Student thesis series INES nr 545

Carta ex Machina

Testing object-based machine learning and unsupervised classification in land use change detection mapping in the semi-arid governorate of Sidi Bouzid, Tunisia

Kristian Emil Havnsgaard Paludan

2021

Department of
Physical Geography and Ecosystem Science
Lund University
Sölvegatan 12
S-223 62 Lund
Sweden

Kristian Emil Havnsgaard Paludan (2021).
Carta ex Machina – Testing object-based machine learning and unsupervised classification in land use change detection mapping in the semi-arid governorate of Sidi Bouzid, Tunisia

Carta ex Machina – En undersökning av objektbaserad maskininläring och icke-vägledd klassifikation för detektering och kartläggning av ändringar i markandvändingen i guvernementet Sidi Bouzid (Tunisien) karakteriserat av halvökenklimat

Bachelor degree thesis, 30 credits in *Physical Geography and Ecosystem sciences*
Department of Physical Geography and Ecosystem Science, Lund University

Level: Bachelor of Science (BSc)

Course duration: *March* 2021 until *June* 2021

Carta ex Machina

Testing object based machine learning and unsupervised classification in land use change detection mapping in the semi-arid governorate of Sidi Bouzid, Tunisia

---

Kristian Emil Havnsgaard Paludan

Bachelors thesis, 15 credits, in *Physical Geography and Ecosystem Sciences*

Supervisor:
Ulrik Mårtensson,
NATEKO, Lund University

Exam committee:
Marko Scholze, NATEKO, Lund University
Abdulghani Hasani, NATEKO, Lund University

Acknowledgements

There are so many people that I could mention in this section! If corona has taught us anything it is to value all the social connections in your life. Thank you. This year has been really tough on me, and there were times where I was not certain I could manage getting my bachelors degree done during these circumstances. where I felt disconnected, frustrated and lonely in my studying. Without all the wonderful people I've met in Lund I would not have completed this thesis.

Firstly I want to thank my supervisor Ulrik Mårtensson for suggesting the topic, and always being supportive and believing in me. Your passionate involvement in the topic has been a great source of motivation, in a time where motivation was sometimes really difficult to keep up.

A special thank you goes to my wonderful student corridor AF200 for being there always, be it for a casual chat, a much needed dinner party, as listeners to my complaining monologues or help with LaTeX, grammar or really everything. I honestly don't know how I would have gotten through the year without you.

I also need to mention my classmates Julia, Marieke and Serena. You girls have meant so much to me both academically as a study group and as friends throughout the entire bachelor and also during my thesis. Thank you for all the fikas and long days and evenings in the computer lab. I will never forget the sunset and the special feeling accomplishment after our thesis all nighter, Serena.

Lastly I would like to send a thank you to my mum. Tak for altid at hjælpe mig med få et overblik og en plan og noget at se frem imod, når motivationen ryger eller jeg føler alting bliver for uoverskueligt. Du er den bedste og sødeste i hele verden.

Ch. 15 Le géographe

- Je suis géographe, dit le vieux Monsieur.

- Qu'est-ce qu'un géographe ?

- C'est un savant qui connaît où se trouvent les mers, les fleuves, les villes, les montagnes et les déserts.

- Ça c'est bien intéressant, dit le petit prince. Ça c'est enfin un véritable métier ! Et il jeta un coup d'œil autour de lui sur la planète du géographe. Il n'avait jamais vu encore une planète aussi majestueuse.

- Elle est bien belle, votre planète. Est-ce qu'il y a des océans ?

- Je ne puis pas le savoir, dit le géographe.

- Ah! (Le petit prince était déçu.) Et des montagnes ?

- Je ne puis pas le savoir, dit le géographe.

- Et des villes et des fleuves et des déserts ?

- Je ne puis pas le savoir non plus, dit le géographe.

- Mais vous êtes géographe !

- C'est exact, dit le géographe, mais je ne suis pas explorateur. Je manque absolument d'explorateurs. Ce n'est pas le géographe qui va faire le compte des villes, des fleuves, des montagnes, des mers, des océans et des déserts. Le géographe est trop important pour flâner. Il ne quitte pas son bureau. Mais il y reçoit les explorateurs. Il les interroge, et il prend en note leurs souvenirs. Et si les souvenirs de l'un d'entre eux lui paraissent intéressants, le géographe fait faire une enquête sur la moralité de l'explorateur.

— Antoine de Saint-Exupéry, *Le Petit Prince*

Contents

Abstract

Sidi Bouzid, Tunisia is an inland governorate in Tunisia that has undergone a rapid agricultural and urban development since the Tunisian independence in 1952 from being a rural and largely nomadic region into a hub of irrigated agriculture. In 2010 Mohamed Bouazizi sparked the Tunisian revolution by lighting himself on fire int he city of Sidi Bouzid, with some blaming the inequality and water scarcity created by this rapid expansion in the irrigation farming as an important cause (Bayat, 2017; Malka, 2018). Sweden has aided in this development from 1972-1992, and as part of an evaluation project by Mårtensson et al (2019) a change detection analysis of the area around the Jabal al Kbar mountain was requested.

In this thesis two different object based classification methods (supervised random forest on a segmented image and unsupervised ISODATA on a segmented image) were tested to map and detect land use changes in a semi-arid environment in 4 areas around the Jabal al Kbar over the years 1972 to 2021 based on 30m and 60m LANDSAT imagery in 10-year intervals. Random Forest proved to have the highest accuracy of the two methods tested, with a total accuracy of 72% and Cohen's κ of 0.57 in a 60m resolution scene and 73% and κ of 0.54 for a 30m resolution scene. ISODATA was not deemed adequate as it produced too few classes with a standard number of iterations to be useful. A map produced with a low number of iterations had an accuracy of only 51% and κ of 0.13. The resolution was a limiting factor as it made identification of features difficult in the sampling process. Nevertheless, the historic development of the study area from the literature in terms of irrigation development and urbanization could be detected using Random Forest. Most notably a rapid development in irrigation farming from 2004 to today in the area Hichria was detected as well as decrease in the same period of irrigated area in the Sidi Bouzid depression. for all 4 areas, starting trends observed from 2004 to to 2011 continued into 2021, thus not showing any noticeable effect of the revolution in the development.

1 Introduction

1.1 Motivation

The governorate of Sidi Bouzid, is an important inland region in Tunisia at the crossroads between the lush Atlasian coastal slopes of the Maghreb and the dry Sahara desert. Today the region is known for being an important agricultural area for Tunisia. Historically, the region was home to nomadic peoples for millennia herding their grazing animals in this water scarce area and having only small scale farming (Abaab et al, 1992).
Since Tunisia's independence from France in 1952, the region has been rapidly urbanised though the region is still mostly rural (Abaab, 1997).
The region has since the 1990s been very intensely farmed with 93% of the total land area being used for some form of agriculture, grazing lands included (Abaab, 1997). Development of this governorate into a rich agricultural region has been named a national priority by the Tunisian Government since the independence in order to secure the coherence of the nation (Abaab et al, 1992). But as precipitation levels are generally very low, though they can range between 60 and 600mm a year (Abaab et al, 1992), this region requires irrigation in order to maintain a steady production for most crops.

Sweden and FAO has helped funding irrigation projects in Sidi Bouzid starting in 1973 and ending in 1992 (Mårtensson et al, 2019). These irrigation projects are reported to have brought improvement of living standards in the region, but has also led to an overexploitation of the limited groundwater resources in the region (Abaab et al, 1992; Malka, 2018; Mårtensson et al, 2019).
This problem is not limited to Sidi Bouzid, as the exploitation of groundwater has increased steadily in all parts of Tunisia (CNEA, 2007), however the exploitation of ground water in Sidi Bouzid is severe and was reported to be 150% In 2013 (Hamdi et al, 2015), which has caused the ground water table to drop, and increased the risk of salinisation of the irrigation and drinking waters. Hamdi et al (2015) furthermore found that the salinity of the ground water basin below the Sidi Bouzid depression had increased from 2 g/l in 1973 to 5 g/l. in 2013.

In 2010 Sidi Bouzid became the centre of the worlds attention as Mohamed Bouazizi lit himself on fire in front of the municipal office in the city (Britannica, T. and Editors of Encyclopaedia, 2021). Bayat (2017) performed a sociological field study in the area after the revolution and accuses the rapid up-scaling and mechanization of the agriculture in the area for having increased the inequality of income in the region and named it a possible driver for the revolution. Malka (2018) also mentions how water scarcity might have been an important contributing factor for the revolution.

Recently, Mårtensson et al (2019) has commenced an evaluation of the impact of the partly Swedish funded development projects, 25 years after the projects ended in 1992. As part of this evaluation it was requested that the spatial land use changes affected by the projects were detected using modern remote sensing methods, but also how the region has develop further since the project ended. Tracking the development is not only important for the evaluation, but can also be of good use for improving management practices in order to make the most of the scarce water resources (Diwediga, 2017).

Okin and Roberts (2004) has pointed out that vegetation mapping in arid and semi-arid regions can be particularly difficult for the following reasons: it is often difficult to get ground truth data, the vegetation in semi arid areas often experience immediate greening after rain events, and the high influence of the soil's spectral signature owing to the vegetation being sparser. In general most of the usual vegetation indices like normalized difference vegetation index (NDVI) tends to underestimate the vegetation in these areas and have difficulties differentiating between different types of vegetation (Okin and Roberts, 2004).

