Monitoring of Open-Pit Mining Areas Using Landsat Series Imagery (1984–2023) and Cloud Processing

Pau Montero 1, Edgardo Bustos 2, Joan-Cristian Padró 3,4 and Vicenç Carabassa 1,2,*

Abstract: While open-pit mining activities represent one of the human-derived most impactful land cover changes, these changes and the linked restoration processes can be challenging to assess. This article presents a reproducible methodology carried out with cloud processing of satellite images (Google Earth Engine (GEE)) to evaluate the evolution of open-pit mining activities and their restoration in a Mediterranean landscape. For this purpose, the calculation of the normalized differential vegetation index (NDVI) was used to obtain a quantitative parameter to monitor vegetation presence in each extractive area. To validate these results, confusion matrices were performed between the classification obtained in the study and the official land cover mapping, using randomly selected mining areas as test points, with an average accuracy of 88%. According to the methodology used, the surface of areas denuded by mining in the period 1984–2023 has fluctuated over time, with a maximum in 2005 coinciding with the peak of the Spanish construction boom, and a subsequent decrease towards the present. From these results, it can be concluded that Landsat-type data processed using GEE provide a quick and useful tool for monitoring the evolution of mining activity, including restoration trends, becoming particularly valuable for public bodies’ inspections or decision making.

Keywords: GEE; open-pit mining; mineral extraction evolution; integrated restoration; NDVI

1. Introduction

Since the 1970s, environmental remote sensing (RS) has been gaining importance in geospatial studies and has been strengthened with the advent of NASA’s Landsat satellites [1,2], providing the longest time series (1972–today) of worldwide satellite data [3]. These new technologies have been a fundamental support for studies of the earth’s surface serving different scientific disciplines [4–6].

Mining industry activities cause substantial land modifications and sometimes irrepairable damage to the natural and social surroundings. Due to the growing regulation pressure and the increasing environmental awareness of mining impacts, there is an increasing need for consistent land-change and impact assessment methodologies. Multiple works have addressed mining land cover changes through the use of RS and geographic information systems (GIS) [7–9]. In particular the free release of Landsat data from the United States Geological Survey (USGS) archive in 2008 triggered a wide spread of land use and land cover change (LULCC) studies tackling the context of mining from global [10,11] to local [12–14]. Moreover, the official launch of Google Earth Engine (GEE) in 2010 opened the cloud processing capacities to all the Landsat users’ community, increasing even more the use of Landsat time series data. In fact, over the past decade, a number of works have shown the promising potential of the GEE tool combined with Landsat data [15–18].
The joint scope provided by RS and cloud processing tools allows the automation of methodologies for image processing, becoming a valuable and an affordable resource to support land surface monitoring tasks at mining areas [15–18].

Thus, these technologies can become a worthwhile complement to the mining industry, as a key methodological asset with the potential of reducing the cost and risks of field prospecting tasks. For example, this scope could allow early identification of new areas prone to be mined without the need of in situ assessments. This can be done through the detection of features of interest, such as the differentiation of land covers based on reflectance capacities of satellite images [19,20], which can shed light onto areas with mining potential. Traditionally, the binary pixel differentiation in two categories (e.g., active mine or non-active mine) was determined by the histogram shape (e.g., Otsu’s method [21]). In Pericak et al. [16], Landsat series and GEE were used to map likely active mine areas or likely non-active mine areas at the Appalachian Basin (USA), with a method based on the normalized differential vegetation index (NDVI) and greenest-pixel composite (maximum NDVI), obtaining good discrimination results.

