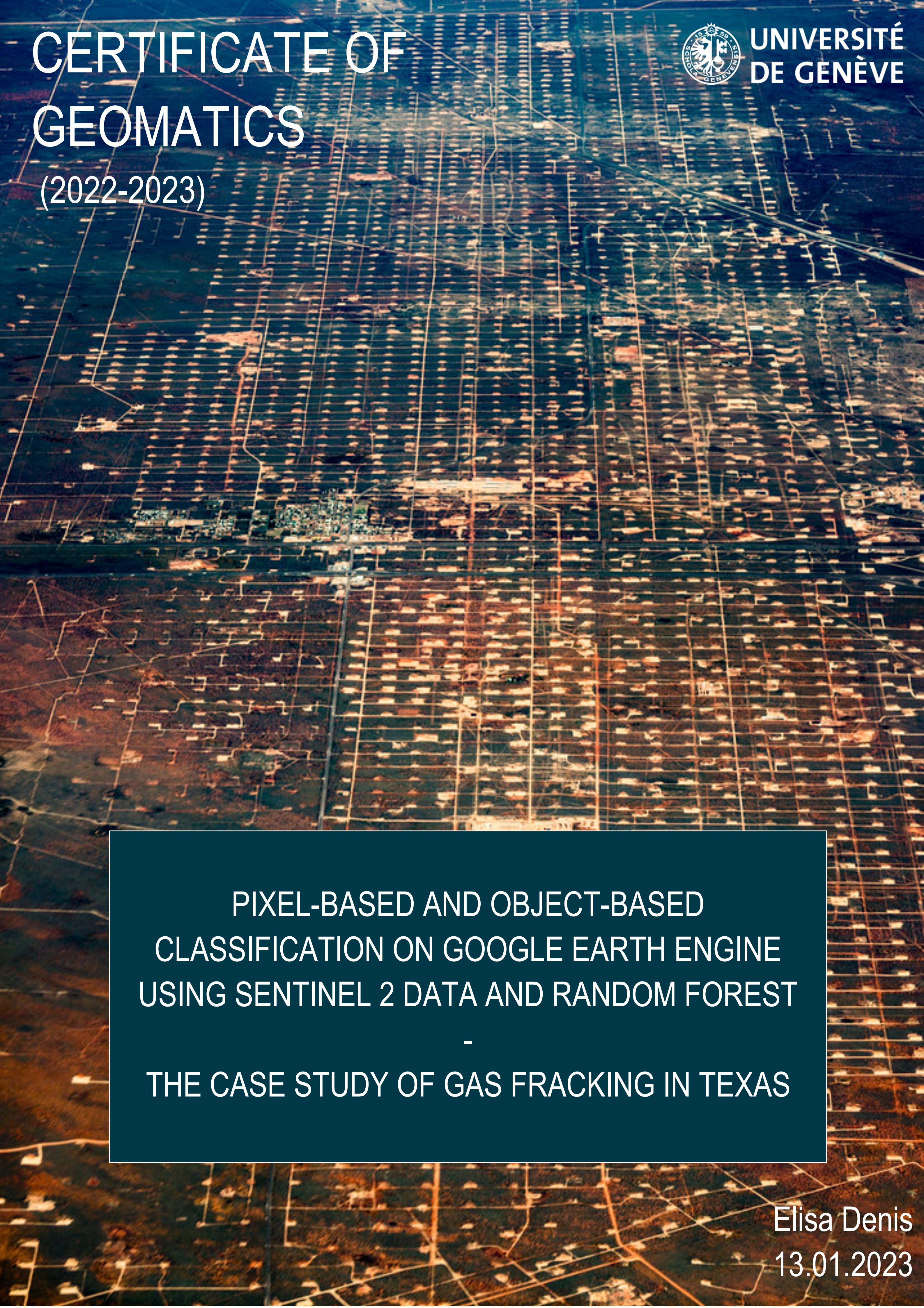

CERTIFICATE OF GEOMATICS

(2022-2023)

UNIVERSITÉ DE GENÈVE

PIXEL-BASED AND OBJECT-BASED CLASSIFICATION ON GOOGLE EARTH ENGINE USING SENTINEL 2 DATA AND RANDOM FOREST

-

THE CASE STUDY OF GAS FRACKING IN TEXAS

Elisa Denis

13.01.2023

Executive summary

Introduction

Fossil fuel production causes a lot of environmental impact and hydraulic fracturing, or "fracking" is among the dirtiest sources of energy. This drilling technique, which injects pressurised liquid in shale or tight rock enables the extraction of natural gas or petroleum. The impacts on the landscape are also noticeable as these extractive pads scatter arid plains with large rectangular areas which can be easily spotted from the sky. Remote sensing and the technological improvement of satellite imagery are efficient techniques to detect this land degradation and its expansion.

Objectives

In response to this, being able to map accurately and rapidly the location of fracking areas provides a cost-efficient monitoring tool. Powerful Geographical Information System (GIS) software now enables experts to map large areas and classify land use using machine learning. The goal of this study is to use Google Earth Engine (GEE) to test different classification methods over distinct areas of Texas to assess which method works best and if a model can be scalable outside of the arid regions of this state.

Method

The proposed method aims to test a pixel-based classification over the entire state of Texas and then three trial areas located in the Permian basin to test a sharpening of the methodology with the addition of Object-Based Image Analysis (OBIA). These two classification techniques are further refined with a mask layer of the Normalized Difference Vegetation Index (NDVI) of the area. The goal of this study is to accurately classify fracking areas with the random forest algorithm, while figuring out which methodology provides the best results.

Results

The pixel classification shows better results than the pixel+obia classification regardless of the area of interest tested with an accuracy of 0.809 and 0.752 respectively. Additionally, results are more precise for the trial zones than for the full state of Texas which goes against the assumption that the model could be scaled up.

Discussion

Several challenges can be highlighted in this study. In the pixel classification, roads and fracking pads get close results due to the similar spectral characteristics they share which reduces accuracy in the results. The segmentation from the OBIA does not sharpen the model as linear features get broken down into a lot of segments that prevents representative clustering. Finally, OBIA was not computed on the entire state of Texas due to computation power limits experimented on GEE. This can be bypassed by exporting results, which was not the aim of this study, as everything sought to be done on GEE only.

List of important acronyms

BAEI - Built-Up Area Extraction Index
DBSI - Dry Bare Soil Index
GEE - Google Earth Engine: Web-based platform which allows users to access extensive geospatial data and perform analysis for free. Most users perform land use classification using the JavaScript language.
OBIA - Object-Based Image Analysis: Method of classification which segments an image into objects of homogeneous characteristics.
PCA - Principal Components Analyis: Statistical tool capable of identifying the variables responsible for most variations within a sample.
NDBI - Normalized Difference Built-up Index
NDTI - Normalized Difference Tillage Index
NDVI - Normalized Difference Vegetation Index
RF - Random Forest: Machine learning supervised classification technique that uses various decision trees.
SNIC - Simple Non-Iterative Clustering is a clustering method which provides a grid of pixels and then expands from the centre of each pixel adding in the closest spectrally matching pixels first.

Table of Contents

List of Tables

List of Figures

Acknowledgement

I am really pleased to have been given the chance to work at UNEP-GRID and meet wonderful experts.

I could not have undertaken this journey without Gregory Giuliani, who generously provided patience, expertise, and invaluable positivity. I also wanted to thank Denisa Rodila for her precious knowledge, feedback and snack offers.

The certificate has been an amazing learning opportunity and a great place to make new friends, which I am truly grateful for.

1. Introduction

Fossil fuel production is a topic scrutinised for different reasons. From sustainable generation of energy to the rise of geopolitical tensions, to socioeconomic considerations, it has gained a lot of interest from the general audience through various angles. Even though the reserves of these non-renewable resources such as crude oil, natural gas or coal are declining, their exploitation remains largely superior to renewable alternatives at the cost of the environment. For instance, natural gas exploration not only affects ecosystems through soil and vegetation loss but also through the contamination of watersheds and the pollution of the atmosphere due to gas flaring on-site. Scientific methods are necessary to monitor and analyse the impacts of this non-renewable way to produce energy.

Remote sensing devices have facilitated the understanding of earth's natural resources management both from an exploration perspective (e.g., discovery of new oil drilling sites) and a conservation one (e.g., identification of areas to be protected). Moreover, technical improvements of satellites have improved the quality, breath and span of aerial imagery, making earth observations more accurate and detailed. The latter provides valuable data which can be processed to evaluate natural resources depletion and monitor the impact of extractive activities on the environment.

Due to an exponential increase in the volumes of spatial data generated (i.e., 10 TB of earth observations data/day, once the Sentinel satellites are completely operational), experts need machine learning instruments to analyse it (UK Parliament, 2020). Artificial intelligence provides undeniable benefits for treating such volumes of information, discovering patterns, generating predictions of the environment and monitoring certain parameters (e.g., drought predictions). Presently, machine learning assists environmental experts in a variety of topics, ranging from weather forecasting, to spotting illegal fishing activities to monitoring land use alterations (e.g., gas fracking, deforestation), etc.

In that context, the objective of this study is twofold. First, it aims to develop and test the Random Forest Algorithm on Google Earth Engine to classify fracking areas. Second, it seeks to compare pixel-based and object-based classification to figure out which technique works best for this purpose. The hypothesis is that object-based image analysis will be able to sharpen the pixel-based classification due to the very distinct pattern of fracking areas. The deliverable highlighted in Table 1 will complement this study for illustrative purposes.