There have been done several land use mapping project in the region, most notably the LADA project by FAO (Biancalani et al, 2013) that covers all of Tunisia, and Diwediga (2017) who tracked the nationwide changes using MODIS data. Higher resolution mapping approaches have been tested in the neighbouring govenorate of Kairouan (Bousbih et al, 2019), where they achieved high accuracy of their maps with a decision tree approach that combined vegetation index data with data from a soil moisture model.

Geographic Object-based Image Analysis (GEOBIA or sometimes OBIA) is an image analysis that involves image segmentation into objects before further classification (Hay and Castilla, 2008). It seeks to mitigate the many problems that can be attributed with pixel based analyses, most notably the inherent neglect of features like texture and shape. The method was developed due to encountered problems with pixel based classification on high resolution data (Castilla and Hay, 2008), but the method has also been proven useful on medium and low resolution data (Blaschke, 2010). E.g. Katagis et al (2014) used an OBIA in a change detection analysis of Mediterranean forest fires using 60m and 30m resolution LANDSAT data with high accuracy. The approach is also tested in many different ecosystems from Chinese wetlands (Dronova et al, 2012) to Brazilian agricultural mapping (Schultz et al, 2015) with great reported accuracies.

1.2 Aim

This thesis aims to investigate the use of GEOBIA remote sensing techniques for making land use maps for a change detection analysis using LANDSAT data.

The goal is produce a time series with scenes roughly every 10 years in the period 1972-2021 to assist in the documentation of some of the land use changes reported by Mårtensson et al (2019) in this area, both during the Swedish development projects and after their termination.

In this thesis two GEOBIA remote sensing approaches are investigated: supervised random forest classification of a segmented image, and unsupervised iterative self-organizing data (ISODATA) clustering of a segmented image. The two algorithms will be tested on both a 60m resolution 4 band raster from 2012 and a newer 30m resolution 9 band raster from 2021 .

2 Method

2.1 Study area

2.1.1 Location

The study area for this thesis is the Sidi Bouzid depression and its immediate surroundings. The area is bounded the governorate border to the west until slightly west of the towns Bir El Huffay and Sidi Ali Ben Aoun where the area is cut in a straight line down to the southern border of the governorate. The study area is then bounded in the south and east by the anticline mountain range that runs from north of the town Majoura to just east of the town Faid. The P13 road marks the edge to the north. A screen shot from Google earth showing the study areas location within Tunisia can be found as figure 10 in Appendix D

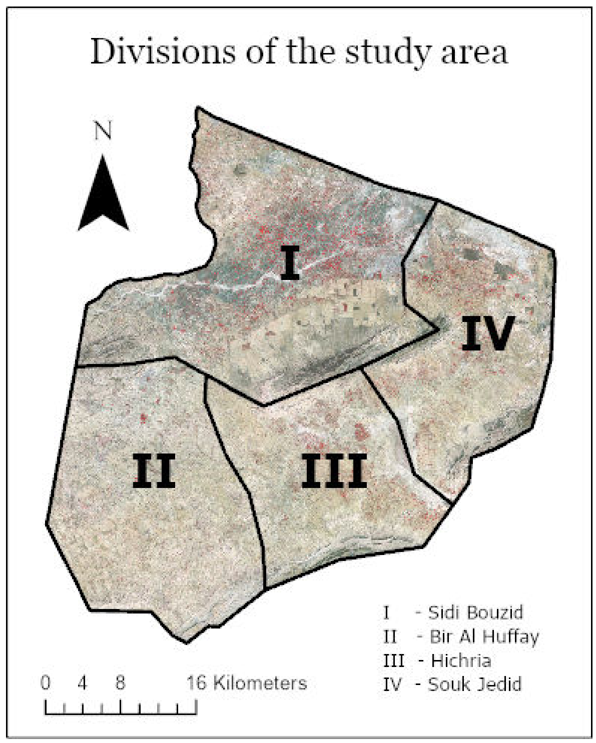


Figure 1: Division of the the study area

The area is further divided into 4 subareas numbered from I to IV and given each area a name corresponding to the largest town in area (figure 1). This was done in order to be able to refer to a specific area in text more easily, but also to provide zonal statistics to uncover local trends within the study area. The borders were chosen as primarily physical borders that divided the study area into 4 similar sized regions.

2.1.2 Climate

The governorate of Sidi Bouzid is a landlocked province in central Tunisia. The region experiences about between 150 to 250mm in a normal year, however some mountainous areas receive considerably more, up to 500-700mm pr year (Abaab et al, 1992). As can be seen on figure 9 in appendix C, The warmest month of the year is July, with an average high temperature of 37 degrees, while the coldest is January with an average high temperature of 15 degrees (Climate-data.org, 2021). On top of being the warmest period of the year, the summer months are also the driest, with July only experiencing less than 5mm of precipitation on average (figure 8 in appendix C). September is the wettest month of the year with an average of 32mm followed closely by March (Climate-data.org, 2021).

Precipitation levels vary a lot from year to year. According to Abaab et al (1992) the annual precipitation varies from 60 to 600mm. As can seen from figure 2 there are actually larger extremes than that in the data. In 1969 Sidi Bouzid experienced 1146 mm of rain the highest ever recorded, due to a very unusual amount of precipitation in the fall, which caused severe flood damages all over Tunisia that year (Poncet, 1970). From figure 2 we can also see that even though the precipitation is highly erratic in the period 1972-1991 it doesn't seem to have become generally neither drier nor wetter in this time period. Soumaia et al (2016) who monitored the lowering of the water table in the Sidi bouzid depression did not find a trend in annual precipitation either in their study period from 1990 to 2015.

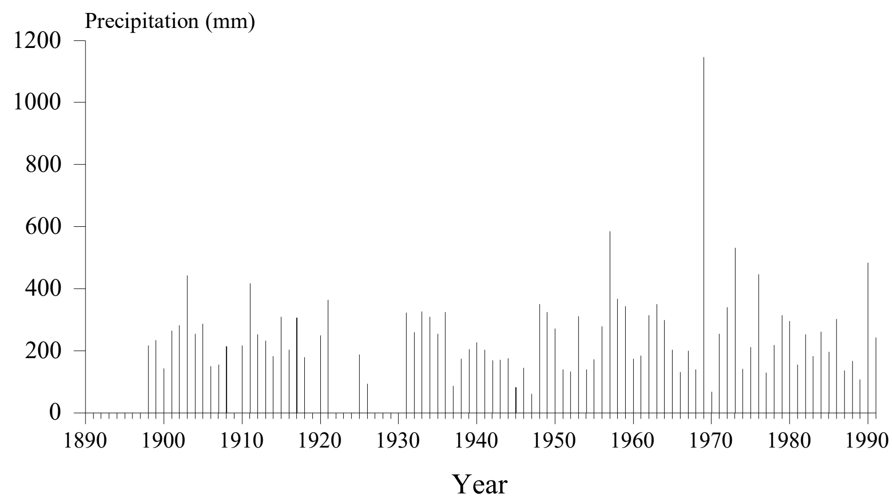


Figure 2: Annual precipitation in Sidi Bouzid from 1898 to 1991. Data from Institut National de la Meteorologie monthly reports 1984-1991

2.1.3 Physical geography

The entire governorate is characterised by relatively flat plains of filled synclines where the anticlines protrude in the form of the Dorsal, a continuation of the Atlas mountain range that runs through the entire country in a SW to NE direction (Abaab et al, 1992). The study area is composed of two such plains, the Sidi Bouzid depression (I in figure 1) and the Hicheria plaine (III and most of IV in figure 1) divided by the Jabal el Kbar mountain (Doukh et al, 2018). The Oued el Fekka runs from east to west in the north of the study area through the Sidi Bouzid depression, which is considered the most important agricultural zone in the governorate (Doukh et al, 2018). South of the mountain there is a Sebkha or Playa, that only rarely is filled with water and most of the time exists as a salt flat with little to no vegetation. The natural vegetation is mostly semi arid steppe with small shrubs and herbs like rosemary and artriplex (Mårtensson et al, 2019). However, the area is very intensely farmed so a lot of the natural vegetation is limited to the mountains. The area spans two groundwater basins divided by the mountain: The Sidi Bouzid basin to the north and the Braga basin south of Jabal al Kbar (Doukh et al, 2018).

2.1.4 Human geography

The population of the governorate of Sidi Bouzid is approx. 430.000 people and is divided into 12 delegations (Statistique Tunisie, 2016), of which the four delegations Sidi Bouzid Ouest, Sidi Bouzid Est, Bir El Hfay and Souk Jedid are located completely or partially within the study area The largest urban area in the area is the eponymous town of Sidi Bouzid, with a population just short of 50.000 inhabitants (Doukh et al, 2018), which is located in the north of the study area in the Sidi Bouzid depression. Sidi Bouzid is both the economic and political centre of the governorate (Doukh et al, 2018).

2.1.5 Agriculture

The study area is quite diverse in terms of crops. According to the inventory by Doukh et al (2018) the dominant types of crops in the Sidi Bouzid depression (I on figure 1) are cereal crops and vegetable crops. The rest of the study area is dominated by olive trees and the most dominant type of irrigation farming here are orchards. The most common types of orchards are almonds, peaches and apples (Doukh et al, 2018).

Mårtensson et al (2019) reported that the dominant vegetable crop in the Sidi bouzid depression used to be tomatoes in the 1980's and 1990's but have today been replaced with slightly less water intensive crops like pimiento, some of which might be covered in plastic greenhouses during the winter.

