Wetland landscape based on Sentinel-2 images and geo-tagged photographs in Centla, Tabasco

Paisaje de humedales basado en Sentinel-2 y en fotografías geoetiquetadas en Centla, Tabasco

Alejandra A. López-Caloca*, Amilcar Morales Gamas, María Gabriela López Aguilar

Centro de Investigación en Ciencias de Información Geoespacial (CentroGeo), Tlalpan-CDMX, México.

Abstract
This is an analysis of the geographic landscape in the Centla wetlands of Tabasco state, Mexico. A map shows the use of remote sensing data combined with easily understood and conveyed visual descriptive data which show the ecological conditions of the landscape. The central map of this article presents a land use and land cover study, obtained from Sentinel-2 MSI data for the Centla wetland zones. The support vector machine algorithm is used to classify Sentinel-2 images. The results show a high general precision of 90.0%, as well as high precision in separating types of wetlands. Information obtained during fieldwork at the ground level is inserted in the map, comprised of photos taken with a mobile application along the Grijalva and the San Pedro y San Pablo rivers. These photos provide visual verification of the map.

Keywords: Sentinel-2; wetland zones in Tabasco; Mexico; remote sensing

1. Introduction
Land use and land cover maps (LULC) are a valuable resource that provides important details on river landscapes and the ecosystems related to them. By comparing this type of maps obtained during several years, it is possible to analyze tendencies of land cover patterns in rivers, since these ecosystems degrade as a result of agricultural, industrial and urban activities (Yang & Liu, 2007). In Mexico, there are several national reference maps with valuable information about LULC (Gebhardt et al., 2014; Gebhardt et al., 2015; INEGI, 2017; and Defourny et al., 2018), which are in line with the Sustainable Development Goals (SDG), specifically with SDG 6.6, which is about the protection and restoration of water-related ecosystems, including wetlands, rivers, aquifers and lakes (SDG, 2020). Their categorization is adjusted to the national level, and is commonly used to identify changes in different types of water-related ecosystems. The derived maps obtained from these, have allowed the
determination of the rate of change in ecosystems over time. However, these maps must be used considering the time window they are representing (sometimes annual, bi-annual, tri-annual, etc.), since cloudy pixels are replaced by mosaics of cloudless pixels from images with different acquisition dates. Issues related to tropical and humid zones are mostly related to cloudy conditions, which tend to be constant during the rainy season. These conditions are also present in our study zone (wetland landscape of Centla, Tabasco, Mexico), and represent a real problem while performing remote sensing studies during the months of heavy precipitation.

Optical sensors such as Sentinel-2 (S2) and Landsat imagery, are often used in LULC studies to determine the extent of water bodies in wetlands (Pena-Regueiro et al., 2020), to delineate the wetland extension and area (Bhatnagar et al., 2020, Perea-Ardila et al., 2019), and to estimate the amount of carbon storage in wetlands (Were et al., 2019, Zamora et al., 2020). Sjögersten et al. (2021) carried out a study to estimate the above and below ground-C stocks in swamp forests, marshes and grazed wetlands in three different regions in Mexico. They reported that freshwater swamp forests and marshlands are important carbon deposits and wetlands that have turned into grasslands, contain a smaller amount of stored carbon.

The use of S2 Multi-Spectral Instrument (MSI) data in wetland monitoring, has been recently investigated, particularly for wetland delineation, to determine wetland vegetation types (Bhatnagar et al., 2020) and in surface water dynamics (Slagter et al., 2020). The mapping and description of wetlands, using S2 MSI data, incorporates a finer spatial resolution of 10 and 20m, with a greater number of multispectral bands, compared to those provided by the Landsat 8 OLI/TIRS satellite. S2 includes three Vegetation Red Edge bands (VRE) with a wavelength of: 0.694-0.714 (μm), 0.731-0.796 (μm), and 0.768-0.796 (μm). With these bands, vegetation indexes have been used to separate the wetlands from other land covers and to identify 2 types of wetlands, bogs and swamps (Kaplan and Avdan 2019). Furthermore, as shown by Valderrama-Landeros et al., 2018, using images other than S2 MSI, exhibiting lower spatial resolution, makes it difficult to differentiate between mangrove species, due to similarities in their spectral reflectance properties (e.g. case of stressed semi-arid mangrove forests located in the Mexican Pacific). Finally, S2 MSI data have demonstrated a worthwhile potential for monitoring water quality. For example, eutrophication of water bodies and wetland types can be determined, as well as the presence of zones with certain plant communities and phenology (Sun et al., 2018).

All these studies, illustrate the potential use of data from S2 MSI, to develop a LULC map in our study area, since there is a large number of continental water bodies that could be analyzed, including very active rivers where many species of wetlands can be found, as well as diverse mangrove strips at the edges of riverbeds.