Also, the availability of RS and GIS could make an important contribution to meet the monitoring and compliance requirements by providing assessments on the progress of a restored or impacted landscape. In fact, in many countries the law contemplates the need to monitor the correct evolution of mining activities, including their restoration once the extractive activity has been closed. This is the case in Catalonia (Spain, NE of Iberian Peninsula), where the Royal Decree 975/2009 of 12 June 2009 [22] establishes the need to monitor the development and expansion in terms of land occupation of extractive activities. In Catalonia, to carry out such monitoring, field inspections are currently executed by public administration technicians. This involves a high expenditure of human resources and time dedicated to field trips, which can become a limiting factor for public administration bodies. As an addition to the field inspections, previous works have assessed the rates of expansion and restoration of mining activities in Catalonia using drone imagery [23–26]. These studies focus on monitoring and planning of open-pit mines by means of obtaining land cover maps, calculation of volumes, dust affectations, and detection of gully erosion.

While the two approaches combined, field inspections and drone imagery, provide robust land-changes information for monitoring purposes, the data collected often show highly irregular periodicity between inspections, challenging the assessment for compliance and decision making. Therefore, the usage of satellite data could automate and provide periodical and reliable data for public bodies to assess the evolution of mining activities and the level of compliance without requiring field inspections or drone flights.

The objective of this study is to propose and validate a replicable method for monitoring the evolution of mining activities using Landsat time series imagery and GEE cloud computing tools. In particular, the methodology has been applied to all the open-pit mining areas in Catalonia using an annual temporal resolution for efficiency, covering a range of 40 years (1984–2023). This timeline covers the regional restoration regulation life-span for which mining sites are required to be restored (Law 12/1981 from Catalan Parliament [27]). This approach has an immense potential to offer significant advantages for regular and cost-effective monitoring of mining activities and restoration processes. Additionally, this work serves as a practical example of how leveraging GEE tools and satellite data allows users to easily access script updates, including new imagery and areas of interest at low processing costs.

2. Materials and Methods

2.1. Study Area

The area of interest (AOI) corresponds to the Catalonia region (NE Iberian Peninsula). The study is focused on the active mining activities found in the EXTCATA cartographic base [28]. This geodatabase is a vectorial file containing 2196 polygons that represent the mining activities’ public concessions. The attributes of the database are as follows: the exploitation code, the concessionary company, the mine name, the main extraction material,
the state of restoration, the authorization typology, and the municipality. The vectorial layer is georeferenced in the UTM reference system zone 31N Datum ETRS89 (EPSG:25831) and is constituted by attributes that allow the identification of the materials extracted in each mining activity (Figure 1). The aggregated area corresponding to all the mining activities’ concessions reach a total of 16,952 ha (0.5% of the Catalonia region).

Figure 1. Authorized extractive activities in Catalonia (in blue). Source: Extractive Activities in Catalonia, EXTCATA [28].

2. Materials

2.2. Materials Series Satellite Images

In this study, images from the Landsat series catalog since 1984 were used to meet annual monitoring. Specifically, Landsat 5 TM sensor imagery covered the periods 1984–2001 and 2003–2012, while Landsat 7 ETM+ sensor imagery was used for 2002 (due to the lack of quality images of Landsat 5). Landsat 8 OLI sensor images were employed for the 2013–2023 period. These images contain data registered by optical passive remote sensors operating in the visible, near-infrared spectra. Additionally, atmospherically corrected products, with pixel values equivalent to surface reflectance, were utilized (Level 2, Collection 2, Tier 1) [29,30]. These datasets include atmospherically corrected and processed to orthorectified surface reflectance imagery, along with quality bands that facilitate cloud and cloud shadow filtering and masking.

Strips of collected data were packaged into overlapping scenes covering approximately 170 km × 183 km using a standardized reference grid, at a pixel size of 30 m. WRS path-row scenes covering AOI are 198–032, 198–031, 198–030, and 197–01.