2.2 Presentation of data

2.2.1 LANDSAT

The LANDSAT-project has been collecting satellite regular imagery of the entire globe since the 1970's and is publicly available which makes it ideal for this thesis' study of long term changes. LANDSAT is the longest running sattelite image collection of earth that we have, starting already in 1972, where the Multispectral Scanner (MSS) was used for data capture, with a spatial resolution of 60m in four bands (Bandwidths in appendix A, table 4). MSS captured data in Sidi bouzid only until 1987 after which it was replaced by the Thematic Mapper (TM) which has a 30m resolution and 7 bands (Bandwidths in appendix A, table 5). During 2012 there was however a LANDSAT mission using MSS again, covering Sidi Bouzid. The scene from this mission will be used in the accuracy assessment.

The most present scene is from Landsat 8 whose scanner is called Operational Land Imager (OLI). OLI has a total of 9 band with 7 bands that are designed to be a continuation of TM as well as two new bands (NASA and USGS, 2010) (Bandwidths in appendix A, table 6).

2.3 Presentation of algorithms

2.3.1 Segment Mean Shift

Segment mean shift is a method of image segmentation that clusters pixels of a raster in to larger uniform object with an average value of the pixels in each object as the object values. Other statistics like size, rectangularity of the object, mean colour and standard deviation are then calculated for the objects and are used to determine whether two objects should merge or stay separate. The input to the algorithm determines how smooth the output should be, i.e. how different the objects have to be, compared to their surrounding objects and how small objects are allowed to be. The more precise mathematical details that goes into this algorithm is beyond the scope of this thesis to explain. The algorithm strives to make objects that are discrete, coherent and contrasting their surroundings. According to Castilla and Hay (2008) one of the strength of segmentation is that it clusters the cells into segments that represents actual features for the mapper which is supposed to ease machine training.

2.3.2 Random Forest

Random forest is a machine learning technique based on the Monte Carlo approach, were the machine constructs a great number of independent decision tree models based on the training input. This results in a normal distribution of outputs which gives a probability for an object to belong to a certain class after which an object is assigned to the most probable class. The way the machine creates its trees is by taking random samples or sections of the training input to use for the making each decision tree. Since each tree is very sensitive to the training input this will result in significantly different trees. This process is called bagging or bootstrap aggregation. The algorithm has two parameters: number of trees, which regulates how many random decision trees are used for the normal distribution and maximum tree depth, which regulates the maximum amount of decisions per tree.

2.3.3 ISODATA

ISODATA is an abbreviation for Iterative Self-Organizing Data Analysis Technique. The algorithm works by choosing random cluster centres. Cluster in this sense being equivalent to classes. Each cell is then classified to belong to the group where the distance in terms of segment statistics or pixel values are the closest to the centre of the cluster/class. Then the standard deviation and distance between cluster centres are calculated. If the standard deviation is greater than a set threshold, the clusters will split in two, and if the distance between two centres is smaller than a set threshold, then the clusters are merged. Furthermore if a cluster is too small it will be deleted and all objects within it are reclassified into the remaining clusters. The iterations are performed until a set number of iterations are performed or the average distance between cluster centres is below a certain value or the average change in cluster centres between two consecutive iterations is too low. Despite the term cluster which might confusingly make the user think of segmentation tool, the ISO cluster tool in ArcGIS is a classifying tool.

2.4 Literature study

The project started out with a thorough literature study, in order to get acquainted with area and to fully understand the premise of the project. This proved to be more time consuming than expected, as the literature concerning Tunisia, be it technical reports, journal articles and even encyclopedic texts are primarily written in French and or Arabic. This was a challenge not only due to a lack of proficiency of scientific terms in French, but also because it made the search itself more difficult and time consuming. Two especially important references for this thesis found in literature search were the two publications from the governorate titled "Atlas du gouvernorat de Sidi Bouzid" (Abaab et al, 1992; Doukh et al, 2018). The publications includes a collection of land use inventory maps, geological maps, thematic maps etc. as well as lexical style background information also on the history of the governorate. The two reports has been a key source of information for understanding the background and present state of the region. The first Atlas was published in 1992, coinciding with the end of Swedish projects and the other was published very recently in 2018. This also made them ideal sources for expected changes in my change detection.

In order to align my interpretation of the satellite images with actual features on the ground. I've gone through all of the ground images available on Google Earth in the entire governorate.

From the findings of Mårtensson et al (2019) it was decided to assume in this study that all tree crops with a spacing of over 24m between each tree is rainfed as that was the usual spacing between olive trees prior to the irrigation projects (Mårtensson et al, 2019). All cereal crops, vegetable crops and tightly packed tree crops are considered irrigated.

2.5 Selection of scenes

A goal for the thesis was to do a time series analysis of rasters from every 10 years starting in 1972 and ending with the most recent in 2021. some years however there were no suitable scenes to choose from ,as some scene would have e.g. band failures on one of the bands or that the study area was cloudy.

From initial testing it seemed like the segment mean shift worked better on some scenes than others. Specifically if the scene appeared relatively humid with a greener soil colour, the segmentation would not be able to create objects that matched the visible fields recognised in the scene. It was also harder to manually detect classes in these scenes. To assure the best possible conditions for the classification I therefore aimed to select the scenes that I suspected to be the driest. Summer is generally the driest season, but as mentioned in section 2.1.1 the precipitation in the region is highly erratic and has high inter-annual variability fig 2, which makes it almost impossible to predict the dryness from seasonality alone. It was therefore not ideal to select scenes purely from time of year, but rather multiple candidate scenes were inspected and judged based on their greenness and my ability to interpret the scene manually. This resulted in a wide range of months in table 1 of selected scenes, though only one scene, MSS 2012, was captured during winter. This scene was not a difficult scene to interpret, which was lucky as it was the only one to chose from.

2.6 Image processing

2.6.1 NDVI

In the image classification wizard in ArcGIS Pro, apart from the three visible bands, one also has the possibility of adding a additional raster to improve accuracy. Despite the findings by Okin and Roberts (2004) who reported that NDVI tends to underestimate vegetation in dry areas, I decided to use this index as my optional layer, as urban areas and humid zones, two classes that were difficult to separate by colour, had slightly different NDVI values. NDVI also fulfilled the criteria of being able to be calculated from MSS data as it only requires the red and NIR bands (eq.1). NDVI has also previously been used to map vegetation in the governorate in Doukh et al (2018), and as part of a decision tree based land use mapping in the neighbouring Kairouan governorate by (Bousbih et al, 2019) et al.

$$NDVI = \frac{NIR - Red}{NIR + Red} \tag{1}$$

2.6.2 Segmentation

Segmented Mean Shift was the algorithm used for segmentation. As band inputs were chosen to be red, green and NIR, which are bands 4-5-6 for the MSS 1 scene, 1-2-3 for

Scenes selected		
Mission number	Scanner	Date
1	MSS	August 11 1972
5	MSS	July 11 1984
6	TM	August 24 1994
6	TM	May 31 2004
7	ETM	September 08 2011
5	MSS	December 31 2012
8	OLI	11 April 2021

Table 1: Table of selected scenes for this thesis

MSS 5, 2-3-4 for TM, and 3-4-5 for OLI (see tables 4, 5 and 6 in appendix A).

According to Castilla and Hay (2008) a good segmentation is a segmentation with no cases of undersegmentation, where the segmentation creates image objects that the mapper interprets as two or more different geo objects, and few cases of oversegmentation, where the segmentation creates too many image objects compared to the identified geo objects. I therefore started with the highest values for spectral and spatial detail to assure not to have cases of undersegmentation. After the segmentation the layer was inspected to check if it succeeded in detecting the image objects I interpreted from the scene manually, if not the input were adjusted to a lower detail until the result was satisfying.

2.6.3 Random forest supervised classification

For the supervised random tree classification each object is classified using a statistical algorithm based on a supervised training through manual classification of sample objects in a training data file.

I used 13 different classes for sampling that I believed to be visually different from each other in the LANDSAT data, but not necessarily representing different land use /land cover types. The 13 different classes represented 5 parent classes: Olive trees, Irrigated agriculture, Urban areas, water and other, where other included the 5 subclasses Humid zones, Dense grazing area, Sparse vegetation, Hillshade and Wadi.

The parent classes were chosen so that I was confident enough that I could distinguish the classes with high precision and confidence in the accuracy assessment. The 13 subclasses with example pictures in TM and Google Earth can be seen as appendix B, table 7. Between 30 and 60 sample objects were selected for each class except for some of the early years when there were to few urban objects and water objects for selecting 30 samples. The objects where sampled based on an interpretation of the unsegmented image with the select object from segmented image tool. This tool allows the mapper to click on the unsegmented image and see the outline of the corresponding object in the segmented image and save it as a sample. This process also double checks the segmentation as the mapper will identify geo objects and click on it to select the sample and check that the image object outline indeed does match the identified geo object.

I chose 30 trees with a maximum tree depth of 10. These values were found using trial and inspection. I gradually increased the numbers in both inputs until the output didn't seem to change. As a post processing, polygons were drawn around identified mountains and the sebkha, and all irrigation class objects inside of the polygons were reclassified using a reclassifying tool to "other".

2.6.4 Unsupervised ISODATA

The other method I'm testing involves using ISODATA cluster unsupervised classification. This an unsupervised method meaning that it does not require training data. Instead the algorithm separates (according to some manually set variables/rules) into a number of classes that it finds significantly different from each other. I aimed to have at least 5 different classes. There is no suggested number of iterations for this algorithm, but according to the documentation for the algorithm provided by ArcGIS Pro, too few iterations will have unstable classes while too many iterations will result in too few and too broad classes (). The standard number of iterations in ArcGIS Pro is 10. I went with the highest number of iterations that produced at least 5 classes, which in the case of the OLI scene proved to be only 2 iterations.