Most prior studies in the study area, have been focused on delineating and describing the mangrove forest since it is possible to differentiate the mangrove forest pattern digitally based on their size, textures, and forms. However, the difficulty lies in separating surfaces with mixed herbaceous classes, since the precise distinction of transition areas from the wetland edges, cannot always be carried out due to a mixture of information within the pixels. S2 MSI data may be used to separate those pixels containing land covers or heterogenous land features, such as those of the study area. For these reasons, the image selection process considered S2 MSI data, to develop a LULC map with six classes in the Centla, Tabasco region, using a Support Vector Machine (SVM) algorithm, which allows the identification and delineation of open water bodies, as well as urban and built-up areas, cropland/grassland and wetlands. Two S2 MSI images were used on this study: S2A MSI for classification (acquired on July 7, 2019) and S2B MSI for validation (acquired July 12, 2019).

Moreover, motivated by the following studies, we propose to use geotagged photos to assist the classification of the terrestrial cover. ElQadi et al. (2020), for example, propose the use of geotagged photos obtained from social networks, even if social networks photos are not always adequate for research, and frequently filtering methods are needed to eliminate irrelevant photos. Also, Stohlgren et al. (2000) suggest the use of Video Mapping System (VWS) 200, combining videography, global positioning systems and geographic information systems (GIS) technologies to evaluate arid landscapes. Other related works are those of the LULC program (Laso Bayas et al., 2020), which have used photos taken with systems such as the FotoQuest mobile application.

Hence, in this work we present an interactive map including the results of the LULC study, obtained from S2 MSI satellite data and supported by geotagged photos describing the natural conditions of rivers and the types of surrounding vegetation in the landscape (e.g. tropical forest, agriculture), along the Grijalva and San Pedro y San Pablo rivers. The photos were obtained with a free mobile application and linked to the routes of the observation field trip.

The final map includes a series of photos or videos taken on the field at the ground level; these photos have geolocation providing information of the zone depicted in greater detail (longitude, latitude, date, hour). Consequently, linking the photos to the LULC map, support the visual verification of remote sensing results, and therefore improve the classification obtained.

The article is organized as follows: Section 2, describes the methodology used for processing the S2 MSI images, and the classification method; followed by the presentation and description of the visualization design, which includes the field trip routes. Section 3, shows the results of the classification for the land use layer, and the resulting map, describing the main results. Finally, in Section 4, we present the main conclusions.
2. Methods

2.1. Study zone

The study zone is located at the Grijalva-Usumacinta hydrological region in Tabasco, Mexico; between (18°40'48.0”N, -92°54’10.8”W) and (18°05’27.6”S, -92°18’54.0”E). The study area was selected by including the protected areas and the areas that are not within any protection scheme but which are transition zones. This includes important parts of the La Victoria and La Frontera-Barra de San Pedro localities, in the Centla municipality in Tabasco, where the zone is rich in biological and ecological diversity (Moreno-Cáliz et al., 2009 and Reyes-Gómez, and Vázquez-Lule, 2009). The study of this type of ecosystem requires constant field visits that imply financial resources. However, it is possible to develop a strategy based principally on satellite data and use complementary visual information collected through field trip routes to perform verification. Natural biodiversity and ecological characteristics are shown in Figure 1.

In the study area, meteorological stations are scarce; however current information was available at the Pantanos de Centla station (18°24’23”, -92°38’47”). Figure 2 shows precipitation data for the years 2018 and 2019. According to the station’s data, monthly accumulated precipitation was 83.2 mm, 0.02 mm and 17.2 mm, for July, August and September, 2019, respectively.

The values show that these were dry months. These values are low, compared to the average value of 139.8 mm/month, the historical record recorded until 2016 in the study area (Medrano-Pérez et al., 2021).

2.2. Field trip route description

The field trip routes included several navigable kilometers in the Grijalva River flow, as well as in the San Pedro y San Pablo river. Along the route it was possible to navigate between zones that were near the natural protected area of the Centla Swamps Biosphere reserve. The first verification route involved a visit to part of the Grijalva delta, mainly navigating the flow of the Grijalva river. It began at the mouth of the Grijalva, up to the Gulf of Mexico, passing by Buey Island. The field trip route gradually allowed for the observation of the natural conditions of the rivers and their surrounding vegetation. It was also possible to navigate the Tabasquillo stream where the following localities were visited: Tres Brazos, Viento lagoon and San Pedrito lagoon (see Figure 3). These zones are connected by water to the Grijalva and Usumacinta rivers. Through navigation along the flow of the Grijalva river, it was possible to reach Boca de Chilapa. The field trip date was July 10, 2019.

The Grijalva and Usumacinta rivers join near Frontera, Tabasco, in the zone known as Tres Brazos, where they share the estuary, which is also a high interest zone because of its biodiversity (INE, 2000). The second field trip route started at the suburbs of the city of Frontera, navigating along the San Pedro y San Pablo river, which is the border between the states of Tabasco and Campeche. The verification of landscape plant covers was carried out during the visit to the main wetland zones. This route went from Frontera city to the Cometa lagoon and back, ending near the city of Frontera, on July 11, 2019. Figure 3 shows the location of the study zone and the field trip routes that were followed.