Table 1: Deliverables associated with this study

Script	Name	Comment	Link
Script 1	Pixel-based classification on a large area: Texas	This script performs a pixel-based classification over the entire state.	https://code.earthengine.google.com/87b80782ce407e34779b665990283e0d
Script 2. a.	Pixel + OBIA - classification on a small area: trial_zone_1	Script 2.a. and 2.b. are the same. Only the area of interest changes.	https://code.earthengine.google.com/97901d0199166ed9727b96b1eff1fd32
Script 2. b.	Pixel + OBIA - classification on a small area: trial_zone_3	-	https://code.earthengine.google.com/bcbeef55b208b9f70cc89b935438ff81
Script 3	Comparison of indices	This script compares the NDVI, NDBI, DBSI, NDTI and BAEI over one reduced AoI (trial_zone_1) to understand which index could provide the best mask.	https://code.earthengine.google.com/5f26b6a73ea5d76b047ee6c776d2dc65

This project was conducted within The Global Resource Information Database - Geneva (GRID-Geneva), a partnership between the United Nations Environment Programme (UNEP), the Swiss Federal Office for the Environment (FOEN) and the University of Geneva (UniGe) which gathers data scientists specialising in the processing of satellite imagery, modelling of geospatial data and creation of visualisations platforms. This project was completed as part of the Complementary Certificate in Geomatics delivered by the University of Geneva.

2. Gas fracking background

In 2021, 82% of the world's primary energy came from oil, gas and coal (BP, 2022). Even though demand for fossil fuel should peak before mid-2030, exploration and production are still growing (IEA, 2022). One technique to recover oil and gas is hydraulic fracturing, or fracking. This geochemical process consists in drilling into shale rock and injecting a high-pressure mix of water, sand proppants and chemicals to extract fossil fuels. Even though this method is banned in several European countries because of the precautionary principle, others such as the USA still rely on this method to extract fossil fuels.

The impacts of gas fracking are numerous and hit many aspects of the ambient world (anthroposphere, lithosphere, biosphere, etc). Table 2 below summarises the main impacts identified in the literature.

Table 2: Compilation of fracking-induced impacts (Meng, 2017)

Area	**Main impact only**	**Explanation**	**Source**
Anthroposphere	Land cover change	Fracking pads and transportation networks alter the landscape. Sites are exploited for a couple of years only which emphasises the spread of fracking sites over large areas.	Meng (2014)
Atmosphere	Greenhouse effect	Loss of carbon dioxide sinks due to deforestation and methane emissions from fracking activities.	Karion et al., 2013
Biosphere	Species distribution and diversity	Fracking sites require a change of land use (clear cut of a forest, paved grasslands, etc) which destroys wildlife and ecosystems.	Meng (2014)
Lithosphere	Soil and ground alteration	Change in geomorphological characteristics due to fluid injections which can result in changes in seismic activities.	Ellsworth et al. (2012)
Hydrosphere	Freshwater consumption	The contamination of groundwater and surface water by hazardous substances such as benzene and toluene is one of the main impacts due to leakages in the installations.	Meng (2017)

Various stakeholders require information to better understand and assess the impacts of gas fracking. Being able to monitor these consequences is key to protect people and the environment efficiently and to implement measures and policies which better legislate fracking operations. While several monitoring activities can happen on the ground (e.g., water quality check), others can be carried out from the sky. In fact, in 2021, the European Space Agency started using satellites to track methane leakages from fracking installations (ESA, 2021). For example, satellites with integrated spectrometers, such as Sentinel-5P, can map atmospheric gases on a daily basis. However, the spatial resolution of this tool is still relatively high (7 km × 5.5 km for methane). It thus requires the participation of on-the-ground experts and airborne instruments to effectively map out the impacts of gas fracking.

Remote sensing is a geospatial technology which gathers reflected and emitted radiations of an area or an object with satellites or airplanes and provides its characteristics without physical contact (USGS, n.d.). Remote sensing has long been recognized for providing insightful measures of the environment such as mapping forest fires or tracking land cover changes (e.g., expansion of a city, deforestation). Land use mapping is a common application since the expansion of remote sensing and Geographical Information System (GIS) instruments, the drop in costs and time-efficiency of the processes (Rawat, 2015). More recently, the increase of open-source (Landsat and Sentinel) and high-resolution data has facilitated the monitoring of land cover and land use changes.

Gas fracking monitoring using remote sensing techniques has been well-documented in the literature. From the observation of land cover ‘s dynamics of shale developments on drylands (Wang, 2021) to the monitoring of pollutants emanating from hydraulic fracture activities (Asrar, 2018). However, much less attention has been given to automating the detection of fracking sites on designated areas to perform real-time monitoring of land use changes. The company Antarctica Capital (previously Descartes Labs) has developed a machine learning algorithm capable of detecting well pads in Eastern USA through Google Earth Engine (GEE) (Thomson, 2021). Their methodology is using Google’s deep learning TenserFlow running with the Earth Engine Python API where the model trains itself and refine outputs over time.

GEE is a web-based platform created in 2010, which allows users to access extensive geospatial data and perform analysis for free. Most users perform land use classification using the JavaScript language (Tassi et al., 2021), however, this platform also supports TenserFlow workflows using the Earth Engine Python API at a cost.

The aim of this project is to use the free functionalities of GEE with JavaScript to detect well pads in Texas and compare different classification options. This workflow seeks to reproduce to some extent the work performed by SkyTruth without the deep learning aspect where the algorithm is training itself after each iteration to improve its classification. Developing an algorithm which uses free functionalities seeks to further increase transparency via an automation of the detection of fracking sites for an easy and open-source access.

3. Methods

Study area

Script 1, 2.a., 2.b. & 3 can be found in the annex file.

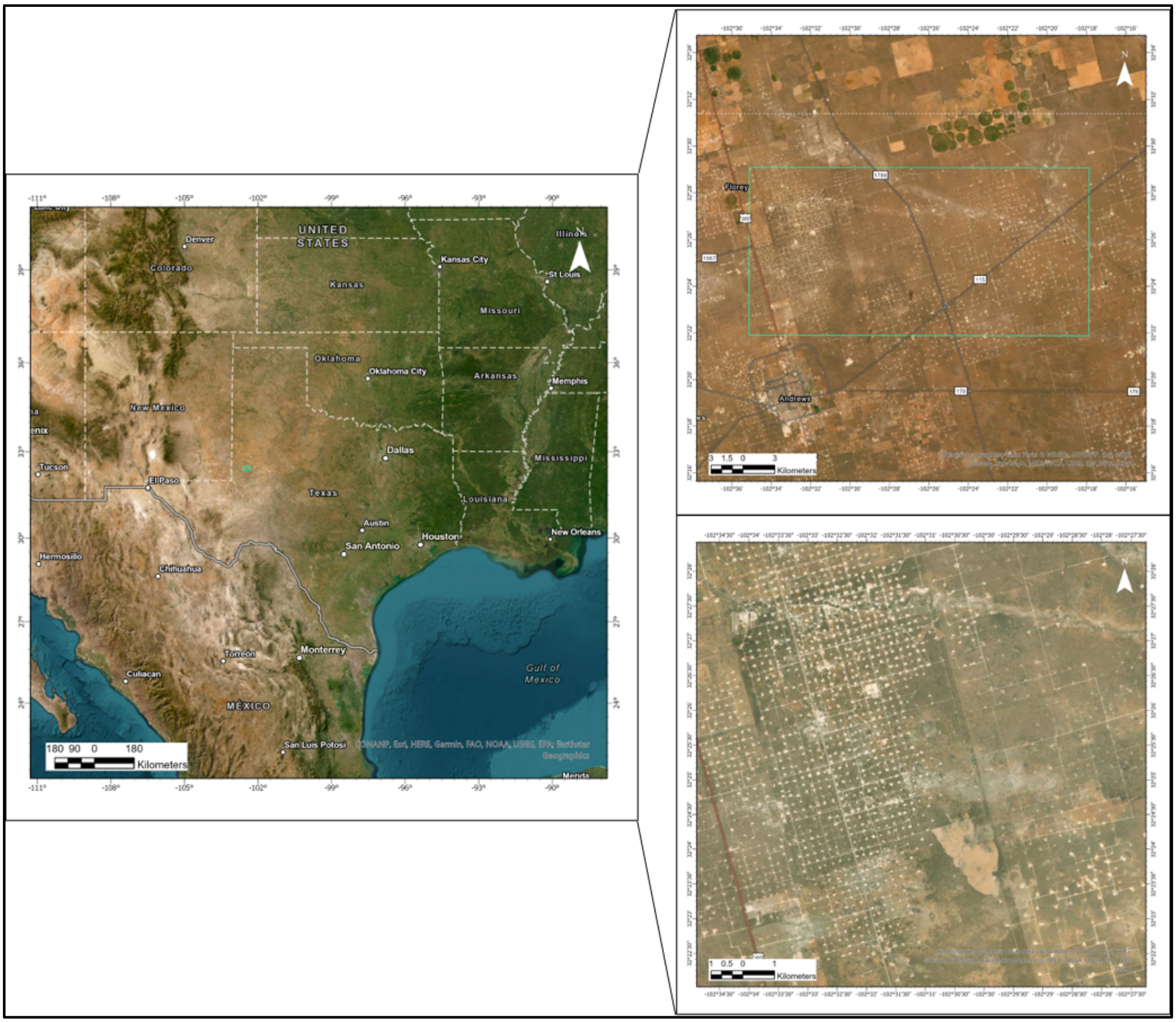

Figure 1: Study area - Texas, USA (left) and the example of one trial zone (right)

The sites chosen (see Figure 1) for this study can be divided into two categories. The first one was the large area that served for the test of the pixel-classification (see Script 1 in Annex). The other ones were the three trial zones that were scattered in Texas and would be used as experiments for the pixel-based + OBIA classification as can be seen in table 3 below (see Script 2.a. and 2.b. in Annex).