Like with the random forest classification, the images were post processed so that any irrigation inside the previously mentioned mountain and sebkha polygons were corrected to the class representing other.

2.6.5 Accuracy assessment

For validation, I generated a stratified random set of 500 points based on my classification of the 2021 OLI and 2012 MSS scenes. I decided to do a stratified set of points in order to make sure that all classes were represented in the accuracy assessment, and 500 points to make sure that there were at least 30 samples within each sub-class. The points were then exported to Google Earth after which I interpreted each point manually. In this interpretation I only divided the points into my 5 parent classes.

This decision was made because it was too difficult to distinguish the sub-classes properly in Google Earth, as they were based on distinctiveness in LANDSAT that includes infrared bands and not in true colour images like Google Earth.

As mentioned in 2.2.1, Sidi Bouzid is one of the areas that were covered by the last LANDSAT 5 mission in 2012. In the interpretation of points for 2012 I used the "show historical imagery" feature in Google Earth as basis for my interpretation. 5 of the locations were however covered by clouds in the 2012 images on Google Earth. In these instances 2013 imagery were used for the for the interpretation.

After all the points were interpreted, the point layer was exported back into ArcGIS and a confusion matrix was generated with interpreted ground truth class on one axis and the mapped class value on the other axis, so that the diagonal contained the amount of correctly classified samples within each class.

For each class a user's and producer's accuracy was calculated (see table 2). User's accuracy measures the accuracy of the mapped class from the user's perspective. If the map says a location is "Class 1", then how likely is this ground truth to agree with this classification . Producer's accuracy is conversely the accuracy viewed from the mapper's perspective. It measures how many of the sample points within a class that were correctly classified.

The total accuracy for the matrix was also calculated. As seen in table 2 it is the total amount of agreements divided by the total amount of samples. Lastly Cohen's κ was also calculated. Cohen's κ is another metric for the accuracy, that corrects the total accuracy for the likelihood of getting an agreement by chance alone. If we consider table 2 and use its values, then:

$$P_1 = \frac{a+b}{a+b+c+d} \cdot \frac{a+c}{a+b+c+d} \tag{2}$$

Where P_1 is the probability of getting "Class 1" by chance. Similarly:

$$P_2 = \frac{c+d}{a+b+c+d} \cdot \frac{b+d}{a+b+c+d} \tag{3}$$

Where P_2 is the probability of getting "Class 2" by chance.

Accuracy assessment			
Ground Truth / Mapped	Class 1	Class 2	User's accuracy
Class 1	a	b	$\frac{a}{a+b}$
Class 2	c	d	$\frac{d}{c+d}$
Producer's accuracy	$\frac{a}{a+c}$	$\frac{d}{b+d}$	Total accuracy $= \frac{a+d}{a+b+c+d}$

Table 2: Example of a confusion matrix, green marks agreement, red marks disagreement

The propability of expected random agreement, P_e is then

$$P_e = P_1 + P_2 \tag{4}$$

Cohen's κ is then defined as

$$\kappa = \frac{\textit{Total accuracy} - P_e}{1 - P_e} \tag{5}$$

The Cohen's κ was the metric I used to judge the performance of each method and against each other. There are no set values for when a κ is sufficient, which is one of the common critiques of the statistic (Delgado and Tibau, 2019), however Henry et al (2016) in their mapping of sugar cane plantations using LANDSAT data interpreted a κ of 0.41-0.6 as "moderate" and everything less than 0.2 as "poor". As random forest had the highest κ of the two in both scenes, this was the classification I used in the change detection analysis.

2.6.6 Change detection

The time series analysis, was split it up into two groups sections before and after 1994 which is the scene that is the closest to both the end of the Swedish development projects and the governorate's inventory report (Abaab et al, 1992). Spatial patterns of the development of selected classes were interpreted manually. The areal sum of each parent class was furthermore calculated for each subdivision (see figure 1 and presented in bar graphs in order to quantify the changes and observe local trends.

3 Results

3.1 Supervised Random forest classification

The classification using Random Forest proved very successful overall with only few noticeable errors to be detected before the accuracy assessment. The most notable error was an overestimation in the detection of urban areas, which have a very similar spectral signature to some of the humid zones with sparse vegetation. The method proved surprisingly effective on MSS data despite the coarse resolution and classified the data very close to how I would have interpreted it manually.

In the confusion matrix it is apparent that some classes were easier to identify than others, with the smallest classes, having the lowest accuracy. In tabel 3 one can read that the total accuracy was 75% with a Cohen's κ of 0.58 for TM data and a little lower accuracy for MSS, whith total accuracy of 73% and a κ of 0.56. This can be interpreted as a moderate accuracy according to Henry et al (2016).

Accuracy assessment		
Method	Total accuracy	Cohen's κ
Random Forest MSS	72%	0.54
Random Forest OLI	73%	0.57
ISODATA MSS	51%	0.12
ISODATA OLI	48%	0.13

Table 3: Table of accuracy assessment statistics for each method and data type

3.2 Unsupervised ISODATA

The ISODATA cluster method separated the scenes into a very different amount of classes depending on the scene. It was not possible to obtain a map that resembled how I would have classified the scenes manually.

Contrary to what I anticipated, the ISODATA-clustring actually resulted in more classes in the MSS data at 2 iterations (14 classes) than from the OLI data (6 classes). After the clustering each cluster were then reclassified to the class that it was interpreted to represent the best.

2 iterations is a very low number of iterations to be the limit for achieving the set number of classes. This fact alone suggest that the method is not adequate in this setting, which also showed in the accuracy assessment. None of classifications performed very well, with both accuracy assessments resulting in a κ value below 0.2 which means that the method is poor according to Henry et al (2016) and only barely better than a random classification. Only one statistic stood out: The irrigated farms class had a 91% user's accuracy on the OLI scene. Based on the comparison of the κ-values I decided to use the supervised random forest classification only for my change detection.

3.3 Change detection

3.3.1 1972-1994

The change detection in this period showed the positive trend in irrigation farming also described by Abaab et al (1992) especially in the Sidi Bouzid depression (I) as seen in figures 4. In 1972 some areas close to Oued el Fekka are still classified as "other", separating the area into two irrigated zones as is highlighted on figure 5. Due to their location I suspect that these were grazing lands. 12 years later these areas seemed to have been converted to irrigated crop lands instead (figure 5). Abaab et al (1992) furthermore described a rapid urbanization of the area, another trend that can be seen clearly in all regions but Hichria (III) (figure 3). The city of Sidi Bouzid was the only urban area that could be detected in 1972 (figure 5). In 1994 the city had already reached almost the same extend as we see in the 2021 map on figure 6. All of the major towns in the area are visible from 1994. As seen on figure 5, the total area of irrigated farmland increased a little in Souk Jedid (IV).

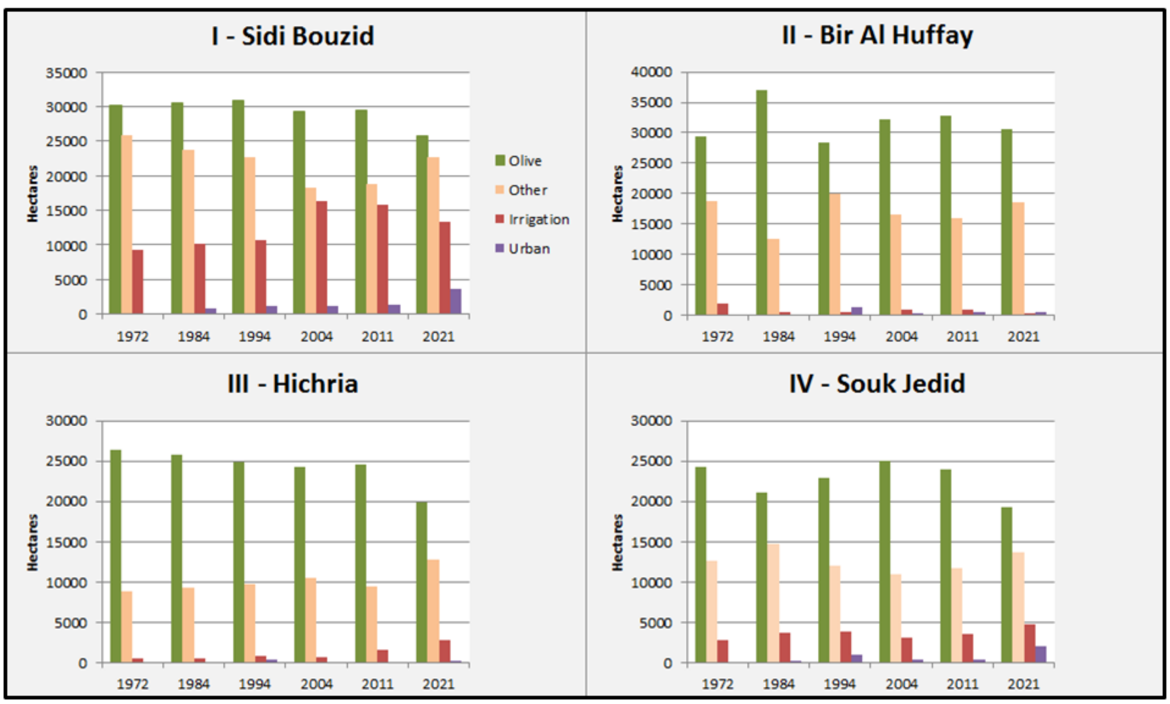


Figure 3: Total mapped area of parent classes within each sub region per year. Water surfaces included in "others" for this chart as it was too small to show

3.3.2 1994-2021

Since 1994 to 2004 a small irrigated cluster was detected south of Bir al Huffay towards Sidi Ali ben Aoun (figure 6). This new agricultural area doubled the total area of irrigated agriculture in Bir al Huffay (II) according to figure 4. This area gradually disappear from the map towards 2021 which is illustrated in both figure 6 and as a decline in total irrigated area for the region in figure 4.