2.3. Data

2.3.1. Satellite Data

LULC studies provide valuable information on the status of natural resources. When providing information on distribution, class type and wetland extension, a quantitative measure of wetland stability is given (Chasmer et al., 2020). In this work the use of S2 MSI images is proposed to update maps of the Tabasco wetland zone. The S2A MSI image that was used for classification was taken on July 7, 2019. The S2B MSI image that was used for validation has an acquisition date of July 12, 2019. Table 1 shows multispectral characteristics of S2.

![Figure 1: Examples of photos taken in the study area, showing the gamut of natural biodiversity and human activities](image1.png)
2.3.2. Reference data

This work used cartography at the national level for LULC. The ESA-CCI Land Cover S2 Prototype of Mesoamerica (ESA-CCI, 2018) was used. This map has a spatial resolution of 10 m and contains 10 cover classes, obtained from S2 MSI satellite data. Also, the CONABIO Monitoring Activity Data for the Mexican (MAD-MEX, 2018) REDD+ program were used. This database contains maps pertaining to a 3-year period, 2015 (RapidEye), 2017 (Sentinel 2) and 2018 (Sentinel 2), including the standardized land cover and land use categories at a 1:20,000 scale. The land cover map uses a classification scheme of 31 categories. Due to the great natural diversity in the study area, the 10 proposed classes in the 2018 ESA-CCI map are limited in terms of the number of types of vegetal cover, when compared to the 2018 MAD-MEX map, so we employed the latter. Table 2 shows the categories in both maps, compared to the ones used in this work.

2.4. Design strategy for the LULC map

The construction of the LULC map was carried out in three consecutive stages: 1. Consulting all essential background information, 2. Field activities and image processing analysis-interpretation and, 3. The creation of the map. During the first stage, available digital reference LULC maps were consulted to recognize the classes present in the study area. Based on these maps, targets of interest were established in the study area for the swift-flowing rivers with surrounding vegetation. During the second stage, fieldwork activities were carried out on July 10 and 11, 2019, in order to collect data on the field trip site. Besides taking digital photos with the Tablet for Fieldwork, a digital field notebook was created using WhatsApp for mobile phone. The photos and videos were identified and labeled, based on types of covers (for example, mangrove, other wetlands, grasslands, agricultural areas, tropical forest trees, urban zone) and visual observations of identified LULC transitions were recorded. The relevant information consulted during the fieldwork included: the S2A MSI image dated July 7, 2019 and the reference LULC maps. All the collected data allowed us

---

**Table 1: Sentinel-2 MSI satellite bands**

| Band  | 1 C/A* | 2 Blue | 3 Green | 4 Red | 5 VRE* | 6 VRE* | 7 VRE* | 8 NIR* | 8a NIR* | 9 WV* | 10 Cirrus | 11 SWIR* | 12 SWIR* |
|-------|--------|--------|---------|-------|--------|--------|--------|--------|---------|------|----------|---------|---------|
| Wavelength | 0.421-0.439 | 0.537-0.582 | 0.646-0.685 | 0.694-0.714 | 0.731-0.749 | 0.768-0.796 | 0.848-0.908 | 0.931-1.338 | 1.539-2.072 | 1.414-1.681 | 2.312 |
| (max-| Res.(m) | 60 | 10 10 | 10 | 10 | 20 | 20 | 20 | 20 | 60 | 60 | 20 | 20 |
| (μm) | 0.457-0.535 | 0.582-0.685 | 0.714-0.749 | 0.796-0.908 | 0.881-0.958 | 1.414-1.681 | 2.312 |

*C/A-Coastal/Aerosol; NIR-Near Infrared; SWIR-Shortwave infrared; VRE-vegetation Red Edge; WV-water vapor.
to understand the condition of the wetland landscapes and water bodies. During the last stage, a LULC map was generated with S2A MSI data (July 7, 2019). For data collected with the Tablet for Fieldwork, once the photographs were geo-tagged, they were organized individually by cover type and used in the subsequent digital classification processes.

2.5. Satellite data processing

2.5.1. Pre-processing stage

Radiometric Correction

Atmospheric correction, using S2 MSI data, has been done with the free Sen2Cor atmospheric correction software in SNAP (Science Toolbox Exploitation Platform), from the European Space Agency (Louis et al., 2016). The Sen2Cor functions are: calculation of surface reflectance in flat and rugged terrain and haze removal in images (Main-Knorn et al., 2017). The atmospheric correction algorithm considers the cirrus band to correct cloud presence, as well as the blue, red and SWIR bands to apply the atmospheric radiative transfer model. In the Sen2Cor, most of the processing parameters values were set to the default value. In Aerosol and the Mid Lat option, the rural option and the summer option were selected, respectively; with the SRTM-1sec digital terrain model. The resulting correction product has all the multispectral bands, which are re-sampled at 10m pixel size. Finally, an analysis was carried out using each band’s histograms, to confirm its initial values started with near zero values; proving that an adequate atmospheric correction was obtained.