Table 3: Characteristics of the two areas used in this study

Name	**Large area**	**Trial zones (1,2,3)**
Description	Texas	Smaller areas of Texas located in the Permian basin
Size of the area (sq.km)	695,662	Approx. 350
Classification performed	Pixel-based	Pixel-based + OBIA

The state of Texas was selected because of its prevalence for fracking. This method of gas extraction has been present in the U.S. since 1860 and has extended to many states within the country (Ridlington et al., 2016). Texas has ranked the highest producers in the country in terms of well numbers with over 80’000 active wells in 2021 (Caldwell, 2021). Additionally, the three trial zones chosen are located in the Permian Basin, in Western Texas and occupy approximately 350 sq.km each (Figure 2). The importance of oil and gas in the Permian Basin is due to the organisms (e.g., coral reefs which covered the seabed over 265 million years ago). This sedimentary basin covers more than 220,000 sq.km, is the largest petroleum reserve of the United States and has produced close to 75 trillion cubic feet of gas since the beginning of its exploitation (Leder, 2021).

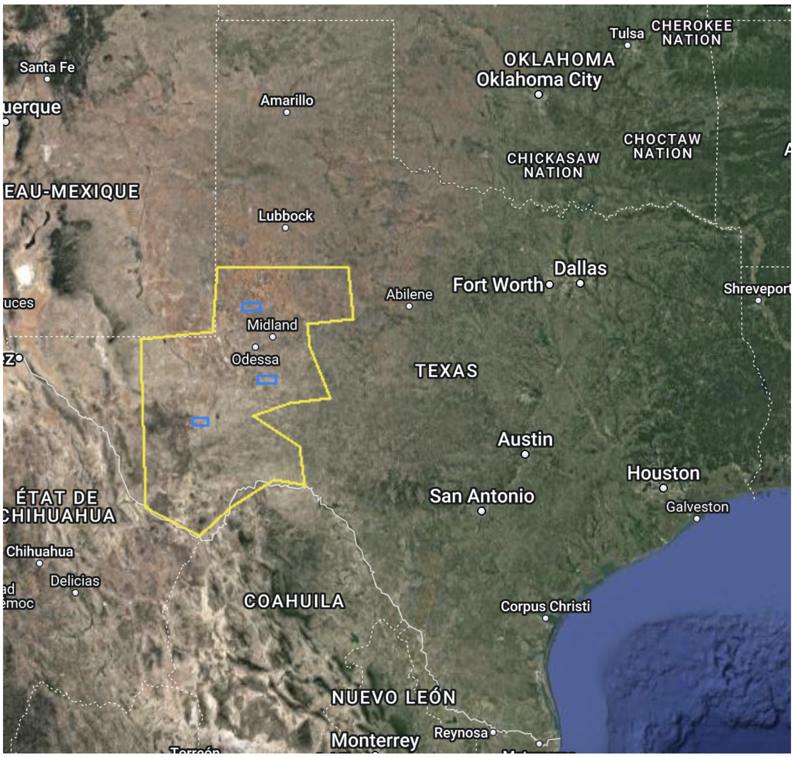

Figure 2: The three trial zones located in the Permian basin

Sentinel 2, cloud-free imagery and pre-processing

A Sentinel-2 image from 2022 was chosen because of its 10-m resolution for certain bands (i.e., blue, green, red) compared to the 30-m resolution of Landsat 8 on GEE. Sentinel-2 was launched in 2015 by the European Space Agency to collect earth observations at a high resolution (10-m to 60-m). It provides spectral data over 13 bands and has an average revisit time of 5 days which provides very accurate and current data (ESA, 2022). The image was obtained directly from the GEE interface and six bands were used for the analysis as can be seen in Table 4.

Table 4: Characteristics of Sentinel-2 MSI used in this study

Sensor	Period	Band	Use	Wavelength	Resolution	Provider
Sentinel-2 MSI - MultiSpectral Instrument, Level-2A	2022-04-01- 2022-09-15	B2	Blue	496.6nm (S2A) / 492.1nm (S2B)	10 m	ESA
		B3	Green	560nm (S2A) / 559nm (S2B)	10 m	
		B4	Red	664.5nm (S2A) / 665nm (S2B)	10 m	
		B8	NIR	835.1nm (S2A) / 833nm (S2B)	10 m	
		B11	SWIR 1	1613.7nm (S2A) / 1610.4nm (S2B)	20 m	
		QA60	Cloud mask	-	60 m	

In preparation of the classification, a cloud mask was created. This serves several purposes such as: reducing noise, avoiding radiometric distortion of the surface and erasing black pixels induced by cloud shadows (Puteri, 2020). The 'QA60' (60-m resolution) band collects both dense clouds and cirrus and classifies pixels accordingly (e.g., bit 10: mask for opaque clouds; 0: no opaque clouds, 1: opaque cloud presents). This classification is then used to detect these pixels and remove them with a mask layer.

To avoid data gaps that could be caused by the cloud mask, images were selected over a period of five months and the image collection was then reduced with the median function to create a new composite. The latter computes for every pixel the median of all values. Finally, the image was clipped to the region of interest (here: Texas) to save processing power. An explanation of the main steps of the methods can be found in Figure 3.

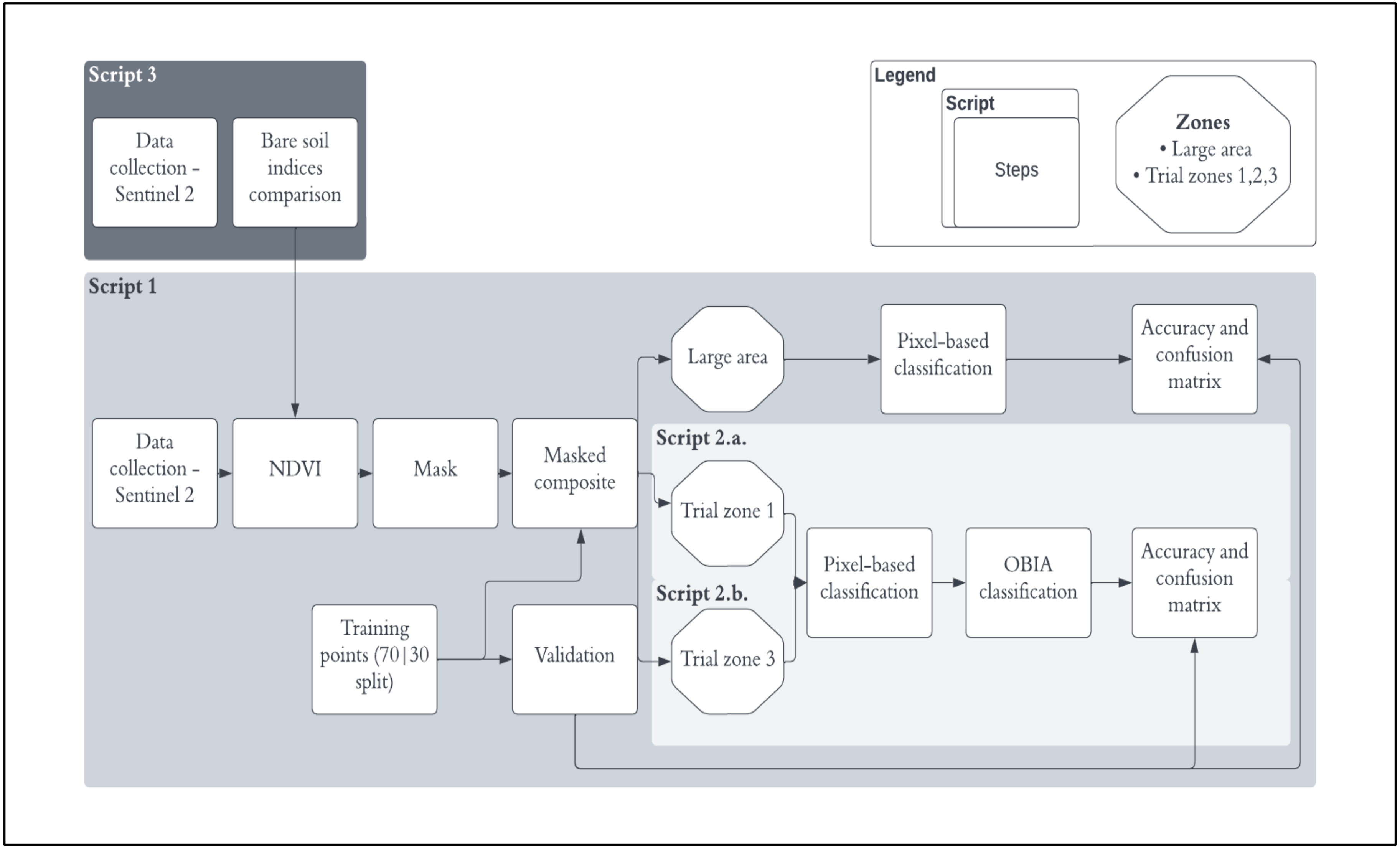

Figure 3: Flowchart of the methodology

NDVI and mask

To improve the results of the classification, it was necessary to calculate an index and use it as a mask to remove certain pixels from the area of interest. This step was crucial to ensure that the classifier would decipher the difference between the spectral properties of pixels of bare land and fracking areas to determine which indices was best to use as a mask, a test was performed on a smaller area of interest (trial_zone_1). Nguyen et al. (2021) used the NDBI to identify bare land features during a fallow period while Osgouei et al (2019) used the NDTI to distinguish bare land from built up areas. Both indices seemed adequate to test on the arid soil of fracking areas in Texas. Additionally, NDVI is commonly used as a threshold and due to its interesting results, it was worth testing it too (Weier and Herring, 2000). Finally, DBSI and BAEI showed interesting results in the context of arid regions, even when applied to built-up areas ((Nguyen et al., 2021; Bouzekri et al., 2015) which also supported an inclusion in the indices test. The script for this step can be found in annex (Script 3).

