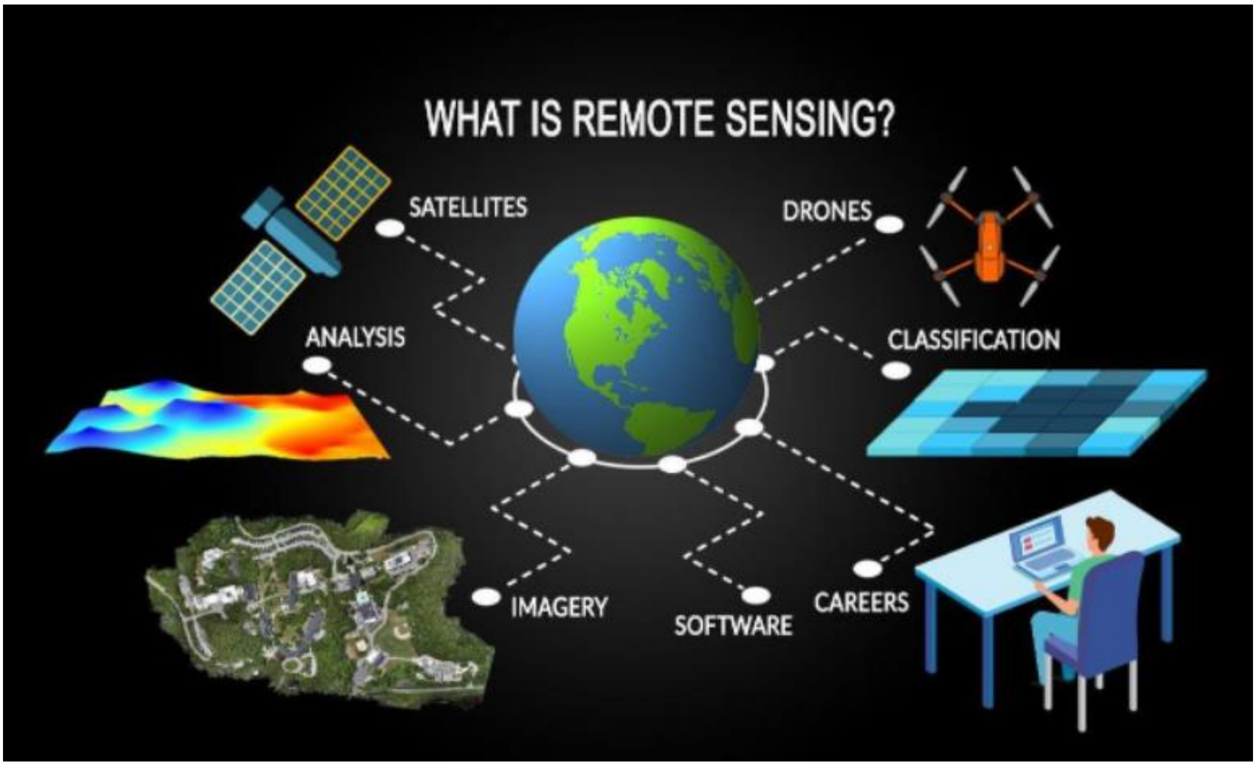

Figure 1 What is remote sensing. Adopted from “GISGeography”

B. The importance of remote sensing [2]

Remote sensing makes it possible to gather data from risky or inaccessible regions (like the photos below for an active volcano). It replaces slower, expensive data collection on the ground, offering speedy and repetitive coverage of extremely massive regions for everyday applications, starting from weather forecasts to reviews on natural disasters or weather change.

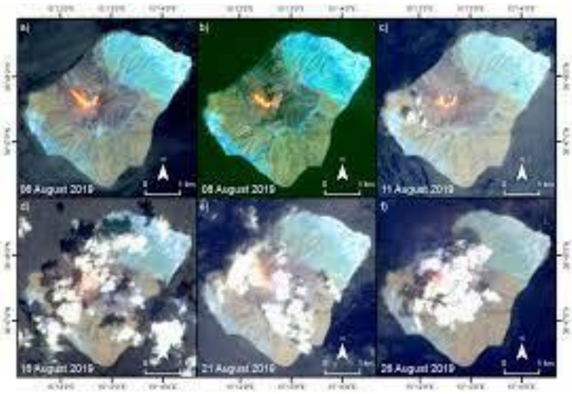

Figure 2 Active volcano. Adopted from "BioMed Central"

Remote sensing is likewise an unobstructed method, permitting users to gather data and perform data processing and GIS analysis from distance without disturbing the target object or region. Monitoring floods and forest fires, deforestation, earthquakes, volcanos, and chemical concentrations are only a few cases in which geospatial remote sensing affords a worldwide perspective and actionable insights that will otherwise be unattainable.

C. Passive and Active sensors [3]

The main difference between active and passive sensors is that an active sensor generates its very own signal that's measured when reflected back by the Earth's floor, but the passive sensor measures sun energy that is both reflected or emitted from features on the Earth's floor.

1. Passive sensors

Although most passive sensors perform in the visible (RGB) and infrared portions of the EMS (electromagnetic spectrum), there are also a few passive microwave sensors in use that measure some of parameters which includes sea surface temperature, atmospheric water vapor, rainfall, soil moisture and wind speed.

A benefit of passive sensors is that most rely on the Sun's energy to light up the target and consequently don't want their own energy source so in general they're easier and simpler instruments. A limitation for most passive optical sensors is they require daylight to perform, despite the fact that there are a few sensors that record night time lights and others that record energy emitted from the Earth's surface. Because most of those sensors perform withinside the visible and infrared wavelengths, they're negatively affected by cloud cover, climate and weather, and because sunlight is basically reflected from the top of a feature, such as a forest, it isn't possible to see below a canopy to measure vegetation structure. To obtain this type of information it's obligated to use active sensors.

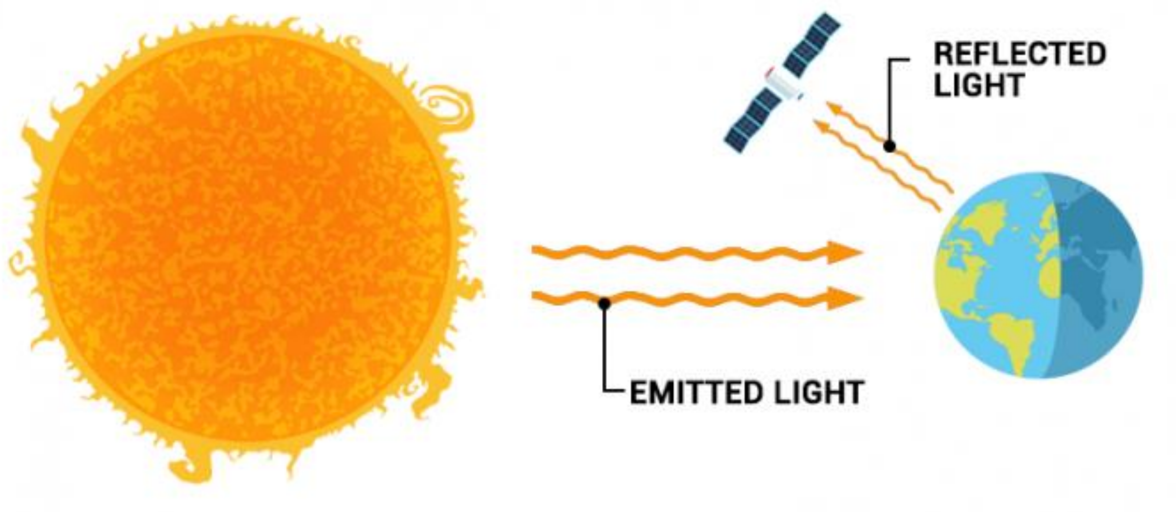

Figure 3 Passive sensor. Adopted from "EARTHDATA-NASA"

2. Active sensors

Active sensors, emit their very own energy to light up a target and are constituted of a signal generator and receiver. They take the measurements of the power of the returned signal and the time delay between emit and receiving the pulse. These sorts of data are used to provide an explanation for vegetation structure. The two main types of active sensors are radar and lidar.

Radar is an abbreviation for "radio detection and ranging". Radar systems operate in the long-wavelength microwave part of the EMS and thus are unaffected by the weather because it uses electromagnetic radio waves to determine the angle, range, or velocity of objects. So, they are considered all-weather systems, so it is a big advantage that makes them very useful in areas that are continuously covered in the cloud or are otherwise not easily acquired by optical remote sensing techniques.

Lidar an abbreviation for "light detection and ranging", is one of the remote sensing methods, to measure earth ranges it uses light in the form of a pulsed laser. These light pulses combined with other data recorded by the airplanes or the helicopters generate precise 3-dimensional information about the characteristics of the earth's surface and its shape.

A lidar mainly includes 3 things are a laser, a scanner, and a specialized GPS receiver. Airplanes and helicopters are the most commonly used platforms to gather lidar data over wide areas. There are only 2 forms of lidar: topographic and bathymetric. First, Topographic lidar uses a NIR (near-infrared) laser to map the land, while bathymetric lidar uses water-penetrating green light to also measure seafloor and riverbed elevations.

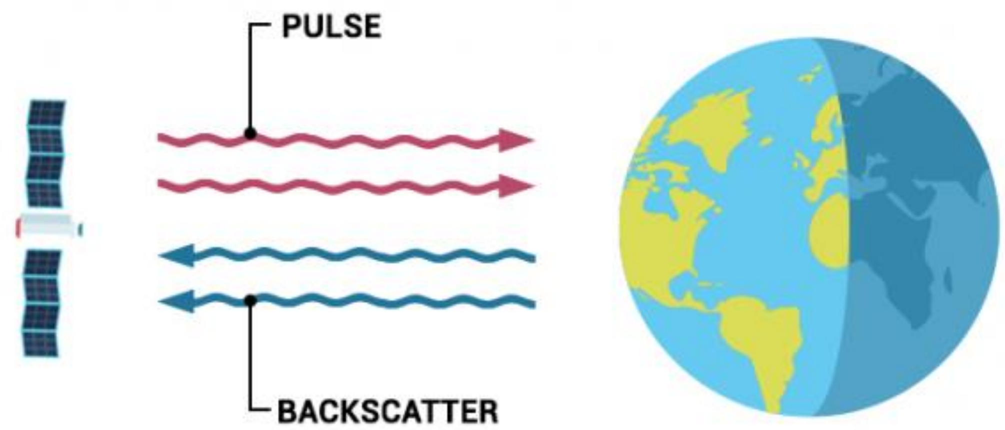

Figure 4 Active sensor. Adopted from "EARTHDATA-NASA"

D. Definition of Sentinel-2 and MODIS

1. Sentinel-2 definition [4]

Sentinel-2 has been developed and is being operated by the European Space Agency, and it is an Earth observation mission from the Copernicus Programme that systematically acquires optical imagery at high spatial resolution from 10 m to 60 m over land and coastal waters. we have two satellites named Sentinel-2A and Sentinel-2B, also the Sentinel-2C is currently undergoing testing to be launched in 2024. The satellites support a wide range of applications like land cover classification (as our topic in this project), also like water quality, and others.

The Sentinel-2 satellite has many characteristics like it has multi-spectral data with 13 bands in the visible, near-infrared, and short-wave infrared parts of the spectrum, 290 km field of view, free data access, revisiting every 10 days under the same viewing angles, also, it provides information on agricultural and forestry practices and for helping manage food security. Satellite images will be used to determine various plant indices such as our index NDVI and many others. Finally, Sentinel-2 is mainly used to monitor changes in land cover and to observe forest characteristics all over the world. also, the same satellite images can be used to provide information on pollution in lakes, coastal waters, and disaster mapping.