Classification Algorithm

In the literature review, two classification methods are found: Random Forest (RF) and Support Vector Machine (SVM), which are used in LULC applications at the global, local and regional levels. Both algorithms can appropriately handle the challenging cases of the high-dimensionality of the input data, the limited number of training samples and data heterogeneity, such as multispectral images. This type of algorithms are considered to be, pixel-wise, spectral classifiers and do not consider the spatial dependence on neighboring pixels (Sheykhmousa et al., 2020). In the case of low spatial resolution images, the RF method consistently offers better results, compared to the SVM method (Liu et al., 2021); while the SVM allows for better results when using medium to high spatial resolution images (Adam et al., 2014, Noi and Kappas, 2018, Sjögersten et al., 2021).
Table 2: Description and equivalent names used for the LULC map proposed in this study which are related to other reference sources.

| Land Cover Class CCI (2017-2018) | Land Cover MAD-MEX 2018 | Categories proposed in this study | Class Description proposed in this study |
|----------------------------------|-------------------------|----------------------------------|-------------------------------------------|
| Open water                       | Water body              | Open water bodies (OWB)          | This class is formed by open water surfaces (inland water bodies), that is, without a plant cover, such as rivers, canals, lakes and lagoons. Human settlements |
| Built-up areas                   | Urban and built-up areas| Urban and built-up areas         | This class groups agricultural-herbaceous areas, those with shrubs, cultivated pastures, induced pasture and agriculture-vegetation. Other classes were bare land and beach sand. |
| Cropland/Grassland               | Cropland and Grassland  | Cropland/Grassland/Others        | The estuary mangroves are forests along riverbeds, canals and streams of the San Pedro y San Pablo river and along the Grijalva river. Coastal mangrove, riverine mangroves, Peten mangroves and inland mangroves are included in this category. Mangroves may be observed which have different heights and types of canopies that are characteristic of the mangrove species (black mangrove, white and red mangroves, and buttonwood mangrove) (Guerra and Ochoa, 2006, Moreno-Cáliz et al., 2009, Kaufman et al., 2016, Torres et al., 2018). In this class one can also find other classes considered to be evergreen tropical forest (for example, natural palms). |
| Aquatic vegetation or regularly flooded | Mangrove and Peten mangrove | Mangrove forest/Forest cover | The category includes complex wetland communities of freshwater marshes (popal-tular), the emergent rooted wetlands, rooted floating wetlands, subaquatic vegetation (Guerra and Ochoa, 2006), and flooded pasture. This is due to the fact that the vegetation and species that make up this category are related to environmental factors such as water availability, water level, flooding time and type of soil. For the study date, two wetland communities were considered and extracted. -OWV I Surfaces with mixed herbaceous and shrubby wetlands, with a damp vegetation spectral behavior. (see Fig.4) -OWV II Surface with mixed herbaceous wetlands, tular-popal wetlands and flooded pasture, with a spectral behavior corresponding to less humid vegetation than OWV I. |
| Tree cover areas                 | Popal- Tular            | Other Wetland Vegetation (OWV)   | This category includes complex wetland communities of freshwater marshes (popal-tular), the emergent rooted wetlands, rooted floating wetlands, subaquatic vegetation (Guerra and Ochoa, 2006), and flooded pasture. This is due to the fact that the vegetation and species that make up this category are related to environmental factors such as water availability, water level, flooding time and type of soil. For the study date, two wetland communities were considered and extracted. -OWV I Surfaces with mixed herbaceous and shrubby wetlands, with a damp vegetation spectral behavior. (see Fig.4) -OWV II Surface with mixed herbaceous wetlands, tular-popal wetlands and flooded pasture, with a spectral behavior corresponding to less humid vegetation than OWV I. |

SVM has been widely used in LULC maps in wetland zones (Shi and Yang, 2015; Noi and Kappas, 2018) since it helps generate a robust model with a small-sized sample dataset, improving performance, classifying types of homogenous and heterogenous land cover types and having separable land cover patterns, specially in the case of wetland covers (Singh et al., 2014).

Classification task consist of three phases: the training data acquisition, the learning algorithm parameter definition and performance of class recognition. In the first one, a set of training data is selected, the values of the attributes and characteristics of the feature space are extracted and the classifier is trained. A balanced sample with all categories and polygons uniformly distributed throughout the whole study area was designed in order to train the SVM. The training data were polygonal data with homogenous areas in the spectral signature. The edition of training areas was done manually on an S2A MSI image dated July, 7, 2019, within the study area. The training sites were digitalized for each class, using visual interpretation. Information from MAD-MEX was selected since its categorization is more...
representative for wetlands in our study area. Also, the points of photographs taken on the field integrated a photographic log-book which allowed us to have greater familiarity with the area in our investigation.

The number of training pixels for each individual LULC class is listed in Table 3. These numbers are also estimated in percentages for individual classes over the whole study area.