From 2004 to 2021 a drastic increase in irrigation farming area in Hichria (III) was observed on the map (figure 6). In the bar chart on figure 4 One can even between 2004 to 2021 observe a seemingly exponential growth, where the total irrigated area have doubled between each map. From 1994 to 2004 in Sidi Bouzid (I) figure 4 shows a very large increase in irrigated farmlands of almost 6000ha. Like Bir al Huffay (II), Sidi Bouzid (I) has experienced a slight decrease in irrigated agricultural area since 2004.

Area of irrigated land

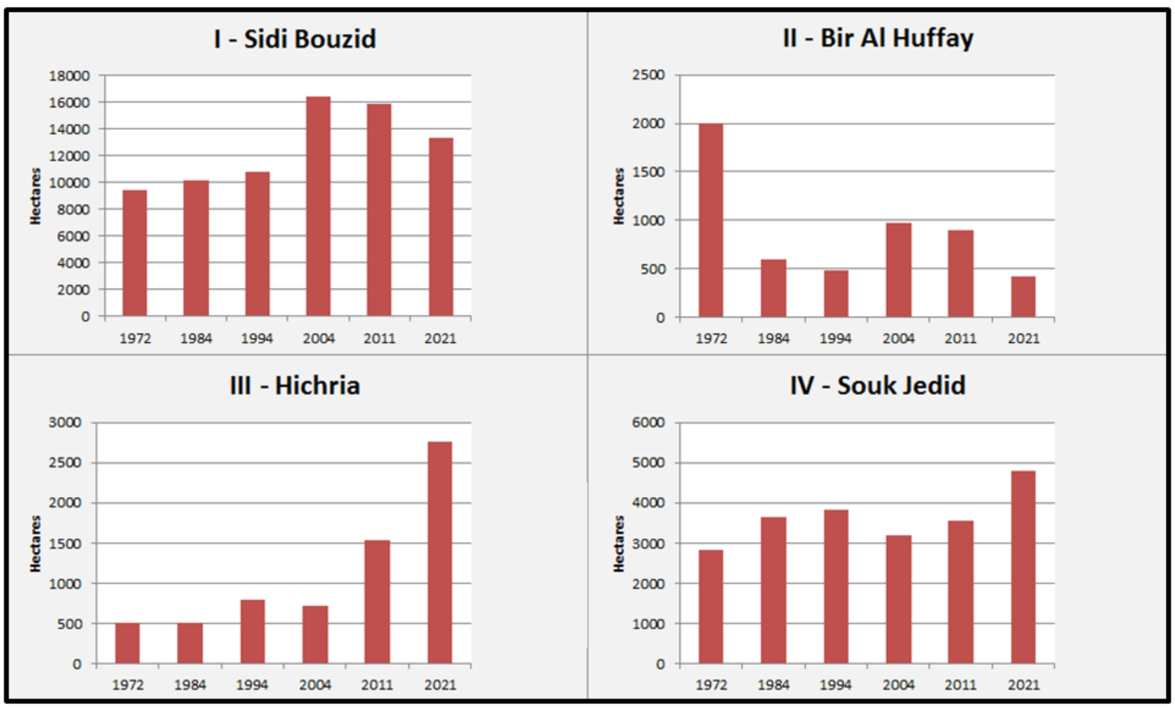


Figure 4: Total mapped area of irrigated land within each sub region per year

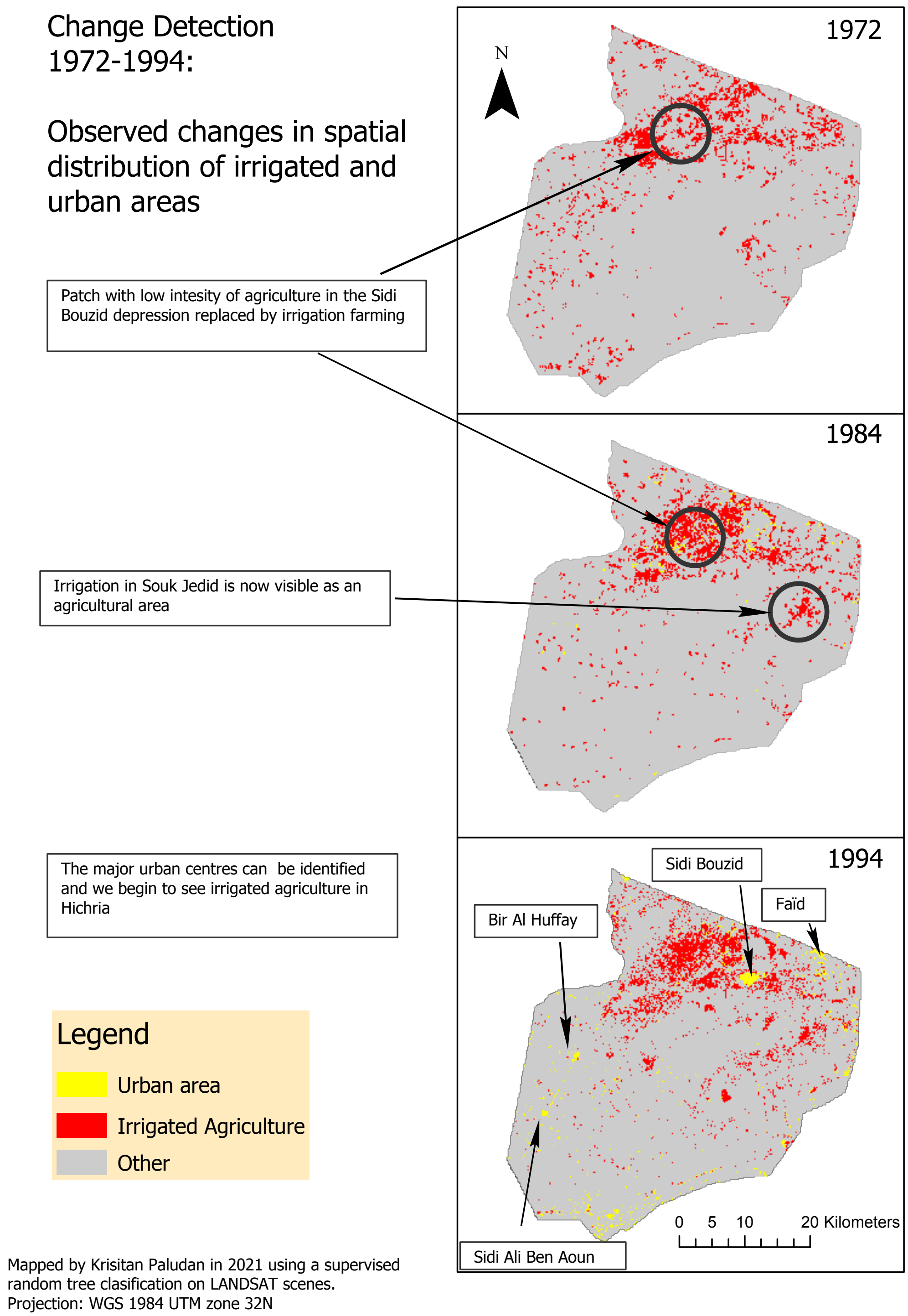


Figure 5: Change detection 1972-1994

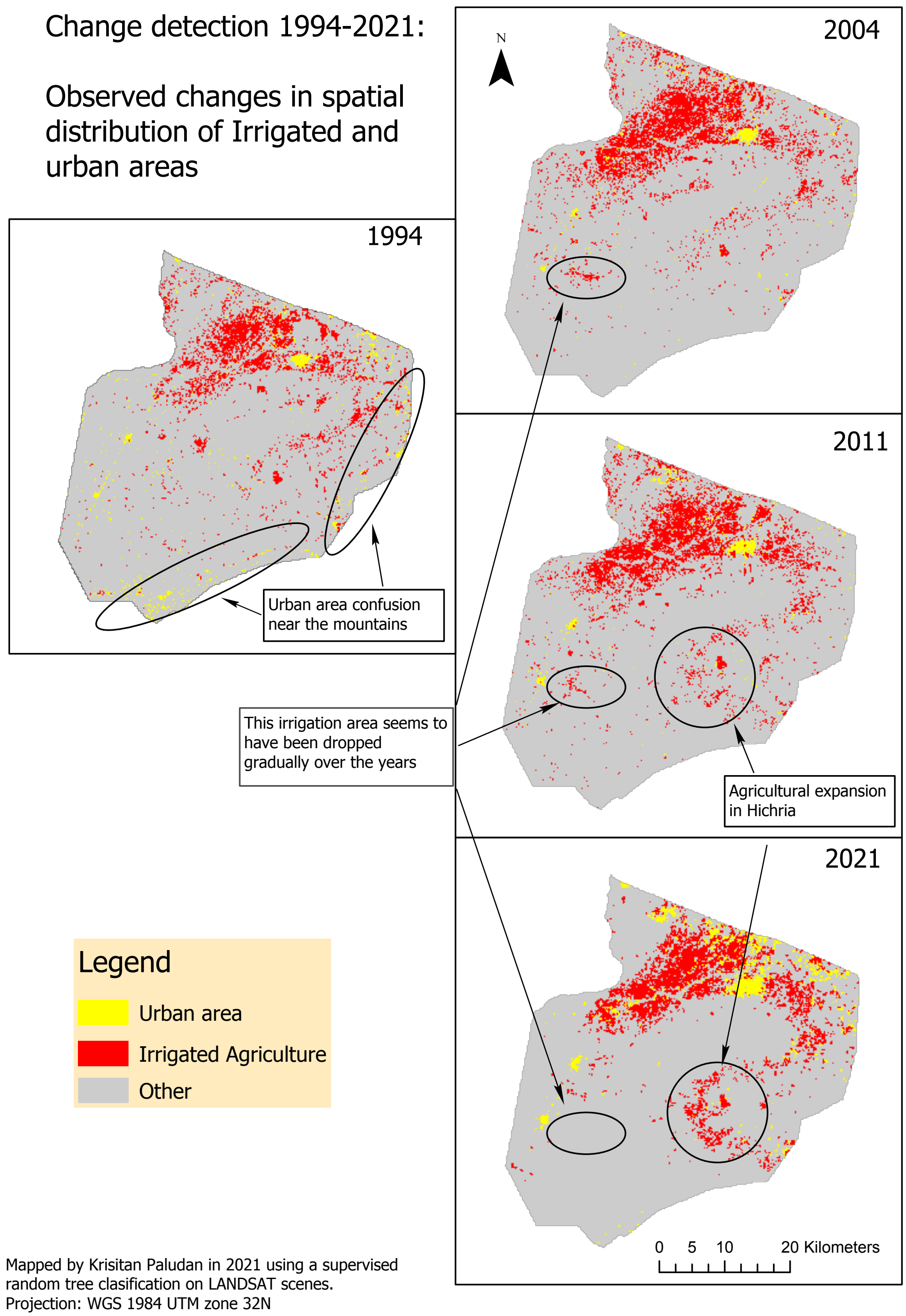


Figure 6: Change detection 1994-2021

4 Discussion

4.1 Discussion of Change detection

The detected decrease in irrigation farming in the Sidi Bouzid depression(I) since 2004 is perfectly inline with both Mårtensson et al (2019)'s reportings of water scarcity affecting especially the tomato production with a subsequent decrease in total vegetable production presented in Doukh et al (2018). A very high increase in the amount of irrigated land were furthermore detected in Hichria (III) which is also described by Mårtensson et al (2019). The irrigated farming in this area is according to the paper short sighted and unsustainable. The extend and growth rate of this trend was not however discussed in that report. Based on the findings in this report it seems that amount of irrigated land in this area was very low and stable until around 2004 after which the area of irrigated farming has roughly doubled every 10 years. In all 4 areas the trends observed from 2004 to 2011 have continued into 2021, and thus does not seem to have been affected by the revolution (figure 4)

The rapid urbanization from 1972-1994 described in the 1992 atlas (Abaab et al, 1992) around the town of Sidi Bouzid was also detected. From 1994-2021 there were some confusion in the maps between urban areas and other classes especially in mountainous areas (see figure 6). The creation of small settlements in Hichria (III) and Souk Jedid (IV) along with the irrigated agriculture expansion could also be detected (figure 6).

4.2 Difficulties in classification

4.2.1 Greenhouses

As mentioned in 2.1.5 Mårtensson et al (2019) reported that there has been a recent increase in uses of greenhouses, which is not accounted for in this study, as they could not be detect in the 30m resolution data. It is therefore likely that some irrigated areas may have been misclassified due to them being covered in plastic. The use of plastic greenhouses was however mentioned to mostly be a case in the winter months, which should make this an insignificant limitation as all but the 2012 MSS scene were captured in late spring, summer or early autumn (table 1).

4.2.2 Urban areas/humid zone confusion

An very difficult separation in the TM and OLI data proved to be urban areas, that had a similar blueish colour to some of the humid range land areas close to the wadis (Classes "Urban areas" and "Humid zones" in table 7, appendix) and in some of the mountainous areas. Especially in 1994 and 2021 there are overestimations of urban areas. In the MSS data however, only the largest towns were detected. I believe the reason for these misclassification to be oversegmentation. In order to avoid undersegmentation in the olive tree areas I had to accept oversegmentation of the urban areas, whose surfaces are much less homogenous than olive tree plantations. When a class is oversegmented it will lose the statistical benefits of GEOBIA, as it will no longer have it's characteristic shape, size and texture. In general Castilla et al (2014) has described how accuracy decreases with object size for object based mapping. An easy way to potentially solve this confusion problem in a future change detection study would be to let both urban areas and humid zones be the same class in the Random forest classification and then reclassify all humid zone cells, within the boundaries of modern day urban areas.

4.2.3 Mixed-use cells

Region II near Bir al Huffay was the hardest to classify. This area seems overall dryer, which can be supported by the lower irrigated agricultural area. The distances between the olive trees here are also generally larger than the rest of the study area, suggesting a higher water scarcity. In the driest parts there tend to be, contrary to the rectangular olive plantations found elsewhere, snaky and long plantation (example figure 7). These snaky plantations are presumably located in depressions with higher water availability. Other areas have small rectangular patches of olives in the middle of an otherwise bare land area. This structure is not as ideal to map for low resolution scenes in particular as it will lead to many cases of mixed use cells.