Figure 5 Sentinel-2 image Adopted from the EUROPEAN SPACE AGENCY

2. MODIS definition [5]

MODIS is the abbreviation of Moderate Resolution Imaging Spectroradiometer, it is an extensive program using sensors on two satellites that each provide complete daily coverage of the earth, and these two satellites are first, on board the Terra (EOS AM) satellite, launched by NASA in 1999 and the second on board the Aqua (EOS PM) satellite, launched in 2002. By using MODIS sensor, it is possible to obtain images in the morning hours from Terra, also we can have them in the afternoon, by using Aqua for any particular location in this world. also, night time images are available, but only in the thermal range of the spectrum. There are many advantages of using MODIS sensor and the most important one is the intensive coverage MODIS sensor provides, in time and space, so MODIS has captured as many as four shots of every place on Earth, every day for the last 20 years. So, by using MODIS we can gather the data we want to calculate the NDVI index for Lebanon.

Figure 6 MODIS image Adopted from GISGeography

E. What is remote sensing used for? [2]

We can use remote sensing technology in a variety of ways, in many different fields like weather, agriculture, forestry, biodiversity, surface changes, and others. To better understand it, here are many real examples of remote sensing: MODIS satellites use thermal sensing and mid-infrared sensing to display active volcanoes; large forest fires can be mapped from space, allowing rangers to see a much larger area than from the ground to help people; as well, tracking the growth of a city and changes in forests or farmland over several years or decades; also, to understand better the way to manage ocean resources; and to assess the effects of a natural catastrophe and create catastrophe reaction techniques to be used before and after a dangerous event; also to decrease the harm that urban growth has on the environment and help determine the excellent way to protect natural resources. Tracking clouds to help predict the weather; the main components of remote sensing in agriculture are Irrigation and soil moisture tracking and management; Remote sensing satellites offer before- and after- remote sensing pictures that allow us to quantify post-earthquake damage, which gives important data for rescue workers. Finally, we can use remote sensing in GIS, and GIS is the abbreviation of "Geographic Information System" and it's a system designed to manage, store, capture, analyse, manipulate, and present geographic or spatial data, remote sensing is the main source of spatial data. Remote sensing and GIS work collectively to gather, store, analyse and visualize data from virtually any geographic position on Earth.

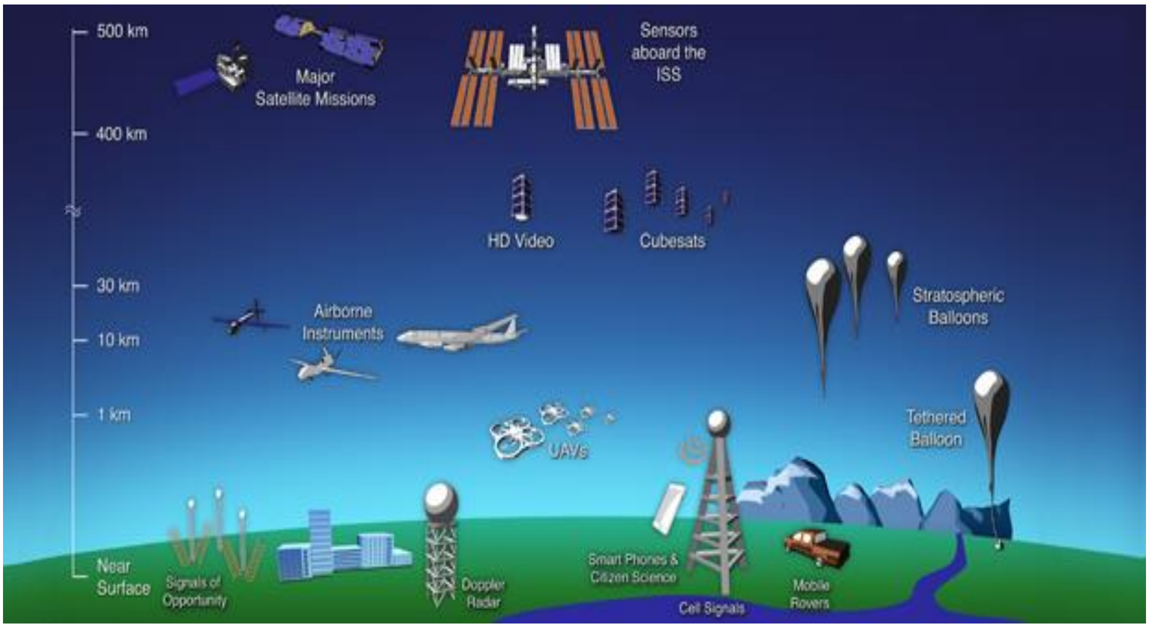

Figure 7 The typical altitudes of major remote sensing technologies. Adopted from "GISGeography"

F. Remote sensing data limitations [2]

Remote sensing is controlled by persons that make important choices concerning which sensors must be used to gather data and when, resolution specifications for the gathered data and sensor calibration, and the selection of the platform that will carry the sensor, all of which expose this technique to a certain degree of human error.

Inaccuracy can also be introduced by the electromagnetic spectrum radiation emitted from powerful active remote sensing systems, which may be intrusive and have an effect on the target phenomenon being investigated. Remote sensing devices may also make contributions inaccurate, un-calibrated data if the hardware device becomes un-calibrated. There can also be price related limitations.

G. Conclusion

Remote sensing is the gathering of data concerning the earth that doesn't need to contact the floor or contact the item that we study. With the support of remote sensing, we can be capable of getting accurate data about the earth which includes landscapes, forests, water resources, and oceans. This data helps the researchers in their research about the earth's components concerning its sustainable management and conservation. So, remote sensing offers the simplest method to detect the degradation of the ecosystem services, which is the focus of the Sustainable Development Goals, especially the SDG goal 15(life on land).

III. SDG 15- Life on land

A. Definition [6]

"Life on land" is the title of the Sustainable Development Goal 15, and this goal is one among seventeen goals established by the United Nations, and therefore the official wording is: **"Protect, restore and promote sustainable use of terrestrial ecosystems, sustainably manage forests, combat desertification, and halt and reverse land degradation and halt biodiversity loss".**

Figure 8 SDG 15. Adopted from "Gothenburg Centre for Sustainable Development"

Goal 15 has twelve targets to be achieved by 2030. during this article, we are going to discuss two targets:

- **Target 15.1**
- **Target 15.3**

B. Target 15.1: [7]

- **By 2020, ensure the conservation, restoration, and sustainable use of terrestrial and inland freshwater ecosystems and their services, in particular forests, wetlands, mountains, and drylands, in line with obligations under international agreements.**

Figure 9 Target 15.1. Adopted from "ICCROM"

-Indicator: 15.1.1: Forest area as a proportion of total land area

1. -Indicator 15.1.1

Why is this indicator important?

Forests are crucial to the survival of our planet and climate extrude mitigation. Forests offer oxygen, keep carbon, are biodiversity reservoirs and critical for humanity. Forest area as a percent of the general land area may be used as a proxy for studying the area to which forests are being conserved or restored, however, it's an indicator of whether or not forests are being sustainably managed. The changes in forest area will help to identify all unsustainable practices by human in the agricultural sector and in the forestry.

For forest policy and planning within the context of sustainable development, the availability of accurate data on a country's forest area is crucial, so this indicator will help to guide all kinds of investment. Adequate forest resources ensure social, financial and environmental balance also adequate sustainable improvement for the next generations.

Measure and monitor the indicator 15.1.1

To measure this indicator, first, we will need to define the forest. so, the definition of the forest according to the UNCCD is "0.5 hectares of land with trees higher than 5 meters and the canopy should cover more than 10%, and it doesn't include land used for agriculture"

Data collection on forest regions draws from the Global Forest Resources Assessment (FRA) which is performed every 5 years. All data are supplied to FAO (Food and Agriculture Organization) by countries in the shape of a standard country report via a web platform. FRA 2020 is the latest evaluation covering 236 nations and territories. Data on the land region is gathered via the FAO Questionnaire on Land Use, Irrigation, and Agricultural Practices that is sent yearly to 205 nations and territories. Although access to remote sensing information has improved, forest regrowth or forest with a low canopy cover density can't be detected every time.

C. Target 15.3: [7] [8]

- **By 2030, combat desertification, restore degraded land and soil, including land affected by desertification, drought, and floods, and strive to achieve a land degradation-neutral world.**

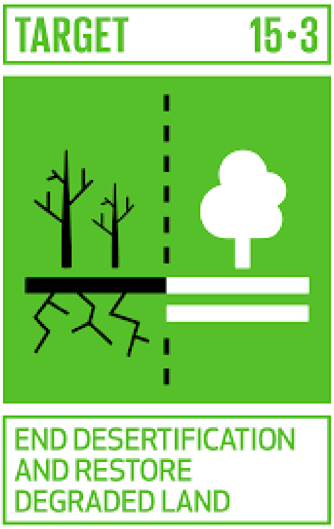

Figure 10 Target 15.3. Adopted from "ICCROM"

-Indicator 15.3.1: Proportion of land that is degraded over total land area

1. -Indicator 15.3.1

Why is this indicator important?

Food production, or the expansion of infrastructure increase pressure on arable lands and forestry. On the other hand, land degradation has terrible consequences on ecosystems, farm output, and the amount and quality of water resources. As a result, food insecurity may also rise and facilitate urbanization and migration. So, Assessing the quantity of land degradation is crucial for the improvement of plans of action to repair land degradation.

Measure and monitor the indicator 15.3.1

The indicator makes use of a binary classification (degraded or not) drawing on the 3 sub-indicators that assess trends in land cover, in land productiveness, and in carbon stocks above and underground. Any bad change or significant decrease in one of the 3 will led to land degradation.

So, what does land degradation mean? **“Land degradation is defined as the reduction or loss of the biological or economic productivity and complexity of rainfed cropland, irrigated cropland, or range, pasture, forest, and woodlands resulting from a combination of pressures, including land use and management practices”.** This definition was adopted by and is used by the 196 countries that are Party to the **UNCCD** (United Nations Convention to Combat Desertification).

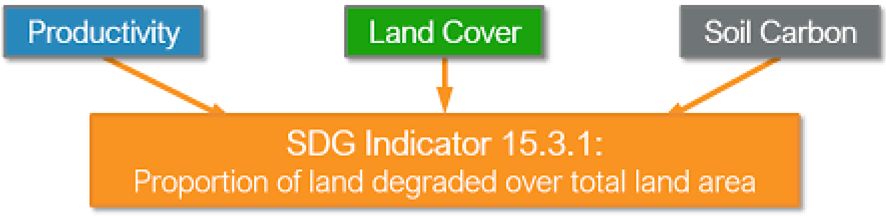

Figure 11 The 3 sub-indicators of SDG 15.3.1. Adopted from "Trends. Earth"

land cover

is defined as the physical cover of the earth's surface, which is for example forests, all vegetation types, sea, all kinds of water, and all the infrastructure made by human. This indicator has 2 main functions for SDG 15.3.1, are: first, changes in land cover may cause land degradation when there is a loss of ecosystem services. Second, a land cover classification system can be used to disaggregate the other 2 sub-indicators (land productivity and soil carbon).

And, LCC is equal to: **LCC = C+PV Recurring – PV Residual Value**

LCC is the abbreviation of life cycle cost. C is the initial construction cost. PV recurring is equal to the present value of all recurring cost, and PV residual value is equal to the present value of residual value.