The training area separability evaluation was carried out with the Jeffries-Matusita distance method (Richards and Jia 1999; Padma and Sanjeevi 2014). An adequate value was considered to be between 1.90-2.00. With the quality analysis of the two-dimensional scatter plot of NIR band and red band, it was possible to evaluate how well representative training pixels were distributed in the scatterplot.

The SVM algorithms are based on data distribution to assign a pixel to a class. The result of the training data is a set of parameters called the weight vector which defines the classifier and the optimal hyperplane that represents the border between classes or support vectors. The definition of an optimal hyperplane is the separation between maximum classes and depends on the multidimensional feature space to be processed. That is, for hyperspectral and multispectral images (Bruzzone and Persello, 2019; Mountrakis et al., 2011), the classification process in the SVM algorithm depends on the number of bands that were included. The hyperplane may be designed with different types of kernel. The kernel that was used in this article was linearly separable.

In the pattern recognition phase, the trained classifier model assigns one of the classes to the new entry data, according to the similarity in their characteristics.

In the SVM classification process, we used the B2 to B8 bands. The SVM algorithm that was used is the one implemented in the ENVI 5.5 software. The description of the generic classes that were used is shown in Table 2, as well as their class label equivalents in works by other authors.

2.5.2. Post classification

The result of the classification process had editions which corrected problems concerning: the misclassified zones (due to class borders with edges that did not separate correctly in some of the interior water bodies), the assignment of cover classes, the unclassified cloud zones and the elimination of isolated pixels with a Majority filter. The classified image included a part of ocean water as water body; to refine the map, this specific zone was eliminated from the image through a specific mask of ocean water. It was also observed that beach sand was classified as an urban area, for which it was selected and re-assigned as another class. The Edit Classification Image was carried out on ENVI 5.5. Finally the classified image was converted from raster to vector shapefile layer, using ArcMap version 9.3, and then included in the interactive map.

2.6. Accuracy assessment

The accuracy assessment methodology was designed to identify classification errors in the land cover and land use map. The classified image was evaluated based on a statistical sampling design. The reference data was the July 12, 2019, S2B MSI image. The design considers the following stages: i) The sampling design method considers the calculation of the reference sample size (n), for stratified random sampling (Cochran, 1977). ii) Classified images. We used a stratified random sampling where once the reference sample size was obtained, a sample allocation to strata was determined for the image classification. The LULC classification was evaluated with approximately 10 collected sample points for small areas, and some classes based on the number of pixels in each class in relation to the total area of classes. The sampling points were obtained using PCI Geomatica software. By means of the “Random Sample List”, each class was assigned a georeference point. This list was exported in Excel format. The reference points per class were Mangrove 54; Grassland-Agri 63; Urban 8; Water 19; 91 and 68 points for OWV-I and OWV-II, respectively. iii) Reference Image. In these terms, image selection was carried out from S2B MSI image (July 12, 2019). The criteria were, the minimal cover without clouds (<1%), and most recent data-date. The proximity of image classification and reference image was 12 days. iv) We obtained the category for the reference image of each number of random points. Using the “Random Sample List” obtained in Excel, each class was assigned with a reference class by visual interpretation. For this, the band combinations were used in the following RGB (8,4,3) y (4,3,2). This was done using ArcMap software. The result of this process was an Excel table with three columns: the georeference points, the class that was obtained from the classified image and the reference class. v) We obtained the error matrix and incorporated area proportions for the new error matrix with an unbiased area estimator. To calculate the map accuracy - errors by omission and commission - the method described by Olofsson et al. (2014) was used. The ArcMap software was used to process Excel data and to prepare the “Error_Matrix” function; the result was an Excel file with an error matrix.

2.7. Photographs and field trip route technology

Using a mobile application as a tool, the collected landscape photographs were obtained along with its geolocation in order to improve cover identification along the course of rivers, canals and during the field trip route along lakes and lagoons.

A Lenovo tablet, Phab2 model, was used with a quad-core processor (1.3 GHz, 3 GB in RAM, storage of 32 GB and Android 6.0 operating system), to which the GPS Tracks Offline application was installed. This application is available at Google Play and is free for Android devices.

This mobile application allowed to generate the geospatial information layers representing the field trip routes that took place on Wednesday 10 and Thursday 11, in July, 2019.

In order to georeference the photographs, the “geo-location” characteristic was added in the photo application of the mobile device. Table 4 shows the number of photos taken for each field trip route and indicate the number of photos available in the Geoviewer.
Table 3: Number of pixels used for training, distributed over the whole study area/Tabla 3: Número de pixels utilizados para el entrenamiento, distribuidos en toda el área de estudio

| LULC classes         | Mangrove forest/Forest cover | Open Water Bodies | Urban and built-up areas | Cropland/Grassland/Others | OWV I | OWV II |
|----------------------|------------------------------|-------------------|--------------------------|---------------------------|-------|--------|
| No. of pixels        | 8788                         | 6560              | 1558                     | 9782                      | 7594  | 6900   |
| Area (ha)            | 87.88                        | 65.60             | 15.58                    | 97.82                     | 75.94 | 69.00  |

2.7.1. The geo-tagging process

The GPS mode of the mobile device was activated and each photograph’s coordinates were saved. The metadata for each data file allowed us to obtain the values corresponding to latitude, longitude, date and hour when the photo was taken, among other characteristics.