2.2.2. Cloud-Based Platform

Google Earth Engine (GEE) [31] is a geospatial information processing tool, either in vector or raster format, which works through a JavaScript or Python programming language. GEE is a cloud-based platform that allows users to analyze geospatial data on a global scale. It leverages Google’s massive computational capabilities making it possible for non-expert users to analyze vast amounts of data [31]. Additionally, GEE grants easy access to a wide data catalog (including Landsat series imagery, MODIS data, or Sentinel
data, among others), as well as allowing code editing and sharing. In this study, GEE was used to massively process RS Landsat data. The generated code aimed at obtaining the calculation of the synthetic annual NDVI for each extractive activity and for the whole mining area, as well as obtaining graphical results and statistical expressions to understand the evolutionary development of mining in Catalonia. These NDVI annual maps were further exported to be edited with a GIS desktop software (QGIS 3.34.8).

2.2.3. Land Cover Maps

For the geographic and reference view on the surfaces, the topographies from the regional cartographic and geological agency (Institut Cartogràfic i Geològic de Catalunya (ICGC)) were used. In addition, for validation purposes, the official land cover maps (MUCSC) [32], acquired in 1987, 1992, 1997, 2002, 2007, 2012, and 2022, were utilized. These land cover maps allow to discriminate the surface of mining soil (bare soil category) of the non-mining areas within the polygons of the EXTCATA cartographic base [28] and then check the accuracy of the NDVI maps obtained with GEE. Most of the unmined land categories correspond to sparsely vegetated, agricultural crops, scrubland, or wooded areas. According to the available land cover maps (1987 to 2022 MUCSC series [32]), in 1987, 43% of the concessions for extractive activities were occupied by rainfed land and only 12% had an exploited surface area. By 2022, the exploited area increases to 19% and the rainfed arable area decreases to 26%.

2.3. Methods

The workflow was addressed combining a cloud computing platform (GEE), a desktop GIS (QGIS), and a spreadsheet (Microsoft Excel), although most of the work is developed in the GEE environment (Figure 2).

Figure 2. Workflow used for images processing and mining areas calculations and validation.
2.3.1. Data Processing with GEE

First, the shapefile with the polygons of all the open-pit mine concessions was imported, corresponding to the AOI. Secondly, the USGS Landsat 5, 7, and 8 Level 2, Collection 2, and Tier 1 [29,30] were imported and filtered by the selected path–rows and dates. In this case, for every year within the 1984–2023 period, the months from April 1st to September 30th were selected, because it is over this period when the vegetation is more active, allowing a better differentiation to bare soil. Images with cloud cover above 90% were discarded. Afterwards, the selected images were cleaned, removing data corresponding to snow pixels, cloud pixels, and cloud shadow mask pixels. These pixels were flagged in the QA band and converted to “no data” values. In a third step, the yearly median of all the images was calculated for each path–row and mosaicked obtaining a synthetic annual image for all the AOI. The annual image was clipped by the AOI polygons. Fourthly, the NDVI for the clipped annual synthetic images was calculated and added as a new band. Finally, the complete imagery is joined in a GEE collection, to plot the NDVI evolution in a linear chart or show the annual NDVI in a GIF. Also, annual imagery is exported in TIFF format to be further analyzed and edited in the GIS desktop. This comprehensive dataset provides valuable insights for end users, who often seek a straightforward interpretation of mining activity evolution, considering both exploited and unexploited or restored areas within the AOI or individual open-pit mine concessions. The code associated with this data processing step can be found at the following link: https://code.earthengine.google.com/350720fd23335ea6478713fb6fa5ab47 (accessed on 18 July 2024).

2.3.2. Data Management

The annual multiband raster obtained from the GEE was processed in a desktop GIS to create detailed cartography for the end user. The NDVI continuous scale values were classified into four categories: <0.1 exploited zone, 0.1–0.2 zone with sparse vegetation, >0.2 zone with vigorous vegetation, and a “no data” category. These thresholds were adapted from the regional cartographic and geological agency (ICGC) categorization [33,34] and other references [35]. Based on the study area characteristics, the NDVI threshold that differentiates likely exploited areas from non-exploited areas was set at 0.1, while the category of 0.1–0.2 is considered a transition zone corresponding to a non-exploited area with sparse vegetation.

