MAPPING HALOPHYTES USING DECISION TREE APPROACH

R.A.Bouchhima[1], M.Ksibi[2]

[1] University of Sfax, Laboratory of Environmental Engineering and Eco-Technology, National School of Engineers (ENIS), Sfax, Tunisia, rim_attya@yahoo.fr
[2] University of Sfax, Laboratory of Environmental Engineering and Eco-Technology, National School of Engineers (ENIS), Sfax, Tunisia, mdh.ksibi@gmail.com

ABSTRACT

Ghannouch in the south-east of Tunisia is colonized by halophytes. Mapping and monitoring with remote sensing approach are previsioned as the ways to trace the spatial and temporal confines of their distribution in order to reveal the extent of salinization and its dynamic. The distinguishing of halophyte vegetation can be done by examining optical remote detecting information as a tree approach applied to European Space Agency Sentinel-2 symbolism. As a result, for an area of interest of 50 × 50 km², at least 68% was classified as halophyte land cover. This mapping exercise represents an important step toward improved halophyte mapping in Tunisia and could be used to monitor the status of other salinity-prone regions in the world.

Keywords: Salinity indices, Halophyte cover, Mapping, Sentinel-2

1. INTRODUCTION

The primary target of this examination is to decide our findings for accurate halophyte cover mapping by using a decision tree algorithm. In order to descry and collude the halophyte land cover in the area of interest (AOI), we compare a selection of 13 Salinity Index (SI), with the end of testing and vindicating which one restitutes the stylish fit in respect to direct saltness compliances on the ground. Climate change causes balance in the ecosystem, drought due to the rise of salts to the surface, and soil degradation due to intense summer evaporation.

In Tunisia, a semi-arid region, salt-affected soils presented 3% of the surface in the south of Tunisia (Gabes Oasis) extended over about 900 hectares. More than 100 ha are affected by salinization because of the high salt content of irrigation water. This oasis constitutes the proper habitat for halophytic vegetation which presents an optimal development in salty or very salty environments. Other forms affected by salinity are Sebkhas, a salt flats area. These ecosystems are generally hypersaline environments and crowded by halophytes. The study of their distribution is necessary to demonstrate the extent of salinization and its dynamics.

2. MATERIAL AND METHODS

The identification and the distribution of halophytes in Tunisia are not fully understood and not much work is reported in the literature. Thus, the limited accessibility and vastness of the south of Tunisia allowed the study of halophytes by mapping and monitoring reliably with a remote sensing approach. The dataset utilized in this study comprises two Sentinel-2 scenes (Level 1C) and 500 focuses on the ground obtained in the AOI, which were utilized as reference data to approve optical outcomes.

100 ground focus were considered for each of the five cover classes considered. The guide of land cover and vegetation type was gathered pixel by pixel utilizing Google Earth. This was created in 2017 and incorporates five classes, which are from now on marked in the text, tables, and figures as W, UBS, V, DV, and H and represent, respectively, water, urban and bare soil, vegetation, dense vegetation, and halophyte vegetation.

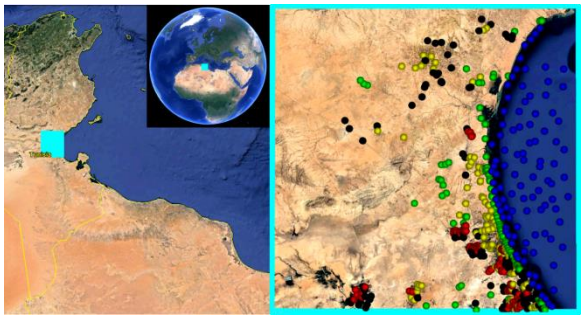

Figure 1: Google Earth pictorial view. a) The cyan box show the selected AOI in Ghannouch, Tunisia, centered at 34° 05’ 44”N, 9° 54’ 16”E. b) Reference pixel land cover map; land cover classes are as follows: ● water; ● urban/bare soil; ● dense vegetation; ● vegetation; ● halophytes

The optical information gives symbolism in the apparent groups, which encase red (R), green (G), and blue (B), in the close infrared (NIR) groups, in the short wave infrared groups (SWIR), and in the

warm groups (TIR), with a spatial goal equivalent to 10, 20, and 60 m for the noticeable and warm groups, separately [1]. The optical pictures we used are displayed in the noticeable band as misleading shading pictures.

Therefore, by using NDWI (Normalized Difference Water Index), NDVI (Normalized Difference Vegetation Index), and other remote sensing indices, such as the SIs, it is possible to map the halophyte vegetation [2,3]. With the purpose of classifying and comparing the land cover types for the two chosen images, we processed Sentinel-2 data by deriving NDWI, NDVI, and different SIs.