Also, the photos taken with the mobile application were identified and the specific visual verification points were marked. The specific sampling points that were recorded are distributed along rivers, canals, lagoons and different types of vegetation. The photos helped with the visual interpretation of the LULC map.

2.8. Interactive map design and implementation

2.8.1. Technological considerations in map generation

To create the online interactive map, these free open software tools were used: the PostgreSQL database with the geospatial PostGIS library, the GeoServer map server, Leaflet JS (JavaScript), HTML 5, CSS and JavaScript. Web geospatial services such as WMS layer (Web Map Services) were published, which use mosaics or vector tiles to facilitate the display of map images at different scales.

2.8.2. Simplifying the classification layer on LULC for the web

The process consists of carrying out a map simplification procedure in order to reduce the number of polygons and the attributes of polygons. The threshold of the area to be eliminated is identified. After carrying out several comparative analyses, a generalization was done, eliminating all polygons under 1,300 m$^2$, since each pixel measures 10m by side. To do this, polygon sizes less than 12m are filtered, whose areas are smaller than 1,300 m$^2$. Once the polygons have been eliminated, the option is applied to attach the eliminated area to the polygon that adjusts the greatest contour. A final revision is done comparing the original layer to the layer obtained by means of this process where small variations are observed which generally do not affect the information at the scale this map is generated.

3. Results and Discussion

3.1. About the Training Data Set

3.1.1. Visual analysis of training areas in the 2D scatter plot

To illustrate the reflectance characteristics of different LULC classes, a scatter plot was created with the values of the NIR band and of the red band, as shown in Figure 4. Figure 4 (a) shows only the training sites, indicating the location of each pixel class within this scatter plot, while in Figure 4(b), the same data are included with the Sentinel-2 data from NIR and Red bands. This scatter plot represents vegetation in conditions of low precipitation (see Figure 2).

The NIR vs Red scatter plots show evidence of the vegetation’s behavior, bare soil, and changes in reflectance, among the 6 classes that were used. The X axis (Red Band), represents different types of bare soil. The red color identifies soil pixels belonging to urban and built-up areas. The Y axis (NIR band) shows dense vegetation such as the mangrove cover/forest cover class. Between mangrove cover/forest cover class and urban and built-up areas, other vegetation domains are present (OWV-I y OWV-II). The open water bodies are concentrated in low reflectance values.

3.1.2. Quantitative evaluation of spectral separability

Table 5 shows the separability values, comparing between pairs of classes. In 12 pairs there is a clear spectral separation between neighboring land covers. However, four pairs have values between 1.68 and 1.80, less than 1.90, as recommended by Jeffries-Matusita; this decreased their separability in the final classification result.
3.2. Classification Accuracy

The unbiased accuracy assessment of the resulting classification is shown in Table 6. The error matrix that was generated considered the unbiased area estimates for each class, including a 95% confidence interval (CI) based on both the reference data and the land cover classification. Table 6 shows the results of the Error Matrix.

All classes have a User’s Accuracy (UA) value between (83.3-100%) and Producer’s Accuracy (PA) of (78.8-100%). Urban and built-up areas, open water bodies, OWV I, and OWV II were higher in UA and PA at 90%. For the mangrove forest/forest cover, the UA was 93.5% and PA was 78.8% because some pixels within the training sites had separability with a low Jeffries-Matusita value (<1.90), compared to the Cropland/grassland/Others class. The Cropland/grassland/Others class has UA of 83% and PA of 94.4% and also Jeffries-Matusita values that are slightly lower than 1.90 Mangrove Forest/Forest cover and OWV II. The overall percent accuracy was 90.0%, indicating that the study showed a good correspondence between the reference pixels and the mapped pixels.

3.3. Classification Result

The result of the LULC classification process, obtained with the SVM algorithm, is the result of the condition of the geographic landscape in summer. Figure 5A shows LULC classification for 6 classes, and two examples of image subsets where the classification detail may be seen. Figure 5B shows a Close-up view of classification map of San Pedro y San Pablo river (Cometa Lagoon), and Figure 5C shows a Close-up view of the Grijalva river, specifically. Figure 5B shows a classification error in the water’s surface at Cometa Lagoon, at the entrance to the different lagoon canals, where class OWV-II was assigned. This is due to the fact that the entrances to these canals are not full open water bodies and part of these water bodies are covered by the tree canopy. The best delineation of water bodies is obtained when these are open water bodies; an example can be seen in Figure 5C. It was not possible to separate agricultural lands and pastures, so they remained grouped in one category type. The water body surfaces were separable and identifiable, using the S2 MSI data directly. The urban areas were differentiated from bare land (beach sand) reassigning them to urban and built up areas class in the post processing phase.