Another challenge for mapping the study area was the great variety of soil colours, which e.g. made for the necessity of creating 4 different olive plantation sub-classes to represent only 1 single type of land use: olive plantation. The soil colours seem in part to be correlated to moisture based on the difference in seasonally depending greenness observed when selecting the scenes (section 2.5)

Figure 7: An example of a tricky mixed use cell

4.3 The mappers influence on the results

Especially on the MSS data I was not able to identify the classes any better than the algorithm but someone more familiar with the area and landscape type with more experience in remote sensing would most probably get a better result than what is presented in this thesis.

The value of on site experience is difficult to measure, but Mogk and Goodwin (2012) stresses the importance of immersion in nature as an important step in acquiring knowledge. This is backed by findings of Fuller et al (2003) who polled university students in geography and environmental sciences during and after an epidemic where field trips were removed from the curriculum. The perceived learning outcome, motivation and interest in the subjects were found to be much lower even though there were no noteworthy change in grades.

As an example of how on-site experience can improve image interpretation, the "water reservoir" shown in appendix B table 7, is not at all a reservoir but in fact a large gulley. In all of the maps in this thesis however, this gulley has both been missampled in the machine learning process and misinterpred from google earth as being a water reservoir. A misinterpretation that could definitely have been avoided through better on site experience with the area.

Another important limitation for this study is its inability of replication due to the subjectivity involved in both the mapping itself and the accuracy assessment. Castilla and Hay (2008) names the mapper's bias an inherent limitation of GEOBIA approaches. The mapper decides the degree of segmentation based on one's own interpretation of the landscape into geo-objects. Geo-objects are however to a large extend fiat objects whose exact borders are vague and defined by the viewer and therefore exist only as a product of human cognition. This also mean that definitions are subjective and contextual for example in terms of scale and resolution (Castilla and Hay, 2008).

These problems also propagate down to the image validation, different interpreters will get different results, as humans interpret images by identifying geo objects whose exact borders are a social construct (Castilla and Hay, 2008).

4.4 Other limitations

4.4.1 Inter annual variation

Seemingly supervised random tree classification performed almost equally well independent on input resolution. This is somewhat in line with the findings of Katagis et al (2014) who also got better results with OLI scenes but also very good results with MSS, when they mapped forest fires in the Mediterranean with a GEOBIA approach. However the differences between the performances between the two resolutions was smaller than expected.

Furthermore the ISODATA cluster method performed better on MSS than on OLI, which was very unexpected. I suspect this to be attributed to the high inter annual variability of this region that might as discussed in section 2.5 make some years easier to classify than others independent of the resolution.

This is potentially an important limitation to the study as that makes the assumption that the accuracy of the 2012 MSS Random forest classification can reasonably be assumed for another MSS scenes less certain.

An investigation into the correlation between "segmentability" and monthly or daily precipitation would therefore naturally be another interesting topic for a future study. Precipitation data for the given months of the scene were not found in the literature study despite searched for as a priority, and therefore a such relationship could not be investigated much further.

4.4.2 Cohen's κ as a metric

Cohen's κ is the metric used in this study to assess the performance of the methods. κ is a widely used metric, that takes into account the probability of getting a result by pure chance. Delgado and Tibau (2019), however, have described the use of kappa as a performance metric as problematic, since it is a relative metric of agreement which makes it's absolute interpretation difficult. In other words it is difficult to assess from κ alone whether an accuracy is sufficient for a given use. Total accuracy is an intuitively easy metric to understand, but since it doesn't take random chance into account it is not an adequate metric on its own either. Delgado and Tibau (2019) suggests using other metrics with better absolute interpretation like Mathew's Correlation Coefficient. This metric is however much more complicated to calculate and is not, unlike Cohen's κ, a standard in the accuracy assessment tool of ArcGIS Pro.