Table 5: Comparative table of indices

Indices	**Name**	**Specificity**	**Sentinel_band_calculation**
NDVI	Normalized Difference Vegetation Index	Vegetation index for all regions (Weier and Herring, 2000)	(NIR-RED)/(NIR+RED) =(B8-B4)/(B8+B4)
NDBI	Normalized Difference Built-up Index	Built-up areas index (Zheng et al., 2021)	(SWIR1-NIR) /(SWIR1+NIR) =(B11-B8) /(B11+B8)
DBSI	Dry Bare Soil Index	Bare soil index for arid climatic regions (Nguyen et al., 2021)	((SWIR1-GREEN)/(SWIR1+GREEN))-((NIR-RED)/(NIR+RED))
NDTI	Normalized Difference Tillage Index	Index which can highlight differences between bare land and built-up areas (Osgouei et al., 2019)	(SWIR1-SWIR2) /(SWIR1+SWIR2) =(B11-B12) /(B11+B12)
BAEI	Built-Up Area Extraction Index	Built-up areas index in arid region (Bouzekri et al., 2015)	(RED+0.3)/(GREEN+SWIR1) =(B4+0.3)/(B3+B11)

Trial_zone_1 was used to test the indices and four classes were designated for the classification (fracking_area = 1; non-fracking/bare land= 2; roads = 3; vegetation= 4). For each test, the index was computed and then the latter was used as a band to perform the classification. Results from the classification can be found in Figure 4.

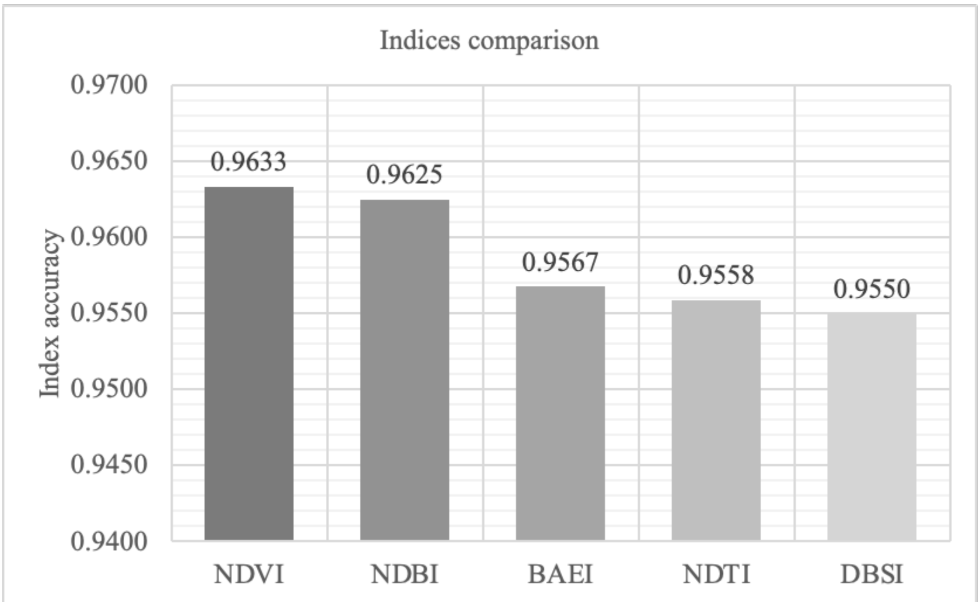

Figure 4: Indices comparison on trial_zone_1 (see below)

The accuracy of the NDVI index was higher than the other indices so this index was selected for the classification. In addition, only using the index as a band in the classification was not sufficient to reduce classification errors. Since it was not necessary for the classification to get all the different classes. The goal to remove everything that was not bare land from the image was twofold. First, it would reduce computation power as all the non-relevant pixels would be masked. Second, it would limit classification errors for the algorithm to only look at the difference between fracking areas and supposedly bare land. To do so, the NDVI was calculated, and the results were re-classified as intervals. Only the interval which took into account the fracking zones was selected to create the mask.

The first step consisted in computing the mean and standard deviation of the NDVI band to ensure that the interval would be appropriate. The second step was to use the vegetation classification according to typical NDVI values from (Dazelios et al., 2001) to narrow the range of the interval to include only bare soil and sparse vegetation. Finally, the third step was to select fracking areas with the inspector tool on GEE to get the pixel values (NDVI band) of fracking zones to further narrow down the interval and remove. The final interval ranged between 0.025 and 0.09. This was an iterative process to find the appropriate thresholds that would strike a good balance between inclusion of all the fracking zones while still excluding enough bare land to avoid classification errors afterwards. All the values outside of this range were masked using the GEE masking function. The outputs of this index test can be found in Figure 5.

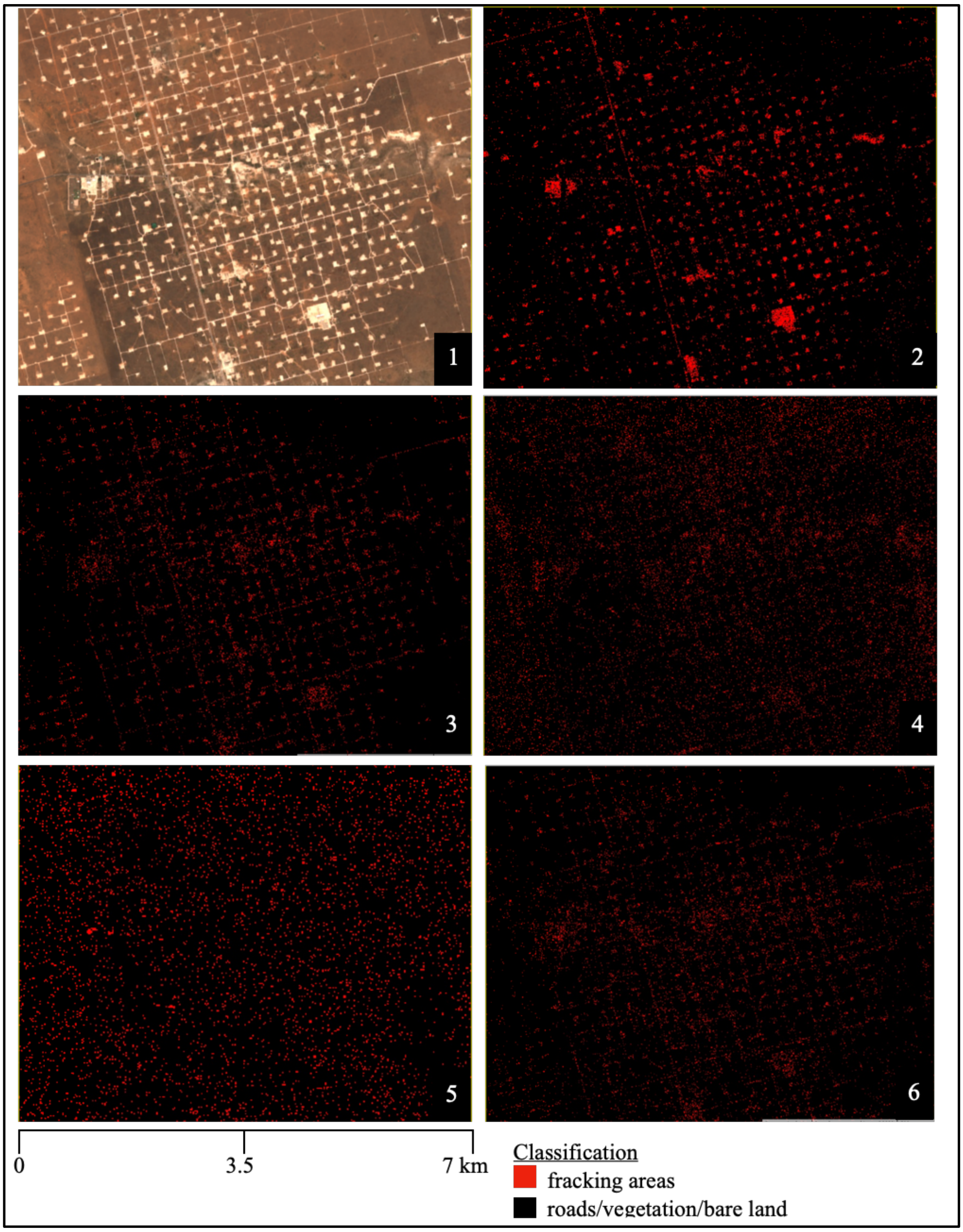

Figure 5: Classification of indices (1) RGB 432, (2) NDVI, (3) NDBI, (4) BAEI, (5) NDTI, (6) DBSI

Training areas

The training areas were separated into two subsets. The first one will now be referred to as the *large area* and the second one will be referred to as the *trial zones* as per Table 6 below. Four categories were created to capture the variety of items that could be found in an arid region such as Texas: fracking areas, bare land, roads, vegetation (see Figure 6).