Our resulting map showed that the Centla swamp area is a dynamic and heterogeneous territory, with respect to its natural protected areas, as well as to the surrounding area. In the maps that were consulted for this work, the natural protected area is represented as a more homogenous zone, perhaps because the design of these maps and the generalization of the legends is based on the work scale, which leads to having more updated classes at the local level. In the analysis of “Other wetland vegetation”, we found that it was possible to identify two-level types of vegetation for humidity conditions. This may be possible, as stated by Guerra and Ochoa, 2006, since certain types of wetlands are emergent, floating and submerged; this is based on the different kinds of surfaces, such as: the permanently flooded surface, seasonally flooded surface, temporarily flooded surface or occasionally flooded surface (Troche et al., 2018). In this study we present a LULC map obtained from a single-day observation that allowed us to find spectral variations in vegetation. In order to perform wetland follow-up, we propose simple digital image processing methods, such as learning machine algorithms like SVM, and assessing training
Table 5: Pair separation data obtained with Jeffries-Matusita (JM) distance

| Pair comparison                                      | JM value     |
|------------------------------------------------------|--------------|
| *Cropland/Grassland/Others and OWV II                | 1.6828 - 1.8030 |
| *Cropland/Grassland/Others and *Mangrove Forest/Forest cover OWV II and OWI I | 1.9233 - 1.9735 |
| Urban and built-up areas and OWV II and OWV I        |              |
| OWV II and OWV I                                     |              |
| Urban and built-up areas and OWV I                   |              |
| OWV II and Mangrove Forest/Forest cover              |              |
| OWV I and Urban and built-up areas                   | 1.9926 - 2.00 |
| Cropland/Grassland/Others and OWV I                  |              |
| OWV II and Urban and built-up areas                  |              |
| Cropland/Grassland/Others and OWV I                  |              |
| OWV I and OWB                                        |              |
| Cropland/Grassland/Others and Urban and built-up areas |          |
| OWV I and OWB                                        |              |
| OWV I and Mangrove Forest/Forest cover               |              |
| OWV I and Mangrove Forest/Forest cover               |              |
| OWV II and OWB                                        |              |
| Urban and built-up areas and Mangrove Forest/Forest cover |          |
| Cropland/Grassland/Others and OWB                    |              |
| OWB and Mangrove Forest/Forest cover                 |              |

data quality. We showed that in the study area there is spectral separability of classes when analyzing mangrove cover/forest cover class, urban and built-up areas and other cover classes in the NIR vs Red Scatter Plot. This chart may result very helpful when analyzing the greenness and wetness of the vegetation, which could indicate water stress in wetland zones and their surroundings (Ahmadian et al. 2016). Although there are not many cloudless optical images, we suggest using S2 MSI images for local wetland monitoring.

3.4. Interactive website

The final map has an interactive version with photos taken during the fieldwork at the ground level; these photos have a geoposition that provides information on the zone in greater detail. As part of the dissemination of the investigation and in order to share the obtained results, an interactive website was designed and developed. This website is for geovisualization and it is based on free and open technologies that show the LULC classification layer, the field trip routes and the photographs of control points with geo-location, which are associated with the zone’s landscape. By clicking on the map it is possible to access the photographs and a brief description of the landscape. As it can be seen in Figure 6, the LULC map obtained with S2 data is available at the visual browser user interface. Figure 6 is an illustration of the general aspect of the visual browser user interface showing 5 panels with in situ photos of the following classes: Mangrove forest/Forest cover, Cropland/grassland/Others, Other Wetland vegetation and Open Water Bodies, which are shown respectively in panels 1-4. Panel 5 illustrates the whole viewer.

The website uses geospatial web services of WMS type, according to the standard emitted by the Open Geospatial Consortium (OGC). It is implemented with tile caching technology to accelerate and optimize the performance of the visualization display. In addition, technologies such as, html5, CSS, JavaScript, Postgre SQL – PostGIS and php are also a fundamental part of the construction of this tool.

The fieldwork demonstrated the importance of having available photos and ground truth information on the landscape. These show the characteristics of the ecosystems that provide the local communities with environmental services. In order to improve the present study, the photos were taken during a trip along the rivers; enhancing the positional precision of the viewpoint and leading to a direct viewing of the terrain’s characteristics. As perspectives of this work, the acquisition of information from geo-photos needs to be strengthened, following the protocol proposed by Fritz et al., 2019, with photos taken in the cardinal directions.