Additionally, the study evaluated the methodology by using desktop GIS to assess the accuracy of the results. Test areas were randomly selected as pixel/point locations within the polygons of the mine concessions [28]. Land cover maps were used to segregate the surface of mining domains into a binary classification between mining soil (bare soil) and non-mining soil. The unmined land categories group includes sparsely vegetated soil, agricultural crops, scrubland, or wooded areas. The binary classification resulting maps allowed the calculation of confusion matrices to determine the accuracy of the two generated classes: mining soils and non-mining soils (corresponding to unexploited or restored soils).

Furthermore, the study quantified the areas in hectares corresponding to mining activity and areas without mining activity. This quantification was performed for all the years covered in the study.

3. Results

The results of the calculation of the denuded area corresponding to the exploited areas show a maximum area affected by the extraction of material/rock around 2005. Subsequently, a downward trend is observed, with small fluctuations in the area mined over the last 18 years (Figure 3). On the other hand, the orange color, in Figure 3, shows the bare soil surface extracted from the MUCSC series [32] within the EXTCATA cartographic base [28] polygons.
2.3.2. Data Management

The annual multiband raster obtained from the GEE was processed in a desktop GIS to create detailed cartography for the end user. The NDVI continuous scale values were classified into four categories: <0.1 exploited zone, 0.1–0.2 zone with sparse vegetation, >0.2 zone with vigorous vegetation, and a "no data" category. These thresholds were adapted from the regional cartographic and geological agency (ICGC) categorization [33,34] and other references [35]. Based on the study area characteristics, the NDVI threshold that differentiates likely exploited areas from non-exploited areas was set at 0.1, while the category of 0.1–0.2 is considered a transition zone corresponding to a non-exploited area with sparse vegetation. Additionally, the study evaluated the methodology by using desktop GIS to assess the accuracy of the results. Test areas were randomly selected as pixel/point locations within the polygons of the mine concessions [28]. Land cover maps were used to segregate the surface of mining domains into a binary classification between mining soil (bare soil) and non-mining soil. The unmined land categories group includes sparsely vegetated soil, agricultural crops, scrubland, or wooded areas. The binary classification resulting maps allowed the calculation of confusion matrices to determine the accuracy of the two generated classes: mining soils and non-mining soils (corresponding to unexploited or restored soils).

3. Results

The results of the calculation of the denuded area corresponding to the exploited areas show a maximum area affected by the extraction of material/rock around 2005. Subsequently, a downward trend is observed, with small fluctuations in the area mined over the last 18 years (Figure 3). On the other hand, the orange color, in Figure 3, shows the bare soil surface extracted from the MUCSC series [32] within the EXTCATA cartographic base [28].

Figure 3. Annual denuded surfaces in the open-pit mining areas of Catalonia according to the proposed methodology (blue color). The orange color shows the surface area of the bare soil category in the EXTCATA polygons according to the MUCSC.

Figures 4 and 5 show, as an example, the temporal evolution of NDVI for the specific extraction activity. In the first example (Figure 4), we can observe that the evolution of the mean NDVI increases since it is an activity that ends its exploitation in the 1980s. Meanwhile, in the example depicted in Figure 5, the mean NDVI trend decreases with time up to 2009, indicating that the mining activity land coverage has expanded over time. However, from 2009 onwards restoration activities began, which slightly increased the mean value of NDVI.

Figure 4. Example of the evolution of a restored extractive activity according to the mean NDVI. Blue dotted line indicates linear trend. Classification according to NDVI thresholds: <0.1 exploited zone, 0.1–0.2 zone with sparse vegetation, >0.2 zone with vigorous vegetation. Orthophotography of the year 2023 [33].
up to 2009, indicating that the mining activity land coverage has expanded over time. However, from 2009 onwards restoration activities began, which slightly increased the mean value of NDVI.