4.5 Comments on ISODATA

Interestingly, despite the generally low κ values and overall accuracy, the ISODATA cluster performed better on MSS than on higher resolution scenes and resulted in more, and more nuanced classes. I don't know why this is, as the there seemed to be no noteworthy difference in the segmentation of these scenes.

In very specific circumstances, where only mapping of a very distinct feature like water or irrigated agricultural fields with crops (see table in appendix), this could be a preferred method as it is very fast, and doesn't require a skilled mapper to train. With more band combinations available in modern data, it is also easier to chose a band combination that highlights a specific land use type, which would allow for a multistage approach, where some classes produced by an ISO cluster are removed from the image after which the band combination is changed to highlight a different feature and so forth. It would be interesting to look into the possibilities of that in a future study on high spectral resolution data.

5 Conclusion

In conclusion, this type of environment was as tricky to map as the literature suggested. The combination of coarse resolution and the erratic nature of the precipitation makes it very difficult for to identify the land use especially in the LANDSAT MSS scenes as some important structural details were not visible in 60m resolution.

The GEOBIA approach proved very helpful as some classes were visually differentiated just as much based on texture and shape as on colour. The more human intuitive approach made it both easier and faster to train the algorithm for an unexperienced mapper. Moisture levels however proved to sometimes make segmentation difficult.

The difference between importantly distinct classes like irrigation farm, shrublands and tree crops on humid soil were however too subtle for unsupervised classification using ISODATA-clustering to be recommended using only 4 bands even when performed object based. A recommendation for a future study, however, would be to investigate multistage classification on TM or OLI rasters as the method was very effective in detecting certain very distinct classes. Despite the faced difficulties it was possible to produce maps with a 73% total accuracy or higher even on MSS data. The change detection anlysis added to confirmation of reported regional trends in the study area most notably the extreme growth in use of irrigation in Hichria and the recent years decline in the vegetable production in the Sidi Bouzid depression, which were both described by Mårtensson et al (2019). Based on the experiences acquired in this study I would suggest future mapping of the Sidi Bouzid depression and surroundings to include some estimation of soil moisture and perhaps a map of soil types that may reduce the amount of classes in the training of the machine.

References

Abaab, A. 1997. Agricultures familiales et politiques agricoles en Méditerranée : enjeux et perspectives. In: . (eds) Abaab, A., P. Campagne, M. Elloumi, A. Fragata, and L. Zagdouni. L'agriculture familiale en tunisie centrale face aux nouveaux défis écologiques et économiques. vol 12. Montpellier: CIHEAM. pp 7–27. *In French*

Abaab, A., S. Gharbi, A. Bousnina, A. Hayder, and R. Lamine. 1992. *Atlas du Gouvernorat de Sidi Bouzid.* Ministere de l'Environnement et de l'Amenagement du Territoire de la Republique Tunisienne. Tunis, Tunisia. *In French*

Bayat, A. 2017. *Revolution without revolutionaries : making sense of the Arab Spring.* Stanford studies in Middle Eastern and Islamic societies and cultures: Stanford University Press

Biancalani, R., F. Nachtergaele, M. Petri, and S. Bunning. 2013. *Land Degradation Assessment In Drylands*. FAO. Rome, Italy.

Blaschke, T. 2010. Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing* 651: 2–16

Bousbih, S., M. Zribi, M. El Hajj, N. Baghdadi, Z. L. Chabaane, P. Fanise, and G. Boulet. 2019. Sentinel-1 and sentinel-2 data for soil moisture and irrigation mapping over semi-arid region. In: editor, T. (ed) IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium Geoscience and Remote Sensing Symposium. . IEEE. Yokohama, Japan: IEEE. p 213

Britannica, T. and Editors of Encyclopaedia 2021. Jasmine Revolution. Encyclopedia Britannica. https://www.britannica.com/event/Jasmine-Revolution. accessed: 2021-26-05

Castilla, G., and G. J. Hay. 2008. . *Image objects and geographic objects*. Berlin, Heidelberg: Springer Berlin Heidelberg. pp 91–110. https://doi.org/10.1007/978-3-540-77058-9_5

Castilla, G., A. Hernando, C. Zhang, and G. J. McDermid. 2014. The impact of object size on the thematic accuracy of landcover maps. *International Journal of Remote Sensing* 353: 1029–1037

Climate-data.org 2021. SIDI BOUZID WEATHER BY MONTH // WEATHER AVERAGES. https://en.climate-data.org/africa/tunisia/sidi-bouzid/sidi-bouzid-47420/. accessed: 2021-20-06

CNEA 2007. *Elaboration d'une étude sur l'état de désertification pour une gestion durable des RN*. Centre National des études agricoles. Tunis, Tunisia.. *In French*

Delgado, R., and X.-A. Tibau. 2019. Why cohen's kappa should be avoided as performance measure in classification. *PLOS ONE* 149: 1–26. https://doi.org/10.1371/journal.pone.0222916

Diwediga, B. 2017. Land cover changes in Tunisia using MODerate Resolution Imaging Spectroradiometer (MODIS) MCD12Q1 yearly products. Tech. rep.. CGIAR. Montpellier, France.

Doukh, M. K., M. D. Ayari, M. Missaoui, K. Zerai, C. Hrizi, and R. Mahfoudhi. 2018. *Atlas du Gouvernorat de Sidi Bouzid*. Ministere de l'Equipement, l'Habitat et l'Amenagement du Territoire du Territoire de la Republique Tunisienne and TOPGEO. Tunis, Tunisia. *In French*

Dronova, I., P. Gong, N. E. Clinton, L. Wang, W. Fu, S. Qi, and Y. Liu. 2012. Landscape analysis of wetland plant functional types: The effects of image segmentation scale, vegetation classes and classification methods. *Remote Sensing of Environment* 127: 357–369. https://www.sciencedirect.com/science/article/pii/S0034425712003781

Fuller, I., S. Gaskin, and I. Scott. 2003. Student perceptions of geography and environmental science fieldwork in the light of restricted access to the field, caused by foot and mouth disease in the uk in 2001. *Journal of Geography in Higher Education* 271: 79–102. https://doi.org/10.1080/0309826032000062487

Hamdi, M., M. Soumaia, D. Lofti, M. Rajouene, and A. Habib. 2015. Effet de l'épandage des eaux de crues sur les ressources en eaux souterraines dans les zones arides: Plaine de sidi bouzid (tunisie centrale). *European Journal of Scientific Research* 1125: 35–36. *In French*

Hay, G. J., and G. Castilla. 2008. . *Geographic Object-Based Image Analysis (GEOBIA): A new name for a new discipline*. Berlin, Heidelberg: Springer Berlin Heidelberg. pp 75–89. https://doi.org/10.1007/978-3-540-77058-9$_4$

Henry, F., D. Herwindiati, S. Mulyono, and J. Hendryli. 2016. Sugarcane land classification with satellite imagery using logistic regression model

Katagis, T., I. Gitas, and G. Mitri. 2014. An object-based approach for fire history reconstruction by using three generations of landsat sensors. *Remote Sensing* 6: 5480–5496

Malka, H. 2018. *Water Pressure - Water, Protest, and State Legitimacy in the Maghreb*. Analysis Paper

Mogk, D. W., and C. Goodwin. 2012. *Learning in the field: Synthesis of research on thinking and learning in the geosciences.* . vol 486. Boulder, Colorado: Geological Society of America Special Papers. p 131–163

Mårtensson, U., A. Rabo, and L. Gammoudi. 2019. Tid och utvickling i Sidi Bouzid, Tunisien. supported by the research funder FORMAS. *In Swedish*

NASA and USGS 2010. Landsat Data Continuity Mission. https://landsat.gsfc.nasa.gov/sites/landsat/files/2012/12/20101119_LDCMbrochure.pdf. Accessed: 2021-26-05

Okin, G., and D. Roberts. 2004. Remote sensing in arid regions: Challenges and opportunities. In: Manual of Remote Sensing. . vol 4. New York City: John Wiley and Sons, Inc. p 111–146

Poncet, J. 1970. La 'catastrophe' climatique de l'automne 1969 en tunisie. *Annales De Géographie* 79435: 581–595. in French

Schultz, B., M. Immitzer, A. R. Formaggio, I. D. A. Sanches, A. J. B. Luiz, and C. Atzberger. 2015. Self-guided segmentation and classification of multi-temporal landsat 8 images for crop type mapping in southeastern brazil. *Remote Sensing* 711: 14,482–14,508. https://www.mdpi.com/2072-4292/7/11/14482

Soumaia, M., F. Lokmen, H. Monji, and R. MAJDOUB. 2016. Etude de l'efficience de l'épandage des eaux de crue sur la fertilité du sol et la recharge de la nappe (sidi bouzid, tunisie)

Statistique Tunisie 2016. Sidi bouzid a travers le recensement général de la population et de l'habitat 2014. Tunis. *In French and Arabic*

A LANDSAT band tables

Landsat MSS		
MSS 1-3	MSS 4 & 5	Colour and wavelength (μm)
Band 4	Band 1	Green 0.5-0.6
Band 5	Band 2	Red 0.6-0.7
Band 6	Band 3	NIR 0.7-0.8
Band 7	Band 4	NIR 0.8-1.1

Table 4: Table of available bands in Landsat MSS scenes, data from the United States geological survey (USGS)