Land productivity

refers to the overall productivity of a system and can be found in an equation where the (NPP) Net Primary Production, is equal to the Gross Primary Production (GPP), minus the Respiration (R). **(NPP = GPP – R).**

The NPP is the overall efficiency of the plants in the ecosystem. This sub-indicator calculates the changes in the health and productive capacity of the land, where the declining of NPP is often a characteristic of land degradation.

Carbon stock

is the amount of carbon in a ‘pool’: a reservoir that has the capacity to release carbon and is constituted of above- and below-floor biomass, soil organic carbon, and dead organic matter.

SOC is an indicator of overall soil quality, is the abbreviation of Soil organic carbon, it is the carbon that stays in the soil after the partial decomposition of any material produced by living organisms. It is a main element of the global carbon cycle via the atmosphere, vegetation, rivers, and the ocean.

Formula: Organic matter (%) = Total organic carbon (%) x 1.72

1.72 is a commonly used conversion factor to convert organic carbon to organic matter. But this factor is not always fixed, it depends on the type of organic matter, soil type and soil depth, the maximum value of this factor is 2.5.

Convert percentage (%) to weight for a given area:

SOC stock in tonnes of carbon per hectare (tC/ha) = (soil organic carbon %) x (mass of soil in a given volume)

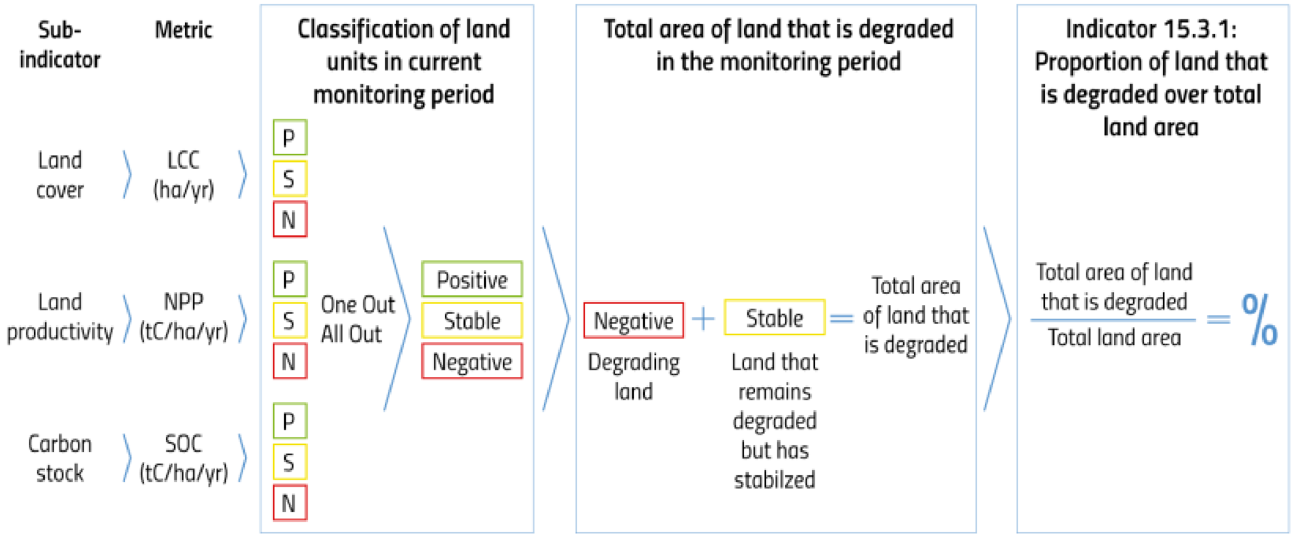

Figure 12 Steps to derive the SDG indicator 15.3.1 from the sub-indicators. Adopted from “UNCCD"

D. Metadata about our Dataset [9]

National data on the three sub-indicators are and can be collected through existing sources (e.g., databases, maps, reports), including participatory inventories on land management systems as well as remote sensing data collected at the national level. Datasets that complement and support existing national indicators, data and information are likely to come from multiple sources, including statistics and estimated data for administrative or national boundaries, ground measurements, Earth observation and geospatial information. A comprehensive inventory of all data sources available for each sub-indicator is contained in the Good Practice Guidance for SDG Indicator 15.3.1.

The most accessible and widely used regional and global data sources for the sub-indicator "Land productivity" are briefly described here.

Land productivity data represented as vegetation indices (i.e., direct observations), and their derived products are considered the most independent and robust option for the analyses of land productivity, offering the longest consolidated time series and a broad range of operational data sets at different spatial scales. The most accurate and reliable datasets are available in the:

(1) MODIS data products

(2) Sentinel-2 data products

Data is gathered every 4 years since 2018 by UNCCD. The Good Practices Guidance for SDG Indicator 15.3.1, gives a detailed overview of computation and monitoring. And as a data compiler, we will use the FAO dataset.

MODIS bands

Band	spectral range [nm]	Description
1	620–670	red
2	841–876	NIR
3	459–479	blue
4	545–565	green
5	1230–1250	SWIR 1
6	1628–1652	SWIR 2
7	2105–2155	SWIR 3

Figure 13 MODIS Bands. Adopted from ResearchGate

Sentinel-2 bands

Sentinel-2 Bands	Central Wavelength (µm)	Resolution (m)
Band 1 - Coastal aerosol	0.443	60
Band 2 - Blue	0.490	10
Band 3 - Green	0.560	10
Band 4 - Red	0.665	10
Band 5 - Vegetation Red Edge	0.705	20
Band 6 - Vegetation Red Edge	0.740	20
Band 7 - Vegetation Red Edge	0.783	20
Band 8 - NIR	0.842	10
Band 8A - Vegetation Red Edge	0.865	20
Band 9 - Water vapour	0.945	60
Band 10 - SWIR - Cirrus	1.375	60
Band 11 - SWIR	1.610	20
Band 12 - SWIR	2.190	20

Figure 14 Sentinel-2 Bands. Adopted from Satellite Imaging corporation

From MODIS bands we will use band 1 (Red) and band 2 (NIR), and from Sentinel-2 bands we will use band 4 (Red) and band 8 (NIR), to calculate the NDVI index which is equal to (NIR-RED) / (NIR+RED).

E. Conclusion

If international economies keep growing and people become more prosperous, then the way of thinking about forests will change. People may be more interested in conservation, and moving towards green growth and bio-economies will reduce pressures on life on land. But, if economies become stagnant from now to 2030, then the reverse may occur: humans will be more worried about their short-term health and might resist conservation measures.

So, combining SDG 15 into all of the processes guided by the other SDGs will be crucial to support life on land to 2030 and further on.

IV. Google Earth Engine

Figure 15 GEE image. Adopted from "GIS Lounge"

A. What is the google earth engine? [10]

Google Earth Engine is an online platform that allows you to explore and analyse satellite imagery, measure changes in the earth's surface, and track changes in natural resources.

Also, it is a tool that allows users to visualize and understand geographic data in a variety of ways. This data includes imagery for the entire globe from various sources such as satellites and maps at varying resolutions that contains a great deal of interpretable visual information.

The platform is built on Google's cloud computing infrastructure, which provides a scalable way to process large amounts of data. It is used by a wide range of organizations, from environmental groups to national governments.

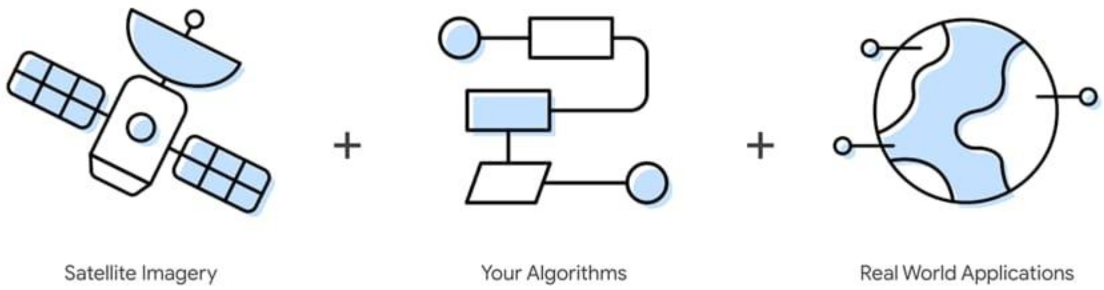

Figure 16 Introducing GEE. Adopted from "Sior Consulting STD"

Google Earth Engine has two main components:

1. The Google Earth Engine API

"API is the abbreviation of Application Programming Interface", is a cloud-based geospatial processing platform that makes large-scale cloud computing functions available to the users. This advanced platform permits the users to process numerous remotely-sensed images, also permits process of vector datasets, and this provides a way for developers to integrate Google Earth Engine functionality into their applications.

2. The Google Earth Engine Processing Service

this allows users to upload data sets and run jobs on them using the power of the cloud.

B. How to use the Google earth engine (GEE)?

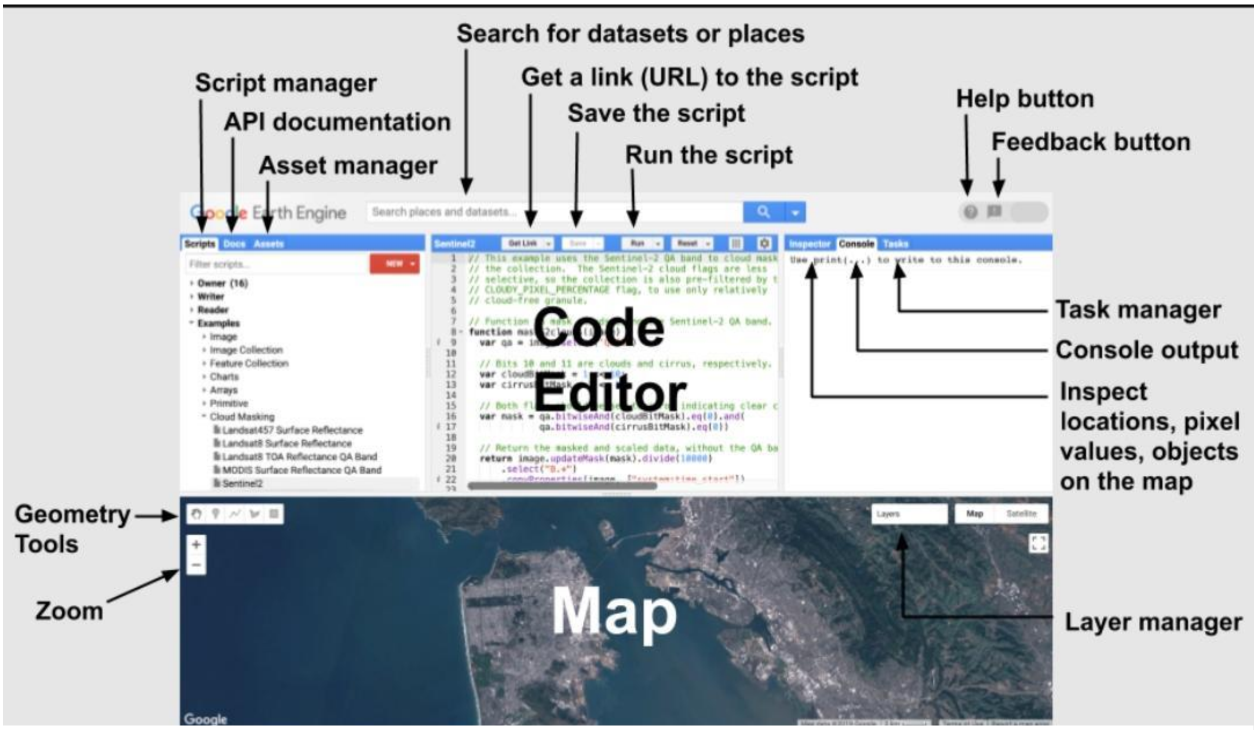

Figure 17 GEE code editor. Adopted from "Google Earth Engine"

1. Accessing GEE [11]

To use Google Earth Engine, request a user account. after approval, you have access to the Google computation infrastructure and the datasets. This access takes the form of a RESTful API. There are now 2 manners of doing so, are:

first Python-based: The Python library is the less popular of the 2 methods, and it permits users to have interaction with Earth Engine using the Python programming language, and due to the fact that the Python API doesn't support any form of visual output, as a result, it's highly recommended to use the Code Editor IDE.

second JavaScript-based: The more popular manner of having access to GEE is through the JavaScript library which is accessed via a web-based IDE, more commonly known as the Code Editor. Even though (not specifically required), it is recommended that you use Google Chrome, as this is the most popular manner of gaining access to GEE.