Table 6: Characteristics of the areas used in this study

Name	**Large area**	**Trial zone 1**	**Trial zone 2**	**Trial zone 3**
Description	Texas	Smaller area of Texas located in the Permian basin	Smaller area of Texas located in the Permian basin	Smaller area of Texas located in the Permian basin
Coordinates	Texas	[-102.58262023267683, 32.370009087637726]; [-102.2784362727159, 32.370009087637726]; [-102.2784362727159, 32.48157979972241]; [-102.58262023267683, 32.48157979972241]	[-103.43638635532736, 30.68230000292648]; [-103.16859460728048, 30.68230000292648]; [-103.16859460728048, 30.797677986617597]; [-103.43638635532736, 30.797677986617597]	[-102.32658964193173, 31.31075549655534]; [-102.01897245443173, 31.31075549655534]; [-102.01897245443173, 31.41570608565165]; [-102.32658964193173, 31.41570608565165]
Specificity	-	Test zone (h0). This zone contains fracking areas. The indices and the mask were calculated using this zone as a reference.	Zone without fracking areas.	Zone with fracking areas. Used to test how the index and the intervals chosen for the index from zone 1 will work on this zone (h1).
Size of the area (sq.km)	695,662	Approx. 350	Approx. 350	Approx. 350
Training zones per category (1- fracking, 2- bare land, 3- roads, 4- vegetation)	2400, 100, approx. 2000, 110	100, 30, 30, 30	100, 30, 30, 30	100, 30, 30, 30
Classification	Pixel-based	Pixel-based, OBIA	Pixel-based, OBIA	Pixel-based, OBIA

For the large area, more than 2000 fracking areas (1400 manually, over 1000 from Descartes Lab) were created as polygons. The roads were downloaded for certain areas of Texas from OpenStreetMap (major roads and secondary roads), classified under QGIS and then imported on GEE and others were directly drawn manually as polylines. Vegetation and bare land were also drawn as polygons manually and areas that were wrongly incorporated in the NDVI mask were used to further refine the classification instead of choosing areas that the mask had already rejected. Fewer zones were selected for the vegetation and bare land category as they were already well excluded with the mask. However, fracking areas and roads were seen as difficult to differentiate on the masked composite, which is why these two categories had the highest numbers of training areas. Finally, due to the large number of training zones drawn, a sample size of 512 pixels was set to avoid computation power issues.

The same categories were used for the trial zones (1,2,3); however, fewer training polygons were drawn as the areas were significantly smaller. Besides, the results from the large area and the small ones were not comparable because the randomization in the classification of the large area did not necessarily consider all the training areas from the trial zones.

Finally, a 70|30 (testing|validation) randomisation was applied to all the sites to have a better machine learning rate (Nguyen et al. 2021) and a confusion matrix was generated to get the results and accuracy of the classification.

Figure 6: Examples of training points used for the classification

Pixel-based classification

For both areas, a pixel-based classification was performed which allocated each pixel to a particular class. This allowed the separation of pixels labelled as fracking zones from others. A Random Forest (RF) classifier was applied, which is a traditional machine learning supervised classification technique that uses various decision trees. The decision trees selected subsets of the training data and made predictions on the results of the decision trees. Other supervised classifiers such as Support Vector Machine and Naive Bayes could have been used but Random Forest is faster and more robust especially if the landscape trained is similar to the training areas and it is efficient on large datasets and maintains relatively good accuracy even if some data is missing (Pelletier et al., 2016). Finally, in this analysis, bands B4 (RED), B8 (NIR), B11 (SWIR1) and the NDVI were used, and 300 trees were computed to improve classification accuracy (Liu & Zhang, 2019).

Despite the benefits of RF, researchers have found that pixel-based classifications can result in a 'salt and pepper' outcome when dealing with high-resolution imagery, which decreases the accuracy of the classification (Weigh & Riggan, 2010). Object-based image analysis (OBIA) can provide an alternative to this by integrating shape identification into the classification and classifying an entire area as a single vector (Gorelick, 2018). Moreover, fracking zones have a particular square shape which provides an interesting testbed for the OBIA method. It was thus decided to run a pixel-based classification on the large area and to test a combined method (pixel-based classification + OBIA) on the zoom area to see if more accurate results were achieved. The OBIA method could not be performed on the large area due to the very high memory and CPU requirements of the classification.

Object-based image analysis

Object-based image analysis has gained rapid momentum in the remote sensing field since the beginning of the 2000s. This method is based on segmentation (Hay & Castilla, 2008) which divides an image into regions of homogenous feature (pixels) characteristics and goes further than pixel-based classification as it considers spatial properties of objects (van der Werff & van der Meer, 2008). This technique solves the limitation that looking at spectral values only has by integrating the shape of objects in the analysis. However, a spectral classification or shape-based approach only is less accurate than a combined spectral-shape classification (van der Werff & van der Meer, 2008).

The most crucial step of OBIA is the segmentation as it directly affects the quality of the classification results (Blaschke et al., 2008). In this analysis, the superpixel seed location spacing, in pixels, was a size of 10. Various segment sizes were considered for the analysis but 10 remained the chosen option due to the small size of the objects considered (fracking zones) and to strike a balance between over-segmentation of the image and the accuracy of the results.

Superpixels put points on the image (seed grid) and then expand to collect pixels around these points to get shapes. Superpixels are not the objects but the reduced object. Simple Non-Iterative Clustering (SNIC) was chosen over Simple Linear Iterative Clustering (SLIC) as it provides better results (Achanta & Susstrunk, 2017) and is the only segmentation method available on GEE. SNIC provides a grid of pixels and then expands from the centre of each pixel adding in the closest spectrally matching pixels first (pixels with the minimum distance, Gorelick, 2018).

Several parameters were then defined for the SNIC applied on the masked composite as can be seen in Table 7. For the size, a large value was set to capture all the homogenous areas (as clusters were relatively homogenous due to the use of the masked composite). Compactness, connectivity and neighbourhood size were set based on an iterative method to visually capture what was most accurate and representative of the area as a lack of literature exists on the topic for the parameter setting of OBIA on GEE. Finally, the seed grid previously generated was used as an input parameter for the classification. Doing so preserved the spatial features of the underlying data.

Table 7: OBIA parameters used in the classification

Parameter	Explanation (Achanta and Susstrunk, 2017)	Value
Size	Seed location spacing of super-pixels. The seed represents the centre of a cluster.	20
Compactness	Compactness factor (distance weighting) (choice between a purely spectral or purely spatial segmentation)	5
Connectivity	Connectivity (if pixels touch each other)	8
Neighbourhood Size	Amount to extend each tile (overlap) when computing the cluster	48
Scale	Resolution	10

After generating the superpixels, clusters and parameters for the classification, spatial information and statistics were collected. The GEE function ee.reduce.components used the objects (their shape) and found in each tile the homogeneous pixels and applied a reducer to everything underneath those pixels. This allowed the collection of the area, perimeter, width and height of each cluster. Finally, this information was used as a layer in the classifier to compute the OBIA.

4. Results

Large area - pixel-based classification (script 1)

The first results generated were from the large area (entire Texas) and were further broken down into three sites. The NDVI results show how the Eastern side of the state has greater levels of vegetation than the Western side, which is more desertic. The classification results in Figure 7 show that most fracking areas can be found on the Western part of the state as well as scattered areas remaining across the state. A lot of areas are not classified on the Eastern part as the masked composite removed a lot of areas already. In addition, despite a relatively high accuracy and low standard deviation of the validation points over ten runs (average = 0.862, stdDev = 0.270) - see Table 8, visual results below show that some results are inaccurate. This can be attributed to the imbalance between the different categories of training sites. This can also be also due to a small number of training / validation points, to the actual selection of training / validation / test dataset, and many other factor. This is very dependent on the use case. It also must be noted that the accuracy metric is not always the best choice to assess the performance of a model.

Table 8: Classification accuracy results for the entire of Texas (large area)

	Run 1	Run 2	Run 3	Run 4	Run 5	Run 6	Run 7	Run 8	Run 9	Run 10	**Average**	**Std.Dev**
Training	0.973	0.976	0.977	0.978	0.978	0.976	0.977	0.977	0.976	0.973	**0.992**	**0.030**
Validation	0.777	0.750	0.744	0.744	0.773	0.753	0.777	0.745	0.767	0.731	**0.862**	**0.270**