Figure 4. Example of the evolution of a restored extractive activity according to the mean NDVI. Blue dotted line indicates linear trend. Classification according to NDVI thresholds: <0.1 exploited zone, 0.1–0.2 zone with sparse vegetation, >0.2 zone with vigorous vegetation. Orthophotography of the year 2023 [33].

Figure 5. Example of the evolution of an extractive activity in exploitation according to the mean NDVI. Blue dotted line indicates linear trend. Classification according to NDVI thresholds: <0.1 exploited area, 0.1–0.2 sparsely vegetated area, >0.2 vigorously vegetated area. Orthophotography of year 2023 [33].

Additionally, GEE allows to generate GIFs that visually represent the evolution of the different covers (NDVI reclassification by categories) (Figures S1 and S2).

The calculated expected error for the methodology used in this study is presented in Table 1 as the global accuracy and Kappa index. For the years 1987, 1992, 1997, and 2002, the overall accuracy is above 90%, while for the years 2007 and 2012, the highest errors in the detection of mining areas were obtained.

Table 1. Global accuracy and Kappa index according to random points of the official land cover maps (MUCSC) [32].

<table>
<thead>
<tr>
<th>Year</th>
<th>Global Accuracy</th>
<th>Kappa Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>1987</td>
<td>0.92</td>
<td>0.62</td>
</tr>
<tr>
<td>1992</td>
<td>0.90</td>
<td>0.61</td>
</tr>
<tr>
<td>1997</td>
<td>0.91</td>
<td>0.68</td>
</tr>
<tr>
<td>2002</td>
<td>0.90</td>
<td>0.67</td>
</tr>
<tr>
<td>2007</td>
<td>0.84</td>
<td>0.54</td>
</tr>
<tr>
<td>2012</td>
<td>0.79</td>
<td>0.46</td>
</tr>
<tr>
<td>2017</td>
<td>0.88</td>
<td>0.59</td>
</tr>
<tr>
<td>2022</td>
<td>0.88</td>
<td>0.60</td>
</tr>
<tr>
<td>Average</td>
<td>0.88</td>
<td>0.60</td>
</tr>
</tbody>
</table>

4. Discussion

The analysis of the average NDVI results for each mining activity domain and each year allows the detection of the evolution of a mining activity. Namely, an increase in the NDVI value would indicate that the mining area is under a restoration process, while a decrease in the average NDVI value shows an expanding mining activity.

The overall accuracy obtained from the results of the study ranged between 79% and 92% and Kappa indices ranged from 0.46 to 0.68 (see Table 1). These ranges indicate that there are, for some of the years, notable differences between the classification made in
the land cover maps taken as reference [32] and the classifications obtained through the proposed methodology. Therefore, the use of the NDVI calculation and the grouping of the data sets with a fixed threshold has presented limitations for an optimal discrimination between the areas corresponding to mining extraction and the areas where mining activity does not take place, neither those that are restored or in the restoration process. This limitation is especially remarkable for the years 2007 and 2012, which correspond to years with negative precipitation anomalies, i.e., “dry” years with precipitation below the climatic mean (average between the years 1961 to 1990) [36].

Moreover, when zooming into the location of the erroneous test points of the years with the worst results coinciding with the “dry” years, we observe that most of them are located in areas corresponding to agricultural fields. This implies that part of the confusion comes from the erroneous detection of agricultural fields with little or no vegetation as mining areas, as it has been observed also in other studies [14]. Thus, the accuracy of the methodology is reduced under the scenarios of dry years and mining domains with agricultural fields affected by drought. On the other hand, in mining activities located around areas with abundant vegetation (e.g., forests, scrubland, or grasslands), it is possible to obtain reliable results on the expansion or contraction of the surface affected by open-pit mining.