In general, there are 3 main types of operations the user can perform in Google Earth Engine, are Methods, Algorithms, and Functions.

First, methods and they require an object to act on and take inputs. Second, algorithms they are objects themselves and take a specific input to return an object. Finally, functions and they take an input and do something within the API.

2. Google Earth can be used: [12]

We (as students) can use it to find our homes, schools, and other locations that are familiar to us. Through Google Earth, we as users can make informed decisions about the world by comparing our current location to nearby ones. We can also learn about the world through rich layers of mappable data offered by Google's server. In addition, we can create and display our data. Also, we can use it as a basis for homework assignments, and inquiries during class presentations. To create imagery and maps for PowerPoint, Word, and other presentation tools, and as data discovery, organization, and distribution tools for research projects. GEE gives us the possibility to rapidly and accurately process huge amounts of satellite imagery, so we can identify when and where tree cover change has occurred, we can analyse forest and water coverage, assess the health of agricultural fields, or predict disease outbreaks, and many other analyses. finally, we can calculate remote sensing indices (NDVI, NDWI, NDBI…) by using the google earth engine.

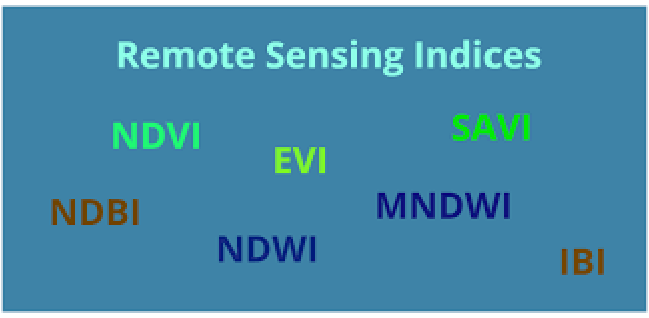

Figure 18 Remote sensing indices. Adopted from "GIS Tutorials"

3. Advantages and Disadvantages of using GEE [10]

There are many advantages of using the google earth engine, like: is easy to install and use, provides easy, web-based access to an extensive catalogue of satellite imagery and other geospatial data in an analysis-ready format, has a large active user community and is an effective tool for integrating the study of multiple disciplines. GEE is a great research tool and is appropriate for educational use in a wide range of subject areas. Also, it enables users to create and display their data, is pre-loaded with a wide variety of useful data, and can work with abundant third-party data that is available on the web. in addition, GEE has up-to-date maps and data, extremely detailed maps with 3D capabilities, despite all these features and we can download it for free.

GEE has some disadvantages, for example, low resolution in some areas, especially less developed countries are poorly presented on maps, and not all the areas of the globe are covered, and of course, it requires internet access.

4. NDVI calculation in GEE [13]

The normalized difference vegetation index (NDVI) is an indicator of calculating vegetation health, so this indicator observes how plants reflect certain ranges of the EMS (electromagnetic spectrum). thus, a healthy plant actively absorbs red light and reflects NIR (near infra-red), in contrast, an unhealthy plant will do the exact opposite.by using remote sensing, we can measure wavelengths of light absorbed and reflected by green plants, so satellite sensors are a superb source of spectral signature data for NDVI analysis. The NDVI index is very useful in numerous ways and for many purposes, for example, NDVI can help farmers do precision agriculture, also this index can allow people who care about conservation to understand and follow ecosystem changes.

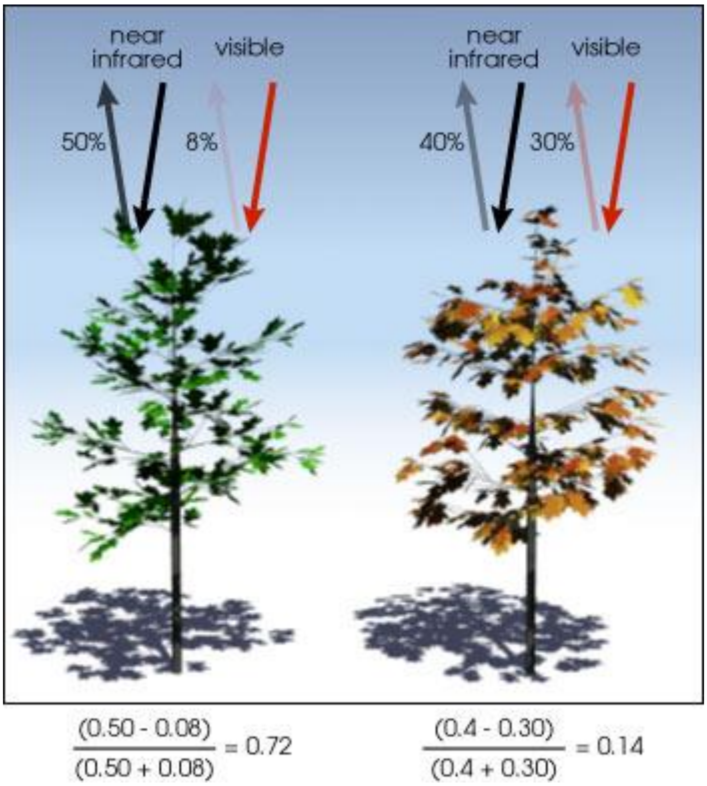

Figure 19 Example of NDVI calculation. Adopted from "Earth Observatory"

The general formula of NDVI is:

NDVI = (NIR-RED) / (NIR+RED)

The value of the NDVI will always be between -1 and +1. If the Value is between -1 and 0, that means the plants are dead, or the objects are inorganic such as stones, houses, and roads. And if the value of NDVI is between 0 and 1 that means the plants are not dead, with 1 being the healthiest and 0 being the least healthy.

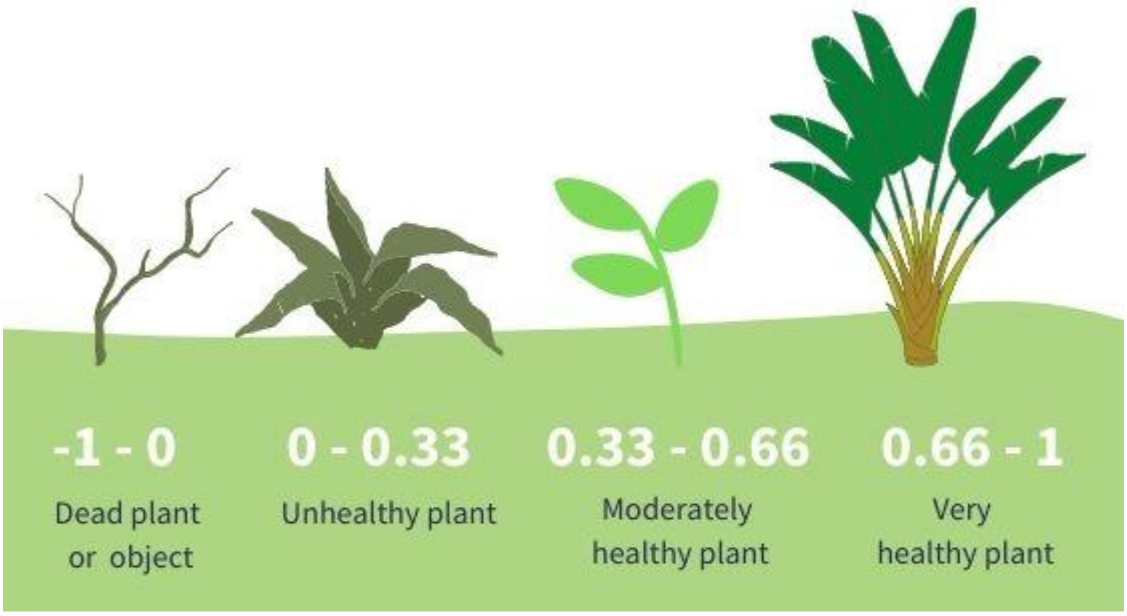

Figure 20 NDVI value meaning. Adopted from "UP42"

C. conclusion: [10]

Google Earth Engine, as a whole, represents the most advanced, cloud-based geoprocessing platform to this date. Even though many other platforms enclose some of these features, no suite of tools presently available can repeat the access to geospatial data, the simplicity of use, and the awesome power of analysis that GEE offers. also, there are new features, and algorithms constantly updated.

V. Contribution

In this chapter we will talk about how we can calculate the NDVI index in the google earth engine, also we will write the code of the running application with the JavaScript programming language.

As mentioned in the previous chapter, the NDVI is the normalized difference between the red and the infrared band, calculated as NDVI = (NIR-RED) / (NIR+RED).

So now, we will calculate the NDVI index and we will have a look at the NDVI of May 2022 for the whole country of Lebanon.

There are 3 kinds of satellites we can use to calculate NDVI, are: Sentinel-2, Modis, and Landsat 8. In this article, we will use just 2 satellites (sentinel-2 and MODIS) and we will make a quick comparison between them and talk about their differences in resolution, and which one is better to use. We choose these 2 satellites because they include data about

Lebanon in May 2022. (Landsat-8 do not include the necessary data we want for Lebanon in May 2022).

we will use a simple code to get all the names of all the countries that have data in the FAO dataset, so we can assume that Lebanon is included in this dataset, also we can see how it is written.

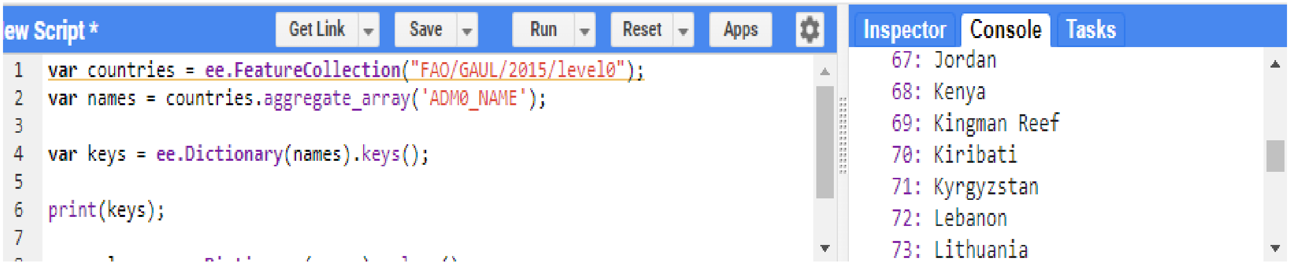

Figure 21 Lebanon in the FAO dataset. Adopted from GEE

This image is the output, so we can now see that there are many countries in the list and the number 72 is Lebanon, so we can now see the necessary schema to calculate the SDG goal 15.3.1.