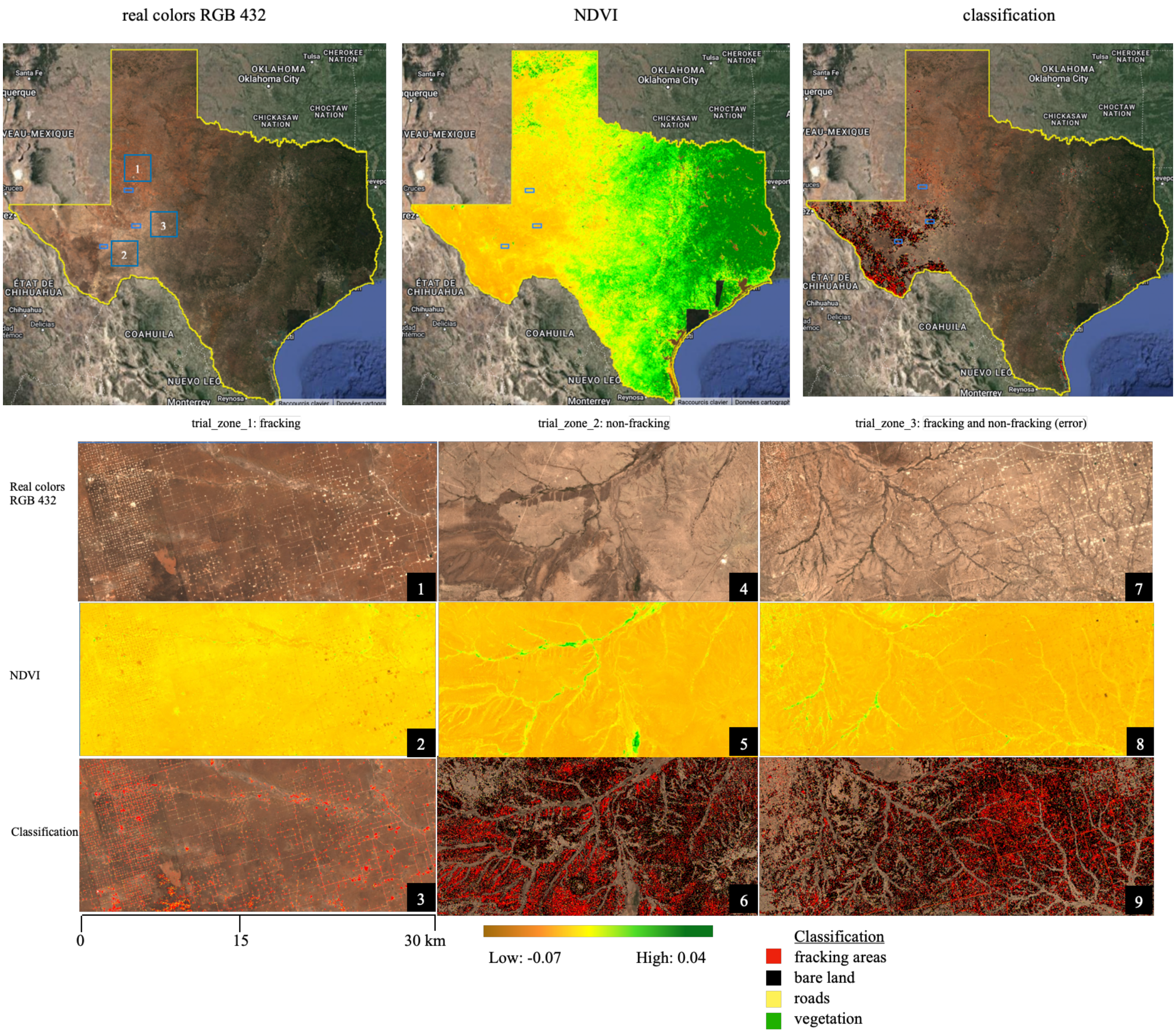

Figure 7: Large area results (top) and trial zones results (bottom): three close-ups of similar arid regions distinguishing the NDVI index and classification results.

Furthermore, trial zones show variations in results for the results of the pixel-based classification. The first trial zone shows relatively accurate results as it is possible to see which zones are fracking areas instantly whereas the second and third trial zones display most of the zone as a fracking zone which is incorrect and is likely due to the pixel values of bare land being close to the ones of fracking zones. Bare land, roads and vegetation do not necessarily appear on the final classification (3,6,9) because the composite mask masked most of the areas that were not fracking zones already, removing pixels that fell outside of the NDVI threshold set. Besides, trial_zone_2 still displays fracking zones even though this area does not have any fracking sites as can be seen in the real colour image, which shows that there is a high likelihood that trial_zone_3 contains errors that will be discussed below.

Trial zones - pixel-based classification and pixel-based classification + OBIA

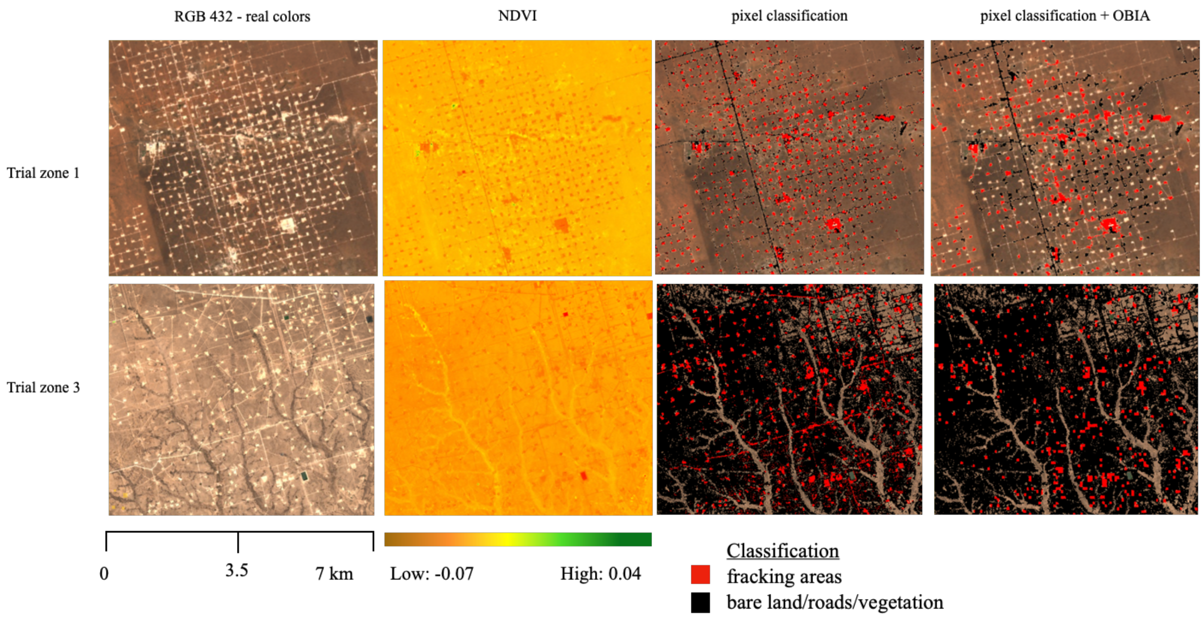

Figure 8: Close-up results of the trial zones (RGB 432, NDVI, pixel classification, pixel + OBIA classification (from left to right). 4 classes are classified but results are aggregated in two categories for visualisation purposes to highlight any error in fracking areas classification.

A second analysis was run only on trial zones 1 and 3 as they contain fracking areas while trial zone 2 does not. The OBIA was added to the pixel classification. Training points were set in each area so that the classification was performed on training sites from the corresponding area. Results show very little difference between trial_zone_1 and trial_zone_3, as can be seen on Figure 8, which means that when appropriate training sites were set, classification was more accurate, regardless of the type of classification performed. In fact, using training sites that were specific to the interest sites was more accurate which underlines the idea that this classification of fracking areas was not necessarily scalable.

Table 9: Trial zones - classification accuracy assessment

Training (validation)-Script 2.a. 2. b and the data for the ten runs can be found in annex

Classification method	**Statistics**	**trial_zone_1**	**trial_zone_3**
Pixel-based	**Average**	0.993 (0.809)	0.995 (0.805)
	StdDev	0.004 (0.016)	0.002 (0.012)
Pixel-based+OBIA	**Average**	0.992 (0.752)	0.998 (0.642)
	StdDev	0.003 (0.021)	0.002 (0.026)

Overall, trial_zone_1 gave better results than trial_zone_3 for the pixel classification (80.9% and 80.5% accuracy respectively), but by a very small margin (see Table 10). In addition, trial_zone_1 performed better than trial_zone_3 for the pixel + OBIA classification (75.2% and 64.2 % accuracy respectively). It is interesting to note that pixel+OBIA classification have more variation in their results than the pixel classification only which is more homogeneous as can be seen in Figure 9.

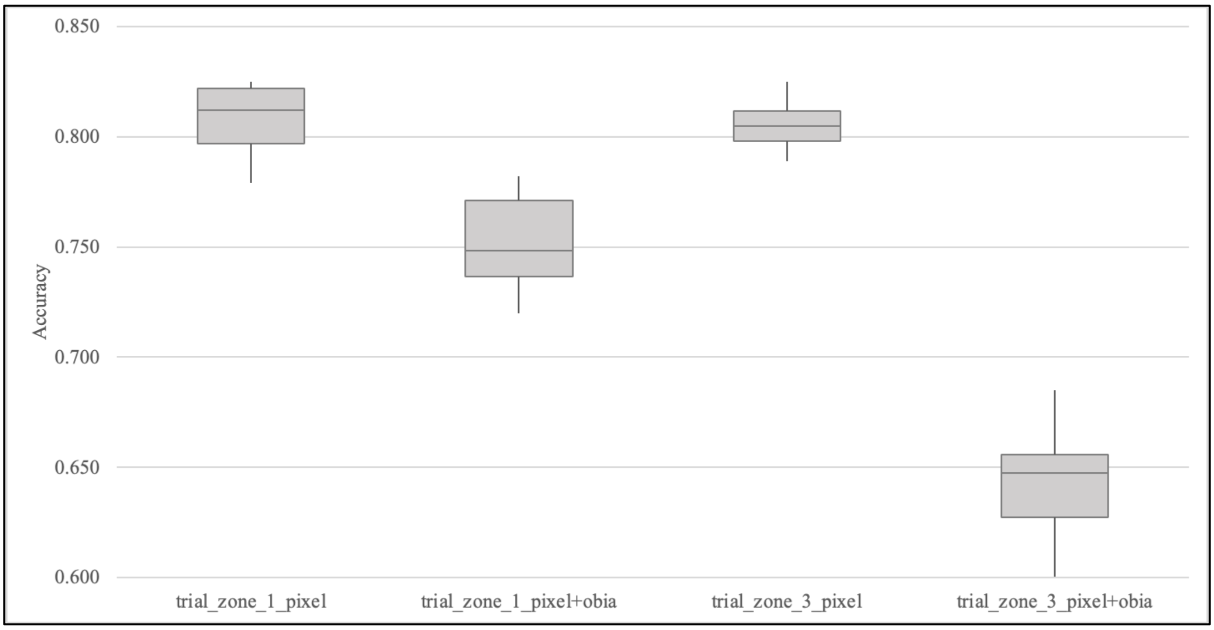

Figure 9: Accuracy of classification methods on different trial zones

These better results for trial_zone_1 could be due to the quality of polygons drawn or for the OBIA classification, it could be attributed to the less homogeneous areas which created high segmented objects. Nevertheless, the variations between the trial zones remained very minimal to draw significant conclusions. This better result in trial_zone_1 could also be attributed to the NDVI mask which was based on this area to determine the interval. Finally, in both scenarios, pixel-based classification performed better than pixel+OBIA, which goes against the hypothesis that OBIA would sharpen the classification. It also highlights the idea that OBIA is not a necessity when scaling up the model to a larger area (i.e., entire of Texas).