Finally, finding halophytic vegetation requires the presence of salt affected soils. Thus, various SIs were used to recognize and to discriminate the salt affected soils. Such indices are especially sensitive to sodium chloride in the VIS–SWIR (400–2500 nm) region [4] and are commonly calculated from the spectral reflectance of two or more bands. In this study, we tested the sensitivity of 13 SIs applied to the Sentinel-2 reflectance image.

According to [5], the difference image $|\Delta NDVI| = |NDVI08 - NDVI03|$ was then reclassified using a threshold value calculated as σ, where $|\Delta NDVI|$ represents the NDVI difference in absolute value between the two acquisition dates; σ is the standard deviation evaluated on the pixels where we are sure of vegetation presence ($0.075 < NDVIxx < 0.3$; the subscript xx stands for 03 or 08 in reference to acquisition sentinel-2 (date March 22, 2017, and August 29, 2017, respectively).

In our case, the threshold identifies two ranges in the normal distribution: (a) the tail region ($|\Delta NDVI| > \mu + n \cdot \sigma$) and (b) the central region of the normal distribution ($|\Delta NDVI| < \mu + n \cdot \sigma$). Pixels within the tail region of the distribution are characterized by significant vegetation changes, while pixels in the central region represent no change. The n factor defines the range of dispersion around the mean. This study considered the positive variation as the area of probable vegetation, when $|\Delta NDVI| > \mu + n\sigma$, in explicit form $|\Delta NDVI| > 1.96\sigma$.

3. RESULTS

The decision tree algorithm based on optical indices and the chosen threshold criteria are summarized in the flow chart in Figure 2. Each criteria consist of a conditional statement. Both criteria and threshold are described in the order they enter in the decision tree. Every time a condition is fulfilled, the resulting class/binary map is excluded from the area of a potential halophyte, resulting in a final map of suitable areas for halophyte cover. In the final decision tree, we included four variables, which we consider as the most efficient in excluding areas not suitable for the halophyte cover.

Here, the steps we followed along the decision tree are scheduled:

- Water and urban/bare soil were delineated using the NDWIxx algorithm, with a threshold of 0 (water or urban/bare soil);
- Vegetation and urban/bare soil were mapped using NDVIxx and $|\Delta NDVI|$ with a threshold of 0.075 and 1.96σ, respectively;
- Dense vegetation and urban/bare soil were delineated using NDVIxx with a threshold of 0.3 and 1.96σ, respectively;
- Vegetation on saline soil or vegetation on not saline soil was delineated using NDVIxx with a threshold of 0.075 ÷ 0.3 and SIs with a threshold SImin.

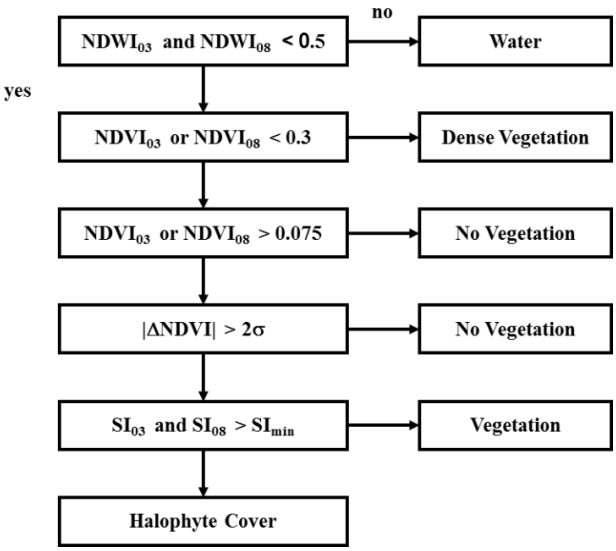

Figure 2: The decision tree based on indices criteria and thresholds

To assess the performance of the estimations, the recaptured classes were compared with the reference land cover chart. A confusion matrix was constructed in order to prize four well- accepted parameters for the evaluation of the bracket procedures(1) overall accuracy (OA),(2) user accuracy (UA),(3) producer accuracy (PA), and(4) kappa(K) measure The OA, PA, and UA are usually expressed as a percent [6]. OA is calculated by summing the number of correctly classified values and dividing by the total number of values. PA is the probability that a value predicted to be in a certain class really is in that class. UA is the probability that a value in a given class was classified correctly. K coefficient was generated from a statistical test to evaluate the accuracy of classification. K measures the agreement between classification and ground truth

pixels. A K value of 1 represents a perfect agreement, while a value of 0 represents no agreement at all. Moreover, the land cover area percentage for each class in AOI was evaluated. The estimated features together with three main classification outputs resulting by applying the decision tree methodology

The comparison between NDVI03 and NDVI08 clearly demonstrates an increased vegetation activity in spring (Marsh) as compared to summer (August).