A. schema

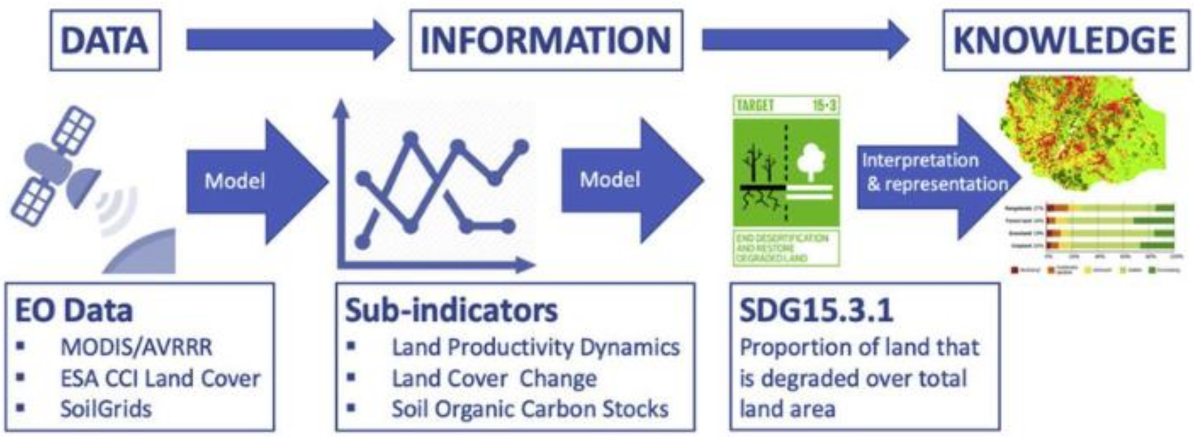

Figure 22 Schema to calculate SDG 15.3.1. Adopted from ScienceDirect

We will not be able to calculate the percentage of land degradation now, but we will calculate the sub-indicator productivity, and the most popular way to calculate it, is by calculating the NDVI index, so we will calculate this index by using two kinds of satellites/data which are MODIS and Sentinel-2, by accessing the FAO (Food and Agriculture organization) dataset. The explanation of this schema and the steps to follow will be in the sections below.

B. Calculating NDVI by using Sentinel-2 [14]

In this section, we will access the whole Sentinel-2 Level-2A collection for May 2022.

1. Ways to calculate the NDVI index [15]

There are three ways to calculate NDVI are:

- The first way uses simple band operations integrated in the GEE to calculate band maths;(Layer name: NDVI Lebanon May 2022 V1)
- the second way uses the function **". expression"** to write a mathematical formula (Layer name: NDVI Lebanon May 2022 V2)
- and the third way uses the function "**. normalizedDifference"**, which calculates the Normalized Difference for two input variables. (Layer name: NDVI Lebanon May 2022 V3)

2. Steps to follow in the NDVI calculation (Sentinel-2)

Step one: data preparation, but first of all we should determine the objective of the project, and here is to calculate the NDVI index for the country "Lebanon" at a specific time "May 2022" by using Sentinel-2 in the Google Earth Engine. Then we gather the necessary data we need to use, here in our case we gather all the data for the country "Lebanon" from the FAO dataset and this dataset provides free access to food and agriculture data for over 245 countries and Lebanon is one of them like we see in the above image (image name: Lebanon in the FAO dataset). Then we prepare the data to calculate NDVI by accessing the Sentinel-2 data and filtering it for all the images of the year 2022 and the month 'May' that lie within the geometry's boundaries. Also, we keep only the suitable bands and filters for cloud coverage. And finally, we clean the data by using the median filter so we can then reduce the noise in an image so it preserves useful detail in the image.

Step two: data processing, now we calculate the NDVI index in three ways, are simple band operations, expression function and normalized difference function. So, by these three ways we calculate the formula of the NDVI index, which is NDVI = (NIR-RED) / (NIR+RED).

Step three: indicator calculation, so by calculating the NDVI index in the previous step, (also we should have calculated the two other sub-indicators "soil carbon" and "land cover"), we can now conclude our SDG indicator 15.3.1.

Step six: is sharing our results, so we visualize all the calculations we made in the previous steps in Google Earth Engine as a map layer, and you can see the results of the above calculations by clicking the below link, and this link will lead you to my Google Earth Engine APP:

https://antonysakr.users.earthengine.app/view/calculating-ndvi-by-using-sentinel-2

C. Calculating NDVI by using MODIS [16]

1. Ways to calculate the NDVI index

Now, we are going to look at how to calculate the NDVI using Modis data of the same area" Lebanon", at the same time" May 2022".

we will look at two alternative ways to access the NDVI from MODIS bands:

- First way, that is again almost identical to that of Sentinel-2 using band maths, (Layer name: MODIS NDVI)
- and second way, that uses a MODIS-dataset that already includes NDVI bands. (Layer name: MODIS NDVI alternative)

2. Steps to follow in the NDVI calculation (MODIS)

Step one: Data preparation, so first of all we should determine the objective of this part, and here is to calculate the NDVI index for the country “Lebanon” at a specific time “May 2022” by using MODIS in the Google Earth Engine. Then we gather the necessary data, as we did in Sentinel-2, in our case we gather all the data for the country “Lebanon” from the FAO dataset. Then we prepare the data to calculate NDVI by accessing the MODIS data and filtering it for all the images of the year 2022 and the month ‘May’ that lie within the geometry’s boundaries (same country and time as we did before). Also, we keep only the suitable bands. Finally, we clean the data by using the median filter so we can then reduce the noise in an image, so it preserves useful detail in the image.

Step two: data processing, now we calculate the NDVI index in two ways, are: using band maths (identical to Sentinel-2) and using the MODIS dataset that already includes NDVI bands. So, by these two ways we calculate the formula of the NDVI index, which is NDVI = (NIR-RED) / (NIR+RED).

Step three: indicator calculation, so by calculating the NDVI index in the previous step, (also we should have calculated the two other sub-indicators “soil carbon” and “land cover”), we can now conclude our SDG indicator 15.3.1.

Step four: it is the final step, and it is about sharing our results, so we visualize all the calculations we made in the previous steps in Google Earth Engine as a map layer, and you can see the results of the above code by clicking the below link, and this link will lead you to my Google Earth Engine APP:

https://antonysakr.users.earthengine.app/view/calculating-ndvi-by-using-modis

VI. Results and Analysis [17]

In this section we will see the results of all the previous code (Sentinel-2 and MODIS).

So first of all, this is the Lebanon map without any changes and without any calculations.

Figure 23 Lebanon map

Also, we should know about the layers, how to choose the best layer and to change its settings. So, after running the code a layer button will appear, like the one in the image below, and after pressing this button each code has its own layers. Also, we can choose if we want to see the output in Map or satellite images. Here in this project, we chose the satellite option.

Figure 24 Layer and Satellite buttons

And after pressing on the layer button, if we want to change the settings of each layer, we should press the settings button.

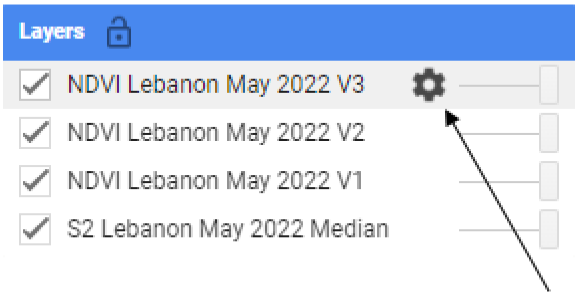

Figure 25 Layer settings button

It will lead you to the settings of the specified layer. For example, like the image below, so we can change the Range, the opacity and the colours of the output, also if we have many indices, we can choose the one that we want, but here we have one index which is the NDVI index. Finally, you should press the apply button to submit your work.

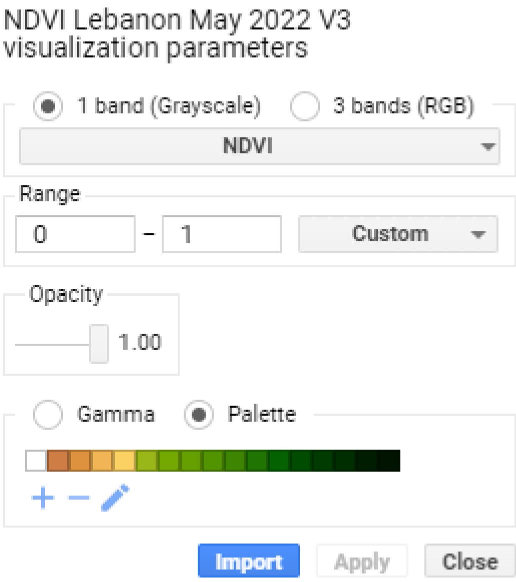

Figure 26 settings for a layer

A. Sentinel-2 results

Now we will see, all the resulting images of the Sentinel-2 satellite, we have multiple layers (four layers), the first layer is the median for Lebanon country in May 2022 and its name here is "S2 Lebanon May 2022 Median", the three remaining layers are for the NDVI index calculation, so each one calculated NDVI index in a different way and by using the different function, and named respectively "NDVI Lebanon May 2022 V1", "NDVI Lebanon May 2022 V2", "NDVI Lebanon May 2022 V3". So, as we can see in the image below, when we press the "Layers" button we will see the 4 layers.

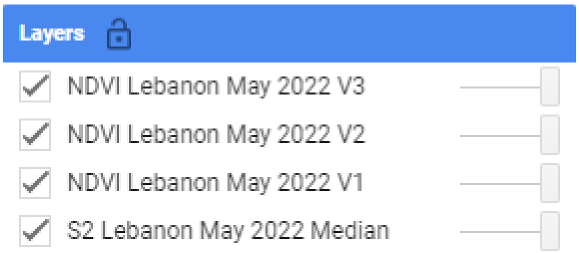

Figure 27 Sentinel-2 layers

And to see each layer alone, for example, if we want to see the “NDVI Lebanon May 2022 V1” layer, we should turn off all other layers by clicking on the box beside each of them, to understand better you can see the image below. So, you can now choose the layer that you want to see.

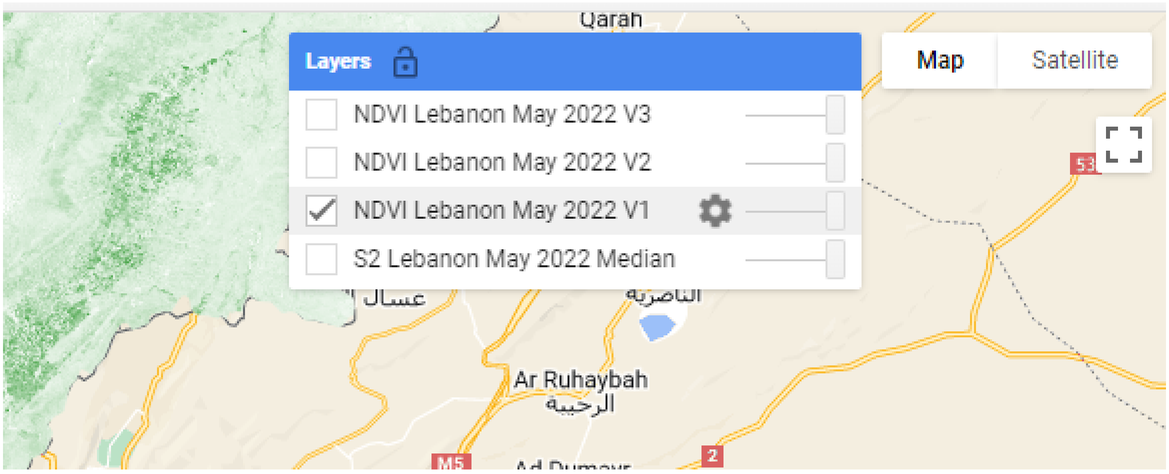

Figure 28 choosing one layer

You can see below all the resulting images of all the layers.