Table 10: Confusion matrices for trial zones 1 and 3 (pixel-based pixel-based + obia classification) – Validation data

trial_zone_1 - Pixel based

	Frackin_areas	Non_ frackin/ bare land	Roads	Vegetation	Total	Accuracy
Fracking_areas	252	0	39	0	291	86.6%
Non_fracking/ bare land	9	51	19	0	79	64.6%
Roads	57	2	200	0	259	77.2%
Vegetation	0	0	0	0	0	0%
Total	318	53	258	0	629	-
Accuracy	79.2%	96.2%	77.5%	0%	-	**77.8%**

trial_zone_1 - Pixel based + OBIA

	Fracking_ areas	Non_ fracking/ bare land	Roads	Vegetation	Total	Accuracy
Fracking_areas	144	0	43	0	187	77.0%
Non_fracking/ bare land	1	39	19	0	59	66.1%
Roads	62	0	86	0	148	58.1%
Vegetation	0	0	0	0	0	0%
Total	207	39	148	0	394	-
Accuracy	69.6%	100%	58.1%	0%	-	**71.5%**

trial_zone_3 - Pixel based

	Fracking_ areas	Non_ fracking/ bare land	Roads	Vegetation	Total	Accuracy
Fracking_areas	300	2	26	0	328	99.3%
Non_fracking/ bare land	0	121	87	0	208	58.1%
Roads	21	63	320	0	404	79.2%
Vegetation	0	0	0	0	0	-
Total	321	186	433	0	940	-
Accuracy	93.4%	65%	73.9%	0	-	**78.8%**

trial_zone_3 - Pixel based + OBIA

	Fracking_areas	Non_ fracking/ bare land	Roads	Vegetation	Total	Accuracy
Fracking_areas	181	7	77	0	265	68.3%
Non_fracking/bare land	0	127	65	0	192	66.1%
Roads	4	108	254	0	366	69.3%
Vegetation	0	0	0	0	0	-
Total	185	242	396	0	823	-
Accuracy	97.8%	52.4%	64.1%	-	-	**68.2%**

Additionally, confusion matrices show that fracking areas were often confused with roads and non-fracking areas were also confused with roads but to a lesser extent (86.6% and 64.6% accuracy respectively for trial_zone_1) as can be seen in Table 10. This pattern could be observed regardless of the classification method used. This is due to spectral information being relatively similar for roads and fracking areas. Indeed, roads are not paved, they are tracks in the sand which have pixel values close to the sandy areas around fracking equipment. Besides, the object-based classification did not manage to remove the roads as linear features can end up broken in different clusters instead of one long object. This is due to the superpixel grid which works better for repetitive patterns such as a large area broken down by fields rather than road networks. The latter also happened on fracking zones as can be seen on Figure 10, which is also why OBIA did not perform better than pixel-based classification. Finally, vegetation did not appear on the confusion matrices as the region is arid and most areas have been masked with the NDVI mask.

Figure 10: From top to bottom: RGB 432, clustering, obia classification (in red, the fracking area is correct, in black the fracking area is misclassified as non-fracking).

5. Discussion and challenges

Methods of classification

In the context of fracking, pixel-based classification provided relatively good results on the trial zones compared to the entire of Texas and this is largely due to tailored and more accurate training zones and the use of the masked composite. Using an NDVI mask was convenient in an area that had very few spectral variations as areas that stood out could be easily excluded (e.g., vegetation patches). However, the homogeneity between classes (i.e., roads and fracking areas) was an issue that the mask itself could not solve. Furthermore, using a mask worked well for a definite area but would not be well scalable as spectral information for the threshold would vary between areas. Thus, expanding the pixel-based + OBIA classification back to the full area (entire of Texas) was not performed. First, because of GEE's computation power limitation. It could have been overcome by using a lower resolution/bigger tile size, however the accuracy would have drastically reduced and not been representative of the objects studied. This is because fracking areas are around 70*70m (4.9 sq.m). Second, because the results on the trial zones showed that no improvements were made when using OBIA over pixel-based classification.

Despite initial belief that OBIA would refine the classification, results showed that adding OBIA to the pixel-based classification did not improve the results but rather worsened them for the small areas. This was surprising as the object studied (fracking areas) were relatively homogenous in their squared shape and provided a pattern that was easily recognizable even with the human eye. However, this relatively poor performance was largely because segmentation can be inconsistent, and one object can be divided into different clusters. Table 11 below highlights some of the advantages and drawbacks of using one method over the other in the context of fracking.

Table 11: Main advantages and drawbacks of classification methods for fracking on GEE

	Pros	Cons
Pixel-based classification	No segmentation errors	‘Salt and pepper effect’
Object-based classification	Segments contains not just spectral information but also spatial (e.g., perimeter) and statistical ones (standard deviation value)	Lack of computation power Cluster heterogeneity

A scalable methodology?

Results showed that scaling up the methodology to expand it to other states would be feasible but not without major modifications that would be time consuming. First, it would require drawing new polygons specific to the area of interest to ensure that training points are representative of the spectral characteristics of the new area studied. Second, the NDVI interval would need to be re-evaluated to fully capture the information of fracking zones and exclude the areas that could cause misclassification. As the goal here is not to classify the entire zone but rather only identify fracking areas, using a mask to exclude certain pixels does not affect the results negatively. However, roads are still problematic due to the similarity of their spectral characteristics with fracking areas. One way to overcome this and automate the classification would be to directly download the road network of the area of interest from OpenStreetMap and either mask them or classify them. This could reduce the noise when classifying the entire area. The limitation to this is that even though OpenStreetMap provides granular information some tracks are not drawn on their system due to their small size.

The number of classes used for the classification could also be discussed. It was decided to choose four classes to take into account fracking areas, bare land, roads and vegetation but perhaps a different classification could have modified the results if roads and bare land were under the same class. However, it was interesting to have roads as separate as to test how OBIA would segment them.

Future improvements

To go one step further and try to perform OBIA again to see if more accurate results can be achieved, a Principal Component Analysis (PCA) could be performed. The latter would reduce redundancy in information and increase computation power which would compensate for GEE's issue. Another way to bypass GEE's limitation would be to export the results to perform the OBIA on another software. However, this was not the goal of this research. Finally, to improve the object-based analysis, a rule could be added to assign pixels to a group if the majority of the pixels of a segment already fall in one class. Xiong et al. (2017) performs this operation to classify croplands in Africa. This could overcome the classification errors of mixing roads and fracking areas due to their similar spectral characteristics.

6. Conclusion

This paper presented the results of a pixel-based classification and an object-based classification over fracking areas in Texas. Overall, the results for the pixel-based classification were better than the ones for the object-based classification. Remote sensing and GEE provide an efficient and quick way to spot fracking areas despite certain accuracy errors. However, computation power remains a big limit to perform object-based classifications and scaling-up of models.

Reference

Achanta, R. & Susstrunk, S. (2017)., 'Superpixels and Polygons using Simple Non-Iterative Clustering', *CVPR*.

Asrar, M., Hui Wen, A., & al. (2018). Space technologies for monitoring health and environmental impact of hydraulic fracturing. *The Lancet, 2* (11), 469-470. https://doi.org/10.1016/S2542-5196(18)30223-7

Bartesaghi-Koc, C. & al. (2019). Mapping and classifying green infrastructure typologies for climate-related studies based on remote sensing data. *Urban Forestry & Urban Greening, 37.* 154-167. 10.1016/j.ufug.2018.11.008

Blaschke, T., Lang, S. & Hay, G.J. (2008). *Object-based image analysis: spatial concepts for knowledge-driven remote sensing applications*. Springer Science & Business Media.

Blaschke, T., Lang, S. & Hay, G. (2008). *Spatial Concepts for Knowledge-Driven Remote Sensing Applications.* Springer.

Bouzekri, S. & al. (2015). A New Spectral Index for Extraction of Built-Up Area Using Landsat-8 Data. *Journal of the Indian Society of Remote Sensing*, 43. 867-873.

BP. (2022). Statistical Review of World Energy*,* 71th edition. Retrieved from: https://www.bp.com/content/dam/bp/business-sites/en/global/corporate/pdfs/energy-economics/statistical-review/bp-stats-review-2022-full-report.pdf

Caldwell, N. (2021). 12 states where tracking is more prevalent. *Stacker*. https://stacker.com/science/12-states-where-fracking-most-prevalent

Thomson, D. (2021). Natural gas and oil extraction with hydraulic fracturing – from exploration to completion. *SkyTruth.* Retrieved from: https://skytruth.org/2021/01/natural-gas-and-oil-extraction-with-hydraulic-fracturing-from-exploration-to-completion/

Dazelios et al. (2001). Cotton yield estimation based on NOAA/AVHRR produced NDVI. *Physics and Chemistry of the Earth Part B Hydrology Oceans and Atmosphere.* 26. 247-251. 10.1016/S1464-1909(00)00247-1

ESA. (2021). OBSERVER: Tracking Fracking: How Copernicus is helping trace methane emissions from hydraulic fracturing. Retrieved from: https://www.copernicus.eu/en/news/news/observer-tracking-fracking-how-copernicus-helping-trace-methane-emissions-hydraulic

ESA. (2022). Sentinel- 2. Colour vision for Copernicus. Retrieved from: https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-2

Gorelick, N. (2018) EEUS 2018- Image Segmentation and object-based methods. Earth Engine User Summit. Retrieved from: https://www.youtube.com/watch?v=2R0aTaMtYTY

Hay, G. & Castilla, G. (2008). Geographic Object-Based Image Analysis (GEOBIA): A new name for a new discipline. In *Object-Based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications* (pp. 75-89). Springer.