Therefore, these results show that the NDVI index is subject to the variations of the climate itself [37,38], and therefore, this effect must be considered when quantifying the total area exploited. In order to integrate the climatic sensitivity into the methodology, a combination of a vegetation index using the shortwave infrared (SWIR) bands could be used. Compared to red and NIR, SWIR bands do not have as much interference between climatic variations due to temperature or precipitation. Previous works have shown higher accuracy of vegetation classification and landscape change detection using the SWIR-based normalized difference vegetation index (SNDVI) compared to NDVI [39]. In addition, another benefit of using SWIR-based indexes is the higher accuracy in identifying organic soils [40], which could improve the accuracy for areas corresponding to agricultural fields within the mining domain.

Alternatively, adaptive thresholds could be used to overcome NDVI threshold limitations and erroneous categorization. For example, thresholds based on $k$-means unsupervised clustering could allow the adjustment of the thresholds according to different climate zones or precipitation regions, as performed in other works [41].

Also, it is important to note that the NDVI averages per year are not calculated with the same number of images. As a result, some years may have distorted values if there are fewer images available. This limitation poses a challenge, as it depends on the cloud presence. In this study, Landsat collection data was filtered from April to September to obtain annual averages. To address this issue and minimize cloud cover, satellite data could be filtered over a longer period to generate a cloud-free synthetic image by calculating the mean pixel values from selected images.

Working with the highest NDVI values according to the available images could serve as a solution for obtaining a better contrast between vegetation and bare soil areas. However, this methodology does not allow detecting the year of the beginning of mining activity with such accuracy [14].

5. Conclusions

This study proposes a methodological approach which combines the use of historical data from the Landsat time series with cloud processing by GEE to track and assess the evolution of land changes in open-pit mining activities. This scope was applied to open-pit mining areas in Catalonia using an annual resolution in a range of 40 years, in order to test and validate the benefits and limitations of the methodology.

The usage of GEE for data processing has allowed a satisfactory monitoring of the changes for the exploited and unexploited or restored areas with $88 \pm 4\%$ of global average accuracy and $0.60 \pm 0.07$ Kappa index.
The methodology presents limitations in accurately detecting mining coverage for specific scenarios. Namely, the results identified a higher incidence of incorrect classification for those pixels corresponding to (1) years with negative precipitation anomalies and (2) agricultural lands found within the mine concession domain. To overcome this limitation, alternative indexes based on other spectral bands (e.g., SWIR) should be also assessed in further research. Also, given the heterogeneity of landscape and precipitation modes, adaptive thresholds could also play as an improvement in accuracy in discriminating the exploited and non-exploited areas.

Overall, the results showed that this methodology can be a promising tool in providing regular and cost-effective monitoring of land changes, being a future cornerstone for decision making at a business, environmental, and administrative level.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/land13081301/s1, Figure S1. Example of the temporal evolution of an open-pit mining activity in restoration process. Classification according to NDVI thresholds: <0.1 mined area (red color), 0.1–0.2 sparsely vegetated area (yellow color), >0.2 vigorously vegetated area (green color), Figure S2. Example of the temporal evolution of an open-pit mining activity under an exploitation expanding process. Classification according to NDVI thresholds: <0.1 mined area (red color), 0.1–0.2 sparsely vegetated area (yellow color), >0.2 vigorously vegetated area (green color).

Author Contributions: Conceptualization, P.M., E.B. and J.-C.P.; Methodology, E.B. and J.-C.P.; Validation, P.M. and E.B.; Formal analysis, P.M. and J.-C.P.; Data curation, E.B. and J.-C.P.; Writing—original draft, P.M. and E.B.; Writing—review & editing, P.M., J.-C.P. and V.C.; Visualization, P.M., E.B. and J.-C.P.; Supervision, V.C.; Project administration, V.C.; Funding acquisition, V.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Government of Catalonia through the project “Research and Innovation in the process and restoration of extractive activities” and by the European Union through the LIFE project “New Life for Drylands” (LIFE20 PRE/IT/000007).

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

References


Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.