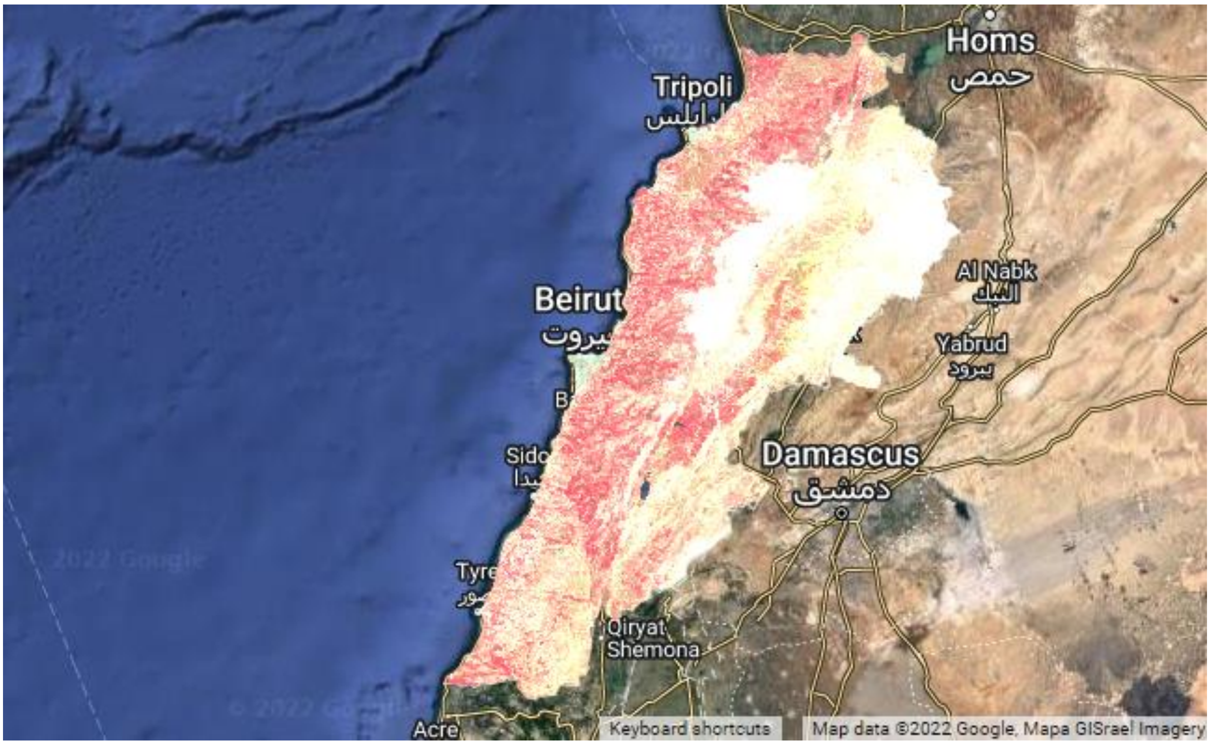

Figure 29 Median Lebanon May 2022 S1

Figure 30 NDVI Lebanon May 2022 V1

Figure 31 NDVI Lebanon May 2022 V2

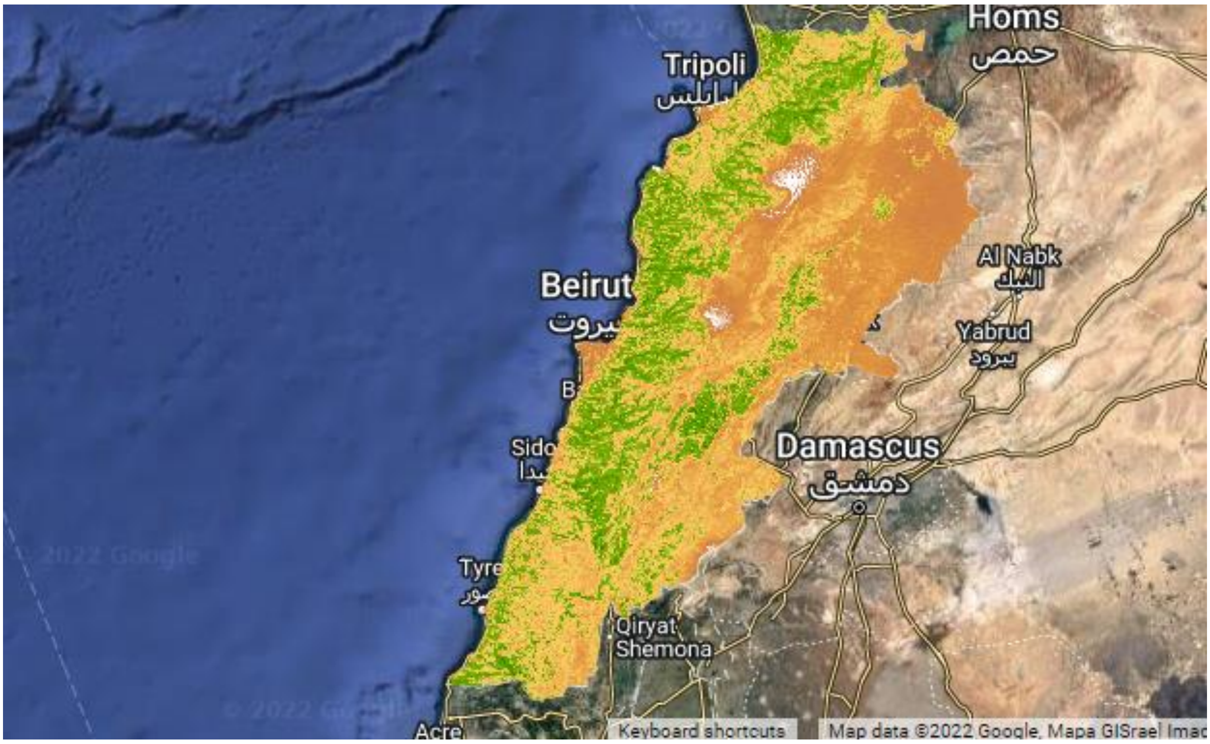

Figure 32 NDVI Lebanon May 2022 V3

B. MODIS results

Just like the results of Sentinel-2, here we have 3 layers for the MODIS satellite are Median and two ways to calculate the NDVI index, so we can see below all the resulting images for this satellite.

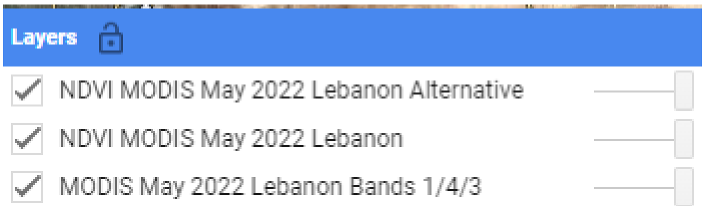

Figure 33 Modis layers

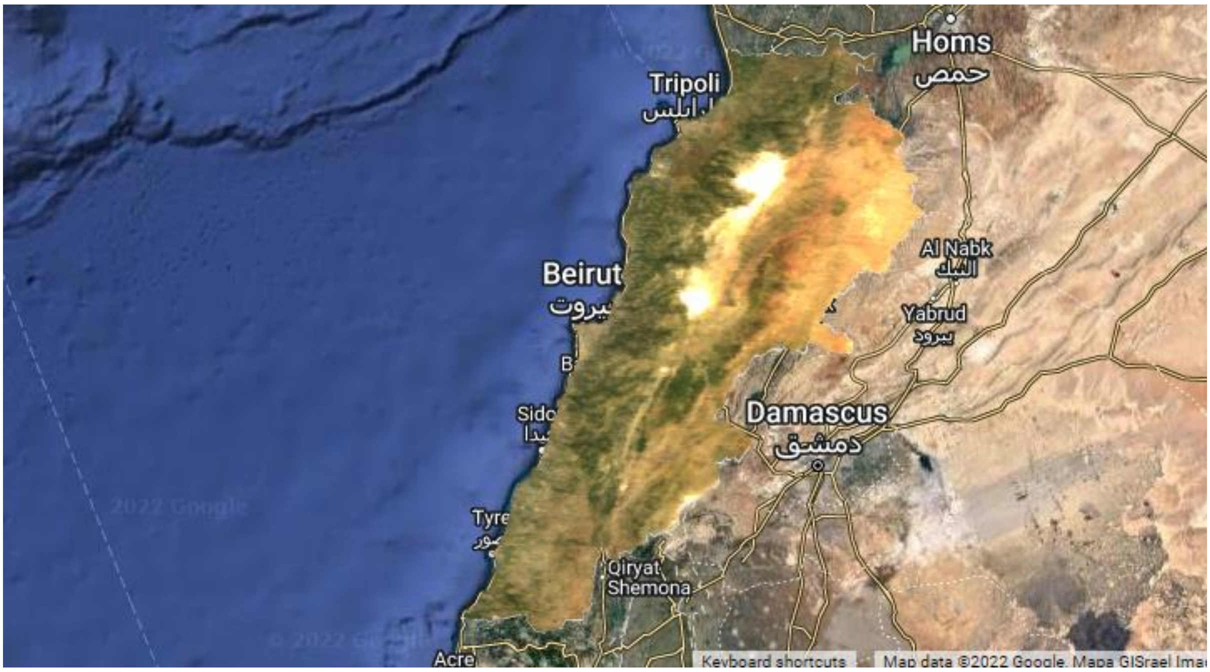

Figure 34 Median Modis

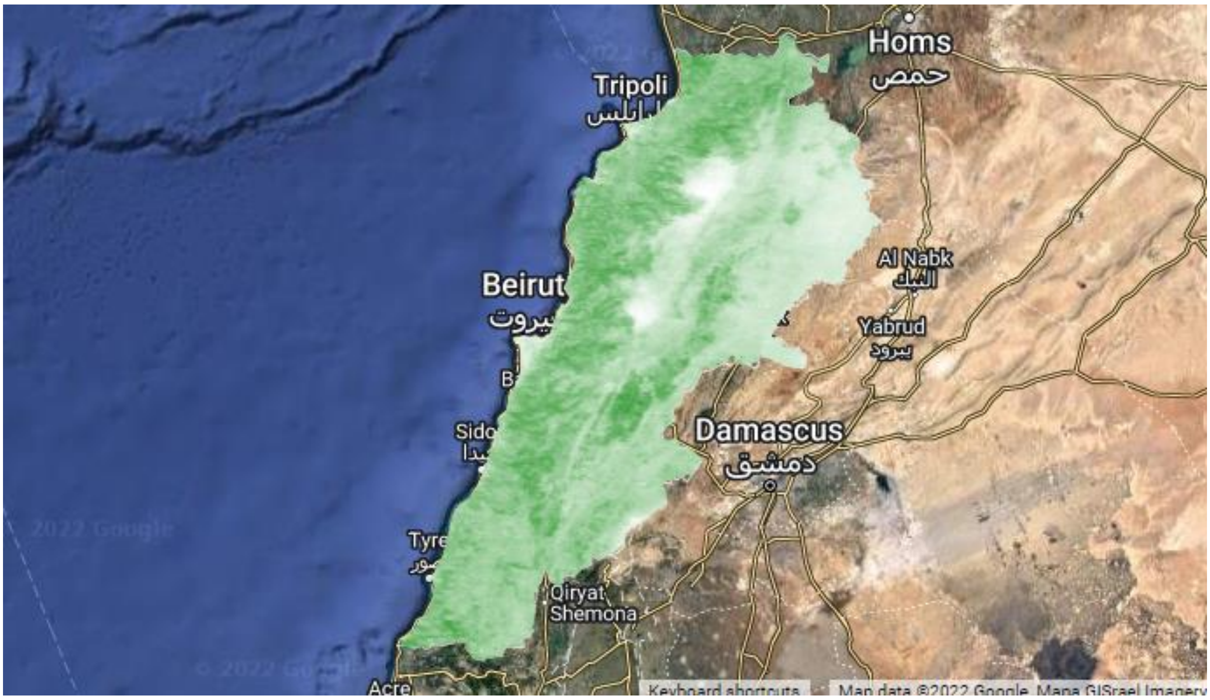

Figure 35 Modis NDVI

Figure 36 MODIS NDVI alternative

C. Comparing sentinel-2 and MODIS [18]

Neither MODIS nor Sentinel-2 is better. The right satellite depends on the application we are working on. so, every satellite is better than the other in a specific feature, for example, Sentinel-2 provides a much higher spatial resolution to get a better view of the targets, so if we are looking at objects such as smallholder agricultural fields or urban buildings, they are too small to be distinguished from MODIS imagery. A very high spatial resolution like Sentinel-2 is then a prerequisite.