IEA. (2022). *World Energy Outlook*. Retrieved from: https://www.iea.org/reports/world-energy-outlook-2022

Leder, R. (2021). There's a ticking climate time bomb in West Texas. *Vox*. Retrieved from: https://www.vox.com/22407581/gas-texas-biden-climate-change-methane-permian-basin

Liu, H. & Zhang, Y. (2019). Selection of Landsat8 Image Classification Bands Based on MLC–RFE. *Journal of the Indian Society of Remote Sensing, 47. 439-446.*

Meng, Q. (2017). The impacts of fracking on the environment: A total environmental study paradigm. *Science of the Total Environment, 580.* 953–957.

Nguyen, C. & al. (2021). A Modified Bare Soil Index to Identify Bare Land Features during Agricultural Fallow-Period in Southeast Asia Using Landsat 8. *Land, 10(3)*. 1-18.

Nguyen, Quang Hung, et al. (2021) "Influence of data splitting on performance of machine learning models in prediction of shear strength of soil." *Mathematical Problems in Engineering* 2021.

Osgouei, P. & al. (2019). Separating Built-Up Areas from Bare Land in Mediterranean Cities Using Sentinel-2A Imagery. *Remote Sens, 11*(3). 34.

Pelletier, C., Valero, S., Inglada, J., Champion, N. & Dedieu, G. (2016). Assessing the robustness of Random Forests to map land cover with high resolution satellite image time series over large areas. *Remote Sens, 187.* 156–168.

Puteri, S. (2020). The Difference of Filtering Clouds and Masking Cloud in Google Earth Engine. Retrieved from: https://sryhandiniputeri.medium.com/the-difference-of-filtering-clouds-and-masking-cloud-in-google-earth-engine-260744bcc600

Rawat, J.S. & Kumar, M. (2015). Monitoring land use/cover change using remote sensing and GIS techniques: A case study of Hawalbagh block, district Almora, Uttarakhand, India. *The Egyptian Journal of Remote Sensing and Space Science, 18(1).* 77-84.

Ridlington, E. & al. (2016). Fracking by the Numbers the Damage to Our Water, Land and Climate from a Decade of Dirty Drilling. Retrieved from: https://environmentamerica.org/wp-content/uploads/2016/04/Fracking-by-the-Numbers-vUS.pdf

Tassi, a. & al. (2021). Pixel- vs. Object-Based Landsat 8 Data Classification in Google Earth Engine Using Random Forest: The Case Study of Maiella National Park. *Remote Sens, 13*(12). 2299.

The Wilderness society. (n.d.). The truth about fracking and the environment. Retrieved from: https://www.wilderness.org/news/article/truth-about-fracking-and-environment

UK Parliament (2020). Remote Sensing and Machine Learning. *PostNote, 628*. Retrieved from: https://researchbriefings.files.parliament.uk/documents/POST-PN-0628/POST-PN-0628.pdf

USGS. (n.d.). What is remote sensing and what is it used for? Retrieved from: https://www.usgs.gov/faqs/what-remote-sensing-and-what-it-used

Van der Werff, H.M.A & Van der Meer, F.D. (2008). Shape-based classification of spectrally identical objects ISPRS. J*ournal of Photogrammetry and Remote Sensing, 63(2)*. 251-258.

Wang, H. (2021). The Impact of Shale Oil and Gas Development on Rangelands in the Permian Basin Region: An Assessment Using High-Resolution Remote Sensing Data. *Remote Sensing, 13*(4). 824.

Weier, J & Herring, D. (2000). Measuring Vegetation (NDVI & EVI). *NASA Earth Observatory*, Washington DC. Retrieved from: https://www.sciencedirect.com/topics/earth-and-planetary-sciences/normalized-difference-vegetation-index

Weih, R.C. & Riggan, N.D. (2010). Object-based classification vs pixel-based classification: comparative importance of multi-resolution imagery. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXXVIII-4/C7*

Xiong et al. (2017) Nominal 30-m Cropland Extent Map of Continental Africa by Integrating Pixel-Based and Object-Based Algorithms Using Sentinel-2 and Landsat-8 Data on Google Earth Engine. *Remote Sensing. 9*(10):1065. DOI:10.3390/rs9101065

Zheng, Y. & al. (2021). An improved approach for monitoring urban built-up areas by combining NPP-VIIRS nighttime light, NDVI, NDWI, and NDBI. *Journal of Cleaner Production, 328*.

Annexes

Annex 1. Script 3 - Confusion matrix obtained with NDVI

	Fracking_areas	Non_fracking/bare land	Roads	Vegetation	Total	Accuracy
Fracking_areas	129	0	11	0	140	92.1%
Non_fracking/bare land	0	610	30	2	642	95%
Roads	5	9	1122	3	1139	98.5%
Vegetation	0	13	10	321	344	93.3%
Total	216	162	579	1	2265	-
Accuracy	96.3%	96.5%	95.7%	98.5%	-	**96.33%**

Script 3 - Confusion matrix obtained with NDBI

	Fracking_areas	Non_fracking/bare land	Roads	Vegetation	Total	Accuracy
Fracking_areas	117	0	23	0	140	83.6%
Non_fracking/bare land	0	621	18	3	642	96.7%
Roads	5	13	1118	3	1139	98.2%
Vegetation	0	0	11	324	335	96.7%
Total	122	634	1170	330	2256	-
Accuracy	95.9%	97.9%	95.6%	98.2%	-	**96.25%**

Script 3 - Confusion matrix obtained with BAEI

	Fracking_areas	Non_fracking/bare	Roads	Vegetation	Total	Accuracy

		land				
Fracking_areas	124	0	16	0	140	88.6%
Non_fracking/bare land	0	615	27	0	642	95.8%
Roads	2	12	1121	4	1139	98.4%
Vegetation	0	10	27	307	344	89.2%
Total	126	637	1191	311	2565	-
Accuracy	98.4%	96.5%	94.1%	98.7%	-	**95.7%**

Script 3 - Confusion matrix obtained with NDTI

	Fracking_areas	Non_fracking/bare land	Roads	Vegetation	Total	Accuracy
Fracking_areas	124	3	13	0	140	88.6%
Non_fracking/bare land	0	611	29	2	642	95.2%
Roads	3	11	1123	2	1139	98.6%
Vegetation	0	12	25	307	344	89.2%
Total	127	637	1190	311	2265	-
Accuracy	97.6%	95.9%	94.4%	98.7%	-	**95.6%**

Script 3 - Confusion matrix obtained with DBSI

	Fracking_areas	Non_fracking/bare land	Roads	Vegetation	Total	Accuracy
Fracking_areas	105	0	35	0	140	75.0%
Non_fracking/bare	0	618	21	3	642	96.3%

land						
Roads	1	18	1118	2	1132	98.2%
Vegetation	0	23	8	313	344	91.0%
Total	106	659	1182	318	2256	-
Accuracy	99.1%	93.8%	94.6%	98.4%	-	**95.5%**

Annex 2. Confusion matrix of the training dataset - script 2.a - pixel based classification trial_zone_1

	Fracking_areas	Non_fracking/bare land	Roads	Vegetation	Total	Accuracy
Fracking_areas	663	0	2	0	665	99.7%
Non_fracking/bare land	0	168	2	0	170	98.8%
Roads	8	0	552	0	560	98.6%
Vegetation	0	0	0	1	1	100%
Total	671	168	556	1	1396	-
Accuracy	98.8%	100%	99.2%	100%	-	**99.5%**

Annex 3. Confusion matrix of the training dataset - script 2.a - pixel based classification + obia trial_zone_1

	Fracking_areas	Non_fracking/bare land	Roads	Vegetation	Total	Accuracy
Fracking_areas	687	0	3	0	690	99.6%
Non_fracking/bare land	0	185	1	0	186	97.3%

Roads	14	0	560	0	574	97.6%
Vegetation	0	0	0	1	1	1%
Total	701	185	564	1	1451	-
Accuracy	98.0%	100%	99.3%	100%	-	**98.7%**

Annex 4. Confusion matrix of the training dataset - script 2.b. - pixel based classification trial_zone_3

	Fracking_areas	Non_ fracking/ bare land	Roads	Vegetation	Total	Accuracy
Fracking_areas	666	0	1	0	667	99.8%
Non_ fracking/ bare land	0	464	2	0	466	99.6%
Roads	3	3	877	0	883	99.3%
Vegetation	0	0	0	0	0	-
Total	669	467	880	0	2016	-
Accuracy	99.6%	99.4%	99.7%	-	-	**99.6%**

Annex 5. Confusion matrix of the training dataset - script 2.b - pixel based classification + obia trial_zone_3

	Fracking_areas	Non_fracking/bare land	Roads	Vegetation	Total	Accuracy
Fracking_areas	680	0	0	0	680	100%
Non_fracking/bare land	0	480	0	0	480	100%
Roads	3	1	897	0	901	99.6%
Vegetation	0	0	0	0	0	-
Total	683	481	897	0	2061	-

Accuracy	99.6%	99.8%	100%	-	-	**99.8%**

Annex 6. Script 2.a/2. b. - Pixed-based results + OBIA - accuracy - trial_zone_1 and trial_zone_3

Validation	**trial_zone_1**	**trial_zone_1**	**trial_zone_3**	**trial_zone_3**
	Small_aoi (pixel-based)	**Small_aoi (pixel-based + OBIA)**	**Small_aoi (pixel-based)**	**Small_aoi (pixel-based + OBIA)**
Run 1	0.779	0.760	0.797	0.627
Run 2	0.811	0.751	0.822	0.588
Run 3	0.824	0.782	0.791	0.685
Run 4	0.813	0.775	0.805	0.652
Run 5	0.822	0.739	0.825	0.628
Run 6	0.825	0.720	0.805	0.651
Run 7	0.822	0.777	0.801	0.657
Run 8	0.789	0.732	0.805	0.664
Run 9	0.811	0.746	0.814	0.644
Run 10	0.792	0.736	0.789	0.627
Median	0.812	0.749	0.805	0.648