The fieldwork carried out with the proposed methodology allowed for exploration of extensive wetland areas in a short time. Currently, more research activities are needed to organize the data obtained through geo-tagged photo mapping fieldwork combined with the information obtained from satellites. Future work will focus on a set of sample photos that will be automatically tagged as a function of their visual content.
Table 6: Results of the Error Matrix / Tabla 6: Resultados de la Matriz de error

| Class                              | Area (ha) | ± 95% CI | User’s Accuracy % | Producer’s Accuracy % | Overall Accuracy % |
|------------------------------------|-----------|----------|-------------------|-----------------------|-------------------|
| Mangrove forest/Forest cover       | 55,310.11 | 108.18   | 93.5              | 78.8                  | 90                |
| Cropland/Grassland/Others          | 58,217.56 | 105.40   | 83.3              | 94.4                  |                   |
| Urban and built-up areas           | 1,235.07  | 94.70    | 100.0             | 100.0                 | 100.0             |
| Open water bodies                  | 18,645.87 | 147.02   | 100.0             | 93.8                  |                   |
| OWV I                              | 98,402.13 | 165.90   | 91.0              | 90.1                  |                   |
| OWV II                             | 78,224.82 | 120.73   | 90.0              | 93.6                  |                   |
We present an interactive map that allows for visualization and easy understanding of the geographic landscape, for visual verification purposes.

The resulting geospatial layers of this research exercise may be accessed remotely, with the added advantage of being able to compare with results of similar works on the study zone, or to cross-check with other types of information and generate one’s own products. The published geospatial layers contain detailed information of the LULC classification type in the study zone. This translates into a visual display of a large number of polygons which, managed by the map server with tile caching, increase the response times of the interactive map display.

4. Conclusions

This paper proposes the use of S2 MSI data to generate the LULC map for natural protected regions in Tabasco, Mexico, accompanied by visual verification information with geotagged photos. For this study zone, the use of 6 categories is proposed for identification of land cover types. The study site is a natural protected area with several aquatic and sub-aquatic plant communities and a great surface of water bodies.

This investigation’s results suggest that, as future work, the identification of more class types is possible. In spite of the fact that our study is a particular case of LULC design obtained with a multispectral snapshot with S2 MSI during the summer, the obtained results indicate a correct LULC classification that can
improve the design with more category levels; this will allow the generation of knowledge on wetland cover types, plant communities and surface water characteristics to complement monitoring information and enhance the conservation of resources in this region.

Presently, freshwater ecosystem monitoring is the basis for the SDG 6.6.1 indicator, for the care of inland water bodies and their relationship with inland wetlands and coastal wetlands. With the new free access satellite data, a methodology was proposed to study water body and wetland extension. The results obtained from digital image processing represent the basis for the study of wetlands in this region, which is feasible during dry season or low precipitation wet season. This preliminary work may contribute to the development of an SDG 6.6.1 indicator focused on change of the different covers for the study zone, allowing decision-makers to locate hot spot zones with water stress and encouraging knowledge in this area to protect, maintain and restore freshwater ecosystems.

Acknowledgments

Authors thank the anonymous reviewers, whose valuable comments helped us to improve the document. The authors wish to thank J.I. Chapela and E. Martínez for the support provided to continue with the research activities. Thanks to El Servicio Meteorológico Nacional (SMN) for pluvial precipitation data. The authors thank F. Marroquín-Zavala, who does sport fishing, and who helped with the fieldwork sample collection. The authors declare no conflict of interest.

Funding

This work began within the framework of the FORDECYT-Conacyt project, CentroGeo grant: 297259.

References

Adam, E., Mutanga, O., Odindi, J., Abdel-Rahman, E. M., 2014. Land- use/cover classification in a heterogeneous coastal landscape using Rapid-Eye imagery: evaluating the performance of random forest and support vector machines classifiers. International Journal of Remote Sensing 35 (10), 3440–3458, https://doi.org/10.1080/01431161.2014.903425.

Ahmadian, N., Demeulê, J. A. M., Xu, D., Borg, E., Zöllit, R., 2016. A New Concept of Soil Line Retrieval from Landsat 8 Images for Estimating Plant Biophysical Parameters. Remote Sensing 8 (9), 738, https://doi.org/10.3390/rs8090738.

Bhatnagar, S., Gill, L., Regan, S., Naughton, O., Johnston, P., Waldren, D., 2020. Mapping VEGETATION COMMUNITIES INSIDE WETLANDS USING SENTINEL-2 IMAGERY IN IRELAND. International Journal of Applied Earth Observation and Geoinformation 88, 102083, https://doi.org/10.1016/j.jag.2020.102083.

Bruzzone, L., Persello, C., 2009. Approaches based on support vector machine to classification of remote sensing data. En: Chen, C. H. (Ed.), Handbook of Pattern Recognition and Computer Vision. WORLD SCIENTIFIC, pp. 329–352, https://doi.org/10.1142/9789814273398_0014.

Chasmer, L., Mahoney, C., Millard, K., Nelson, K., Peters, D., Merchant, M., Hopkinson, C., Brisco, B., Niemann, O., Montgomery, J., Devito, K., Cobbaert, D., 2020. Remote Sensing of Boreal Wetlands 2: Methods for Evaluating Boreal Wetland Ecosystem State and Drivers of Change. Remote Sensing 12 (8), 1321, https://doi.org/10.3390/rs12081321.