But, if your project doesn't need high spatial resolution, in this case, MODIS provides very powerful data. So, MODIS is better in many features like higher temporal resolution, while Sentinel-2 takes at least five days to cover the entire globe, MODIS takes only a day or two. Also, MODIS has Higher Spectral Resolution, it measures 36 spectral bands compared to only 13 measured by Sentinel-2, and of course, Modis has Longer time series, it is a big negative point of newer high-resolution satellites is that they are too recent for our data-starved models. MODIS has been running since 1999 and Sentinel-2 was launched in 2015. Since a satellite can only provide images from its day of launch, MODIS provides a much longer history of data. This restricts the number of data models that can train on, making it harder to identify patterns. Modis also provides images more frequently, which can be essential for urgent warning systems.

And here in our case, to compare sentinel-2 and MODIS, we want to show the results of the previous code of each of them for the same place (Beirut) at the same time (May 2022) so we can then conclude which one is better to use for our topic.

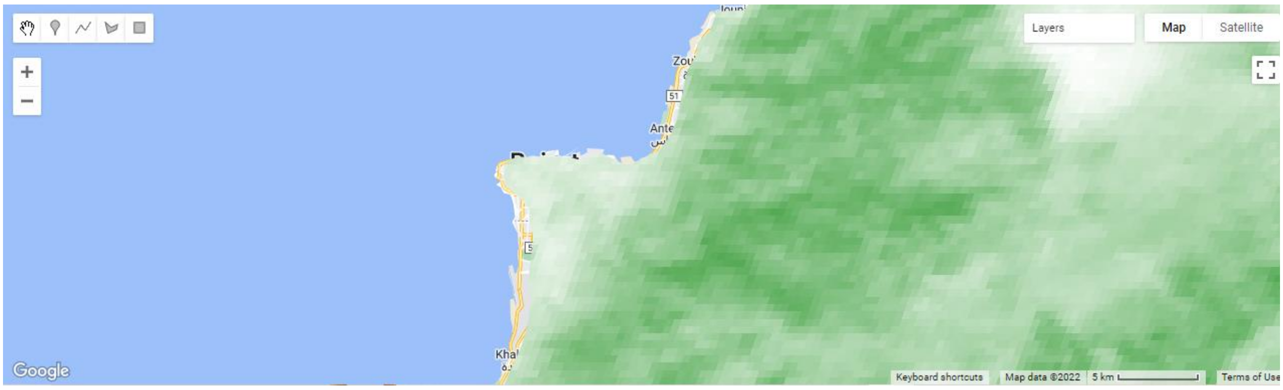

Figure 37 result of calculating NDVI by using MODIS. Adopted from GEE.

Figure 38 result of calculating NDVI by using Sentinel-2. Adopted from GEE.

In our case, using the Sentinel-2 satellite is far better than using MODIS, because the resolution is the most important factor we want to focus on, so we can see clearly that the resolution of the second picture is much better than the first one. We can conclude that the accuracy of Sentinel-2 is higher than the accuracy of MODIS satellite.

D. Analysis of the results [19]

In this part, we will do an analysis of the results, we will use one of the previous layers which is the "NDVI Lebanon May 2022 V3", it is the clearest one because we use the Sentinel-2 satellite, so it has the best resolution (you can see the comparison part to make sure), also its colours are very expressive, the green colour is for the presence of plants, if the green colour is darker it means that there is a forest, and If we have light brown it means that we have a small forest or a little vegetation, and finally if we have dark brown so we do not have any kind of vegetations.

In this project, we will divide Lebanon into five parts, for each part we will show its image and we will make an analysis of the results. These five parts are, first is "Tyre and its suburbs", the second is "Sidon and its suburbs", third is "Beirut and its suburbs", fourth is "Baalbek and its suburbs", and finally "Tripoli and its suburbs".

For the first part which is "Tyre and its suburbs", to specifically identify what is these suburbs, we will count some of the cities and towns that are included in it, so this part is constructed from Tyre, Naqoura, Al tiri, Marjaayoun, and many other small towns. You can see clearly all these cities and towns in the image below.

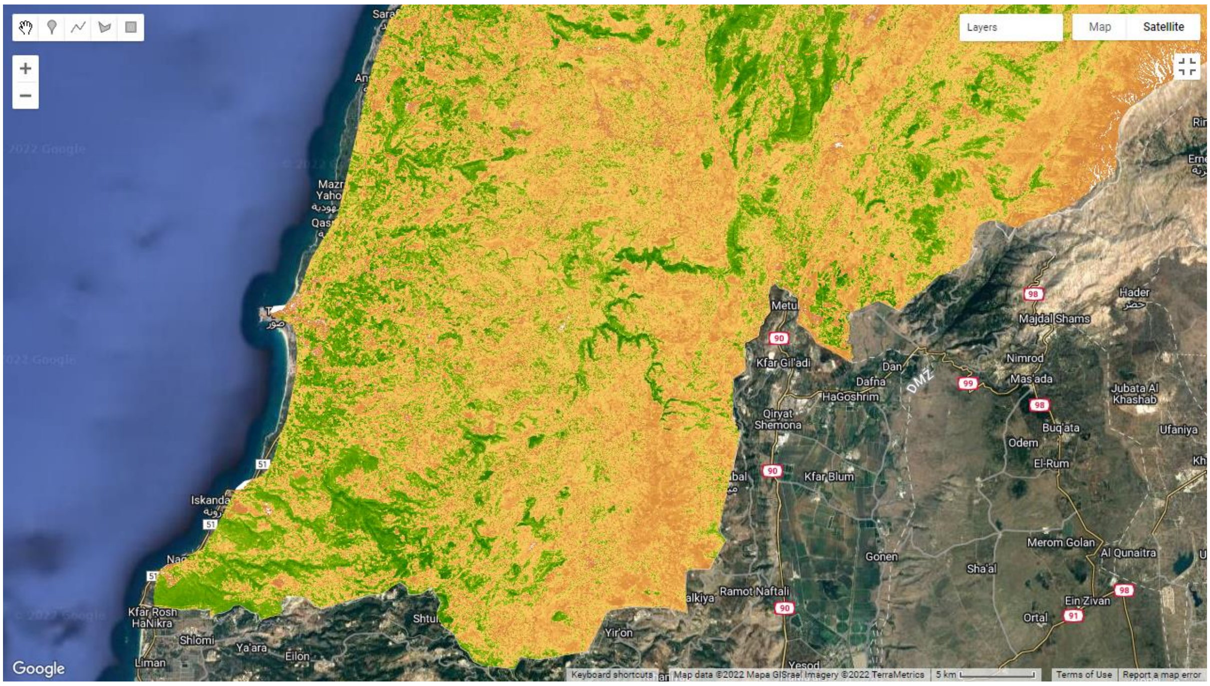

Figure 39 Tyre and its suburbs. Adopted from GEE.

As we can see the closer to the sea the higher the NDVI index is, this is because the citrus trees are planted along the shore of Tyre, so we can conclude that the sea has a positive correlation with the NDVI index, when there is more water there are more vegetations, and the altitude plays a role, the closer to the coast, the higher the proportion of plants, this is due to the moderate temperature on this part for most of the year, and there is a good percentage of rain. also, another reason why we have some points with a lot of vegetation is that the suburbs of Tyre are small towns (Like my village Debel), and in these towns, there are not a lot of people, so there are not a lot of constructions. The last reason is that a lot of people survive from agriculture, so we have more green spaces and higher NDVI.

The second part is "Sidon and its suburbs", to specifically identify what is these suburbs, we will count some of the cities and towns that are included in it, so this part is constructed from Sidon, Hasbayya, Rashaiya, jezzine, joub jannine, and many other small towns. You can see clearly all these cities and towns in the image below.

Figure 40 Sidon and its suburbs. Adopted from GEE.

Here there is more vegetation than the previous one ("Tyre and its suburbs"), although this region has very close characteristics to the previous region, the sea has a positive correlation with the NDVI index, when there are more water there are more vegetations, also it is approximately the same altitude as the previous region, so has the same temperature and same climate, also an important feature of this region that it contains soil suitable for the growth of plants, especially citrus trees and banana trees.

The third part is "Beirut and its suburbs", to specifically identify what is these suburbs, we will count some of the cities and towns that are included in it, so this part is constructed from Beirut, Baabda, Aley, Jounieh, and many other small towns. You can see clearly all these cities and towns in the image below.

Figure 41 Beirut and its suburbs. Adopted from GEE.

Beirut is the capital of Lebanon, it is normal that there is overcrowding, that will lead to a lot of constructions, small spaces of vegetation, its people never care about agriculture and if they care there is no big spaces they can plant in it, although Beirut is located on the coast of the sea, but it never makes use of the water that it has easily in agriculture. As we can see clearly, the further away from the city, the higher the percentage of vegetation, so the NDVI index is low in Beirut (almost negative in 95% of it) and it increases as we get closer to the suburbs. The temperature is another reason why the NDVI index is low, it is very high in Beirut, and it does not rain a lot in most of the year, except the four months (January, February, November and December).

The fourth part is "Baalbek and its suburbs", to specifically identify what is these suburbs, we will count some of the cities and towns that are included in it, so this part is constructed from Baalbek, Aarsal, Maalqa, Qaa, Hermel, and many other small towns. You can see clearly all these cities and towns in the image below.