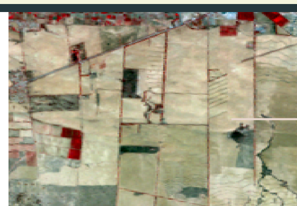
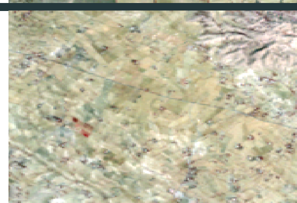
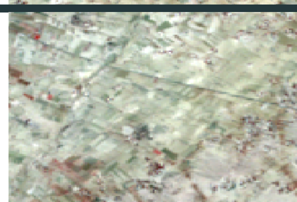
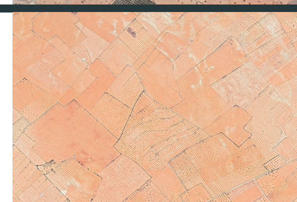
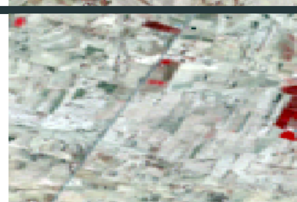
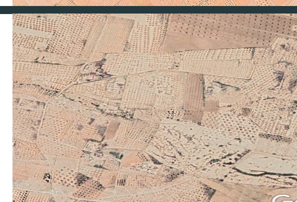
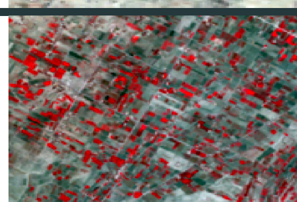
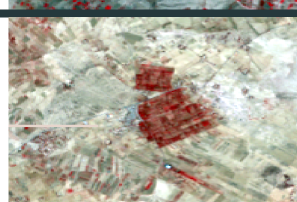
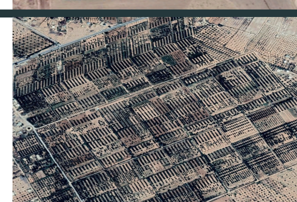
Landsat TM		
TM	Colour	Wavelength (μm)
Band 1	Blue	0.45-0.52
Band 2	Green	0.52-0.60
Band 3	Red	0.63-0.69
Band 4	NIR	0.76-0.90
Band 5	SWIR	1.55-1.75
Band 6	TIRS	10.40-12.50
Band 7	MIR	2.08-2.35

Table 5: Table of available bands in Landsat TM scenes, data from USGS

Landsat OLI		
Band	Colour	Wavelength (μm)
Band 1	Coastal Aerosol	0.43-0.45
Band 2	Blue	0.45-0.51
Band 3	Green	0.53-0.59
Band 4	Red	0.64-0.67
Band 5	NIR	0.85-0.88
Band 6	SWIR	1.57-1.65
Band 7	SWIR	2.11-2.29
Band 8	Panchromatic	0.5-0.68
Band 9	Cirrus	1.36-1.38

Table 6: Table of available bands in Landsat TM scenes, data from USGS

B Classification scheme with picture examples

Classification scheme and picture examples		
Class	Example in Landsat TM	Example in Google Earth
Colonial style olive plantation		
Small scale olive		
Small scale olive on red soil		
Small scale olive on pale soils		
Irrigated field in rotation		
Irrigation farming		

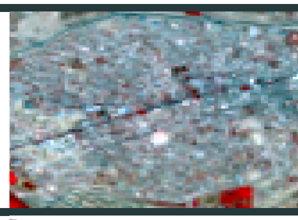
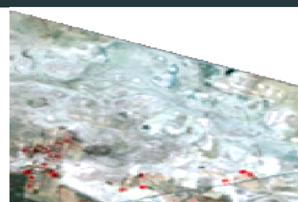
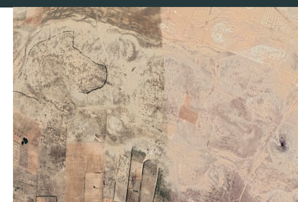
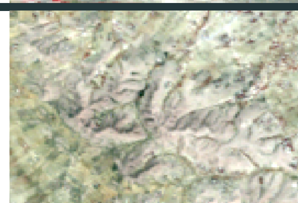
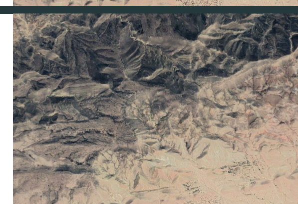
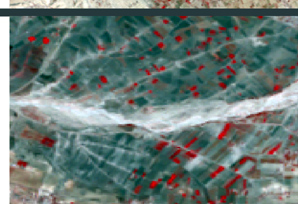
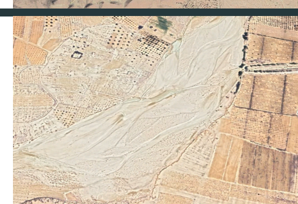
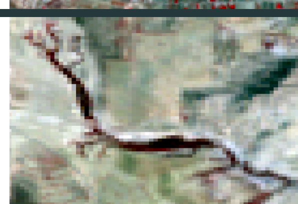
Urban area		
Humid zone		
Dense grazing area		
Sparse vegetation		
Hillshade		
Wadi		
Water reservoir		

Table 7: Classification scheme

C Climate data for Sidi Bouzid

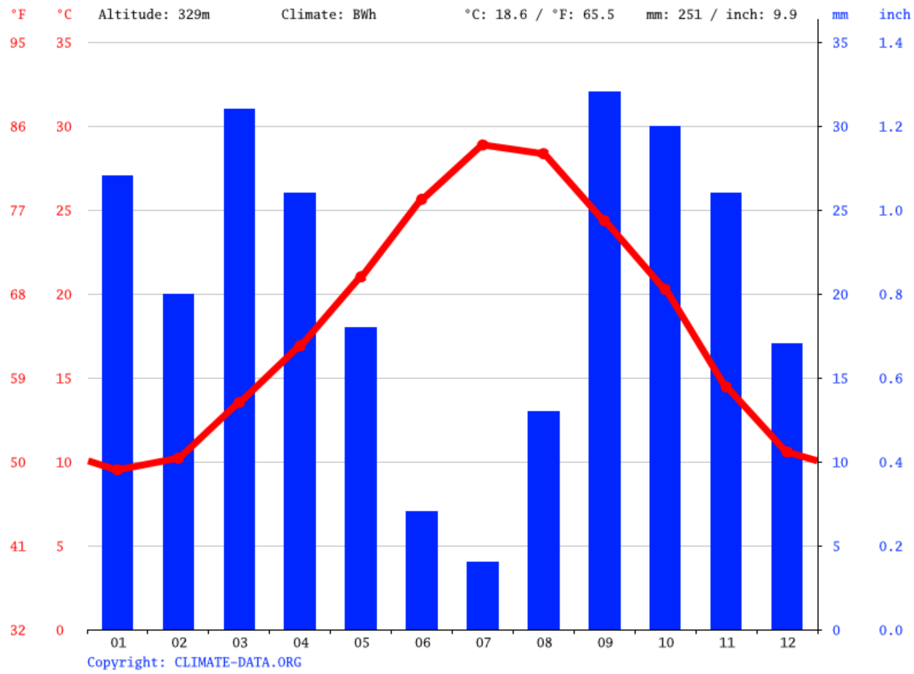


Figure 8: Monthly mean precipitation and temperature in Sidi Bouzid from: https://en.climate-data.org/africa/tunisia/sidi-bouzid/sidi-bouzid-47420/

	January	February	March	April	May	June	July	August	Sep-tember	October	November	December
Avg. Temperature °C (°F)	9.5 °C (49.1) °F	10.2 °C (50.4) °F	13.5 °C (56.3) °F	16.9 °C (62.4) °F	21 °C (69.8) °F	25.6 °C (78.1) °F	28.9 °C (84) °F	28.3 °C (83) °F	24.3 °C (75.8) °F	20.3 °C (68.5) °F	14.5 °C (58) °F	10.6 °C (51) °F
Min. Temperature °C (°F)	4.8 °C (40.7) °F	5.1 °C (41.1) °F	7.5 °C (45.5) °F	10.5 °C (50.8) °F	14 °C (57.2) °F	18 °C (64.3) °F	21 °C (69.7) °F	21.1 °C (70) °F	18.6 °C (65.5) °F	15.1 °C (59.1) °F	9.9 °C (49.8) °F	6.2 °C (43.2) °F
Max. Temperature °C (°F)	15.2 °C (59.4) °F	16.1 °C (61) °F	20.1 °C (68.1) °F	23.8 °C (74.9) °F	28.2 °C (82.8) °F	33.3 °C (91.9) °F	36.7 °C (98.1) °F	35.8 °C (96.4) °F	30.8 °C (87.5) °F	26.2 °C (79.2) °F	19.9 °C (67.8) °F	15.9 °C (60.7) °F
Precipitation / Rainfall mm (in)	27 (1.1)	20 (0.8)	31 (1.2)	26 (1)	18 (0.7)	7 (0.3)	4 (0.2)	13 (0.5)	32 (1.3)	30 (1.2)	26 (1)	17 (0.7)
Humidity(%)	60%	55%	51%	48%	44%	37%	35%	40%	52%	57%	60%	62%
Rainy days (d)	3	3	3	3	3	1	1	3	4	4	3	2
avg. Sun hours (hours)	7.6	8.4	9.4	10.6	11.9	12.9	12.8	11.9	10.3	8.9	8.1	7.5

Figure 9: Climate table of Sidi Bouzid from: https://en.climate-data.org/africa/tunisia/sidi-bouzid/sidi-bouzid-47420/

D Location of study area within Tunisia

Figure 10: Screen shot from Google Earth showing the location of the study area within Tunisia