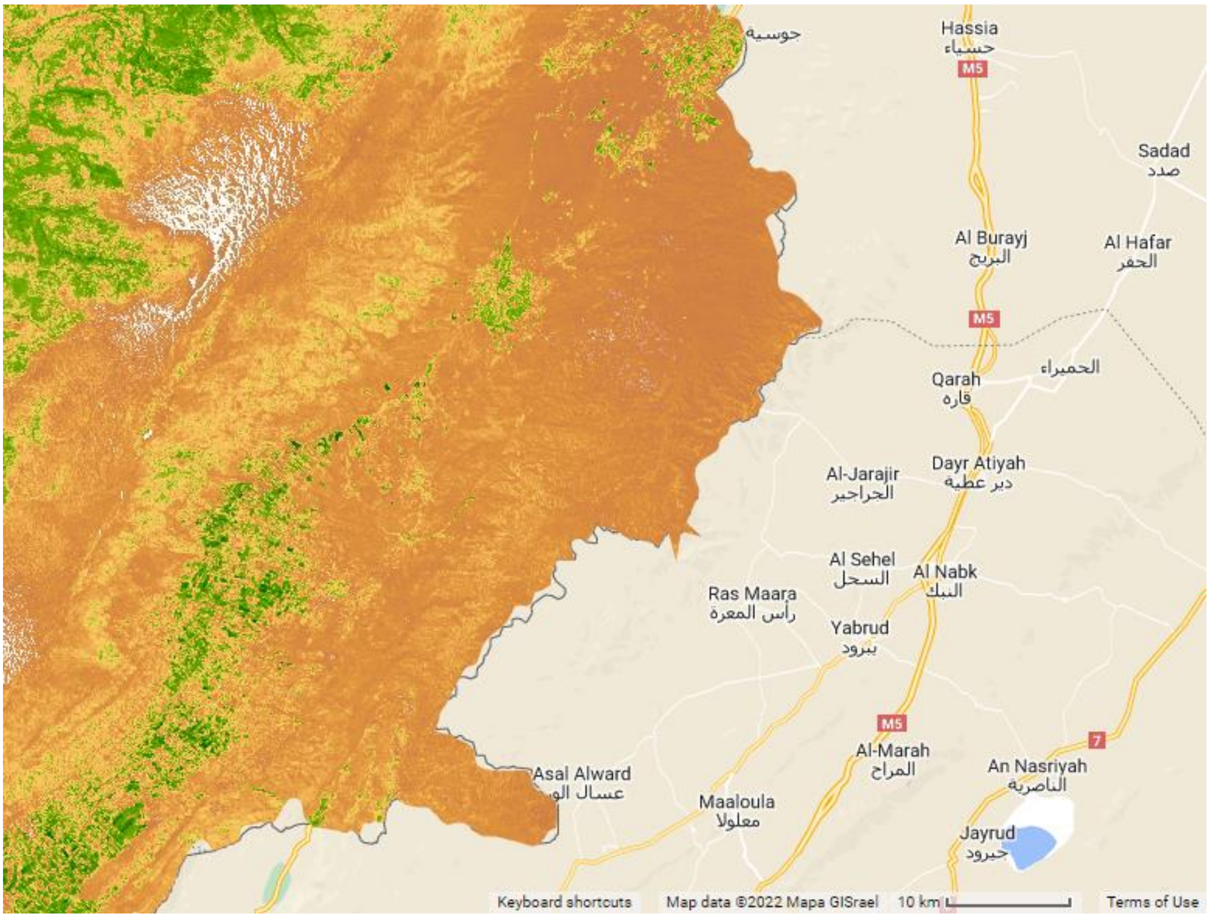

Figure 42 Baalbek and its suburbs. Adopted from GEE.

This area is divided into two parts, the first part, where there are many plants and a high value of NDVI index, which is near Baalbek (the Beqaa valley), this area is abundant with its plants because of its mild weather, and the quality of its soil. Also, the people who live in this area are very poor so they depend mainly on agriculture for living.

And the second part, which has a few plants and a small value of NDVI index, which is the east Lebanon mountains chain and Aarsal. Naturally, the eastern Lebanon mountain chain is not suitable for agriculture because it is a rugged land. And Aarsal is far away from the sea, also its soil is not the best for agriculture, so the main reason for the lack of plants in this area is the desertification, its weather is very similar to the weather of the desert, during the day the temperature is high and at night it drops sharply, and also there are some of security problems in this area.

The fifth part is "Tripoli and its suburbs", to specifically identify what is these suburbs, we will count some of the cities and towns that are included in it, so this part is constructed from Tripoli, Ain El Qatroun, Bsharri, Zgharta, Beit Haouch, khouchah, and many other small towns. You can see clearly all these cities and towns in the image below.

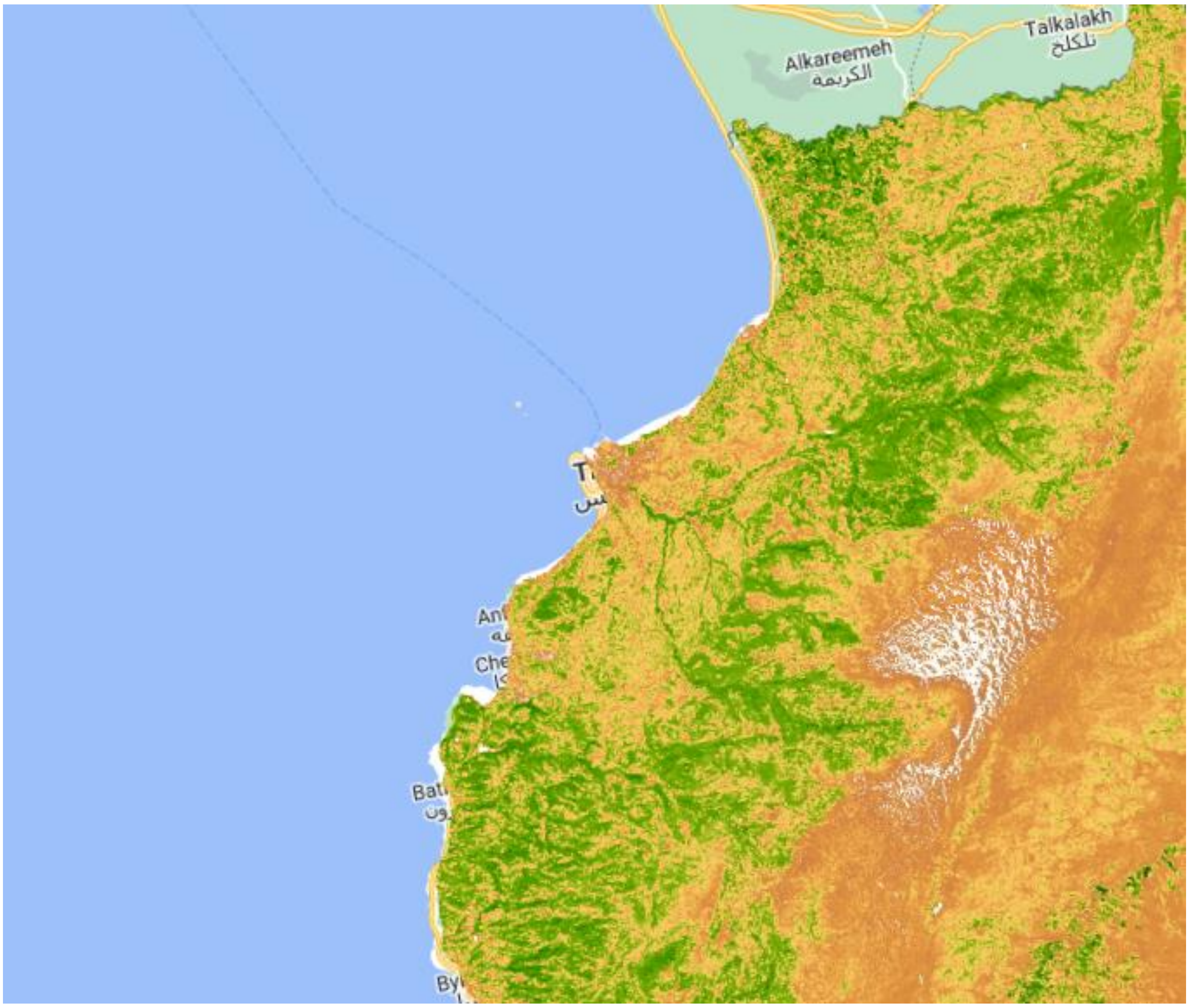

Figure 43 Tripoli and its suburbs. Adopted from GEE.

As we can see the closer to the sea the higher the NDVI index is, as we concluded before there is a positive correlation between the presence of the water and the NDVI index. This is the poorest area of Lebanon, all people here, specifically after the economic crisis, started dependent more and more on agriculture, the moderate heat and the heavy rain help them plant more.

E. Future work and conclusion

Our main objective was to calculate the SDG indicator 15.3.1 and to obtain results as a percentage of how much the land is degraded, we should calculate three sub-indicators, as we have seen before, productivity, soil carbon, and land cover. But here in this project, we calculated one of these sub-indicators, which is productivity, as we know, one of the most commonly used surrogates of NPP is the Normalized Difference Vegetation Index (NDVI), computed using information from the red and near-infrared portions of the electromagnetic spectrum, by gathering data from remote sensing satellites (MODIS and Sentinel-2), so by calculating the NDVI index we tried to calculate the SDG indicator 15.3.1, and as future work, we can calculate the two other sub-indicators to get more developed results.

Also, we tried to build a web app using python API, but for some required infrastructure problems, we don't. but we replace it with two GEE apps, one for MODIS and the other for Sentinel-2, you can see them in the contribution part.

Lebanon is a green country, because of its geographical location (located on the eastern shore of the Mediterranean Sea), its beautiful weather, the rate of rainfall in the year is high and many other reasons. Excepting some areas where the NDVI index is low for some reasons discussed above, as a whole country, Lebanon has a high value of NDVI index in most of its areas and has many vegetation spaces, every day these spaces are increasing, and a lot of tourists from all the world, especially from other Arab countries, come to see its beautiful nature. Because of this project, we can confirm that Lebanon is a green country by calculating the NDVI index for all the country areas and by seeing the NDVI maps.

VII. Bibliography

[1] "earthdata.nasa," [Online]. Available: https://earthdata.nasa.gov/learn/backgrounders/remote-sensing.

[2] T. Mostak, "heavy.ai," [Online]. Available: https://www.heavy.ai/technical-glossary/remote-sensing.

[3] S. W. Running. [Online]. Available: https://www.sciencedirect.com/topics/earth-and-planetary-sciences/remote-sensing.

[4] "THE EUROPEAN SPACE AGENCY," [Online]. Available: https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-2/overview.

[5] "Yale university," [Online]. Available: https://yceo.yale.edu/what-modis.

[6] "unstats," [Online]. Available: https://unstats.un.org/sdgs/report/2018/Goal-15/.

[7] "landportal," [Online]. Available: https://landportal.org/node/52265.

[8] D. Murphy, "soilquality," [Online]. Available: https://www.soilquality.org.au/factsheets/organic-carbon#:~:text=Organic%20matter%20(%25)%20%3D%20Total%20organic%20carbon%20(%25)%20x%201.72&text=However%20this%20can%20vary%20with,vary%20in%20size%20and%20decomposability..

[9] "SDG Indicators Metadata repository," unstats.un.org, [Online]. Available: https://unstats.un.org/sdgs/metadata/?Text=&Goal=15&Target=15.3.

[10] "MDPI," [Online]. Available: https://www.mdpi.com/2072-4292/10/10/1509/htm.

[11] J. M. a. L. X. Coll, "gistbok.ucgis," [Online]. Available: https://gistbok.ucgis.org/bok-topics/google-earth-engine.

[12] "serc carleton," [Online]. Available: https://serc.carleton.edu/introgeo/google_earth/how.html.

[13] "informatics and GIS," [Online]. Available: https://igis.ucanr.edu/Tech_Notes/EarthEngine_NDVI/#:~:text=NDVI%20is%20calculated%20as%20the,in%20the%20Sentinel%2D2%20imagery..

[14] "gis stack change," [Online]. Available: https://gis.stackexchange.com/questions/395713/gee-extracting-country-names-from-fao-dataset.

[15] A. Taylor, "Informatics and GIS Program," [Online]. Available: https://igis.ucanr.edu/Tech_Notes/EarthEngine_NDVI/.

[16] "freie universitat Berlin," [Online]. Available: https://www.geo.fu-berlin.de/en/v/geo-it/gee/2-monitoring-ndvi-nbr/2-2-calculating-indices/ndvi-s2/index.html.

[17] "GEE," [Online]. Available: https://earthengine.google.com/.

[18] "cervest earth," cervest, [Online]. Available: https://cervest.earth/news/remote-sensing-of-planet-earth-part-3-modis-vs-sentinel-2.

[19] "Google Earth Engine," GEE, [Online]. Available: https://earthengine.google.com/.