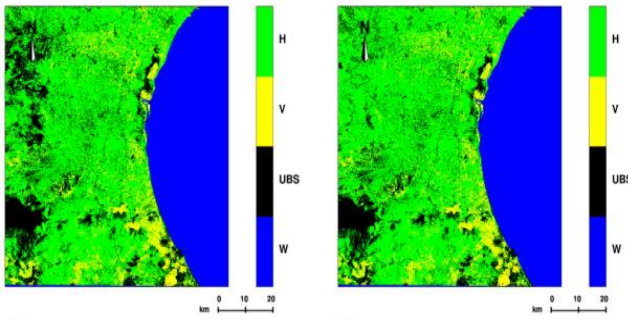

Figure 3: Halophyte cover map relative to SI selector: a SI1, b SI9. Land cover classes: W: water; UBS: urban and bare soil, DV: dense vegetation; V: vegetation; H: halophyte.

Finally, concerning the main aim to classify halophyte vegetation, the results reported in Table 1 for SI1 show that 86% of the halophyte areas have been correctly identified as halophyte land covers. Particularly, more than 73.5% of the areas called halophyte correspond actually to halophyte vegetation cover, while the 14% of halophyte areas were mistakenly classified as urban/bare soil (11%) and vegetation (3%).

Table 1: Confusion matrix related experiment SI1. Columns represent true classes, while rows represent the classifier's predictions. The matrix is square, with all correct classifications along the upper-left to lower-right diagonal.

		True data					
		W	UBS	DV	V	H	Total
Classified data	W	100	0	0	0	0	100
	UBS	0	86	0	0	11	97
	DV	0	0	89	50	3	142
	V	0	0	11	33	0	44
	H	0	14	0	17	86	117
	Total	100	100	100	100	100	500

The confusion matrix, relevant to all SI features, is shown in Table 1, where also UA, PA, OA, and K are listed, while the land cover percentage for each class is shown in Table 2. Note that the entries in Table 1 are descending-ordered according to their H. The confusion matrix shows that: a) UA and PA for W, UBS, and V show the same value for all features; b) UA estimates were higher than 73.1% per class. c) PA estimates were higher than 86% per class. d) OA for all the features is larger than 91%; e) K value for all the features is considered B almost perfect because in according to [7] is between 0.81 and 1.

Table 2: UA, PA, OA, and K related experiment. Features are listed in descending order according to their OA

	PA (%)				UA (%)				AO (%)	K
	W	USB	V	H	W	USB	V	H		
SI1	100	86.0	91.5	86.0	100	88.6	98.3	73.5	91.0	0.875
SI6		85.0		87.0		89.4		73.1		0.875
SI7		85.0		87.0		89.4		73.1		0.875
SI8		85.0		87.0		89.4		73.1		0.875
SI9		85.0		87.0		89.4		73.1		0.875
SI10		85.0		87.0		89.4		73.1		0.875
SI12		85.0		87.0		89.4		73.1		0.875

Table3: Land cover area percentage related experiment Features are listed in descending order according to their H

	Land cover area (%)		
	UBS	V	H
SI6	13.7	10.6	75.6
SI8	13.8		75.5
SI10	14.0		75.2
SI9	14.5		74.8
SI7	14.8		74.4
SI12	18.3		71.0
SI1	23.2		66.1

The classification results show that the decision tree algorithm applied to the integrated layer stack composed of Sentinel-2 VNIR based on SI6, SI7, SI8, SI9, SI10, and SI12 data outperformed the other configurations of classifiers: the halophyte cover percentage showed in Table 3 was higher than 50% of irrigated land. We were more confident about the low value of halophytes land cover (around 66%), thus we discuss here detailed results only from this classification approach. A visual inspection of the decision tree classification maps versus a digital elevation model (DEM) of our AOI shows that:

a) The main percentage of UBS class represents the main percentage of cover of AOI. This matches the results from the digital elevation model (DEM)

b) The eastern part of the study area is more heterogeneous and characterized by a mosaic of the four classes UBS, V, and H.

c) The major concentration of H falls in the northeast and in the south of the AOI. This result is in agreement to [8]. They reported that the spatial distributions of electric conductivity, total dissolved salts, and major ions are rather similar. Moreover, they showed an increase in salinity from the mountainous regions (defined as recharge areas) toward the depressed Sabkhas areas.

Finally, concerning the main aim to classify halophyte vegetation, the results reported in Table 2 for SI1 show that 86% of the halophyte areas have been correctly identified as halophyte land covers. Particularly, more than 73.5% of the areas called halophyte correspond actually to halophyte vegetation cover, while 14% of halophyte areas were mistakenly classified as urban/bare soil (11%) and vegetation (3%) (see Table 1). Thus, the critical point in this land cover analysis is represented by the discrimination between urban/bare soils and halophyte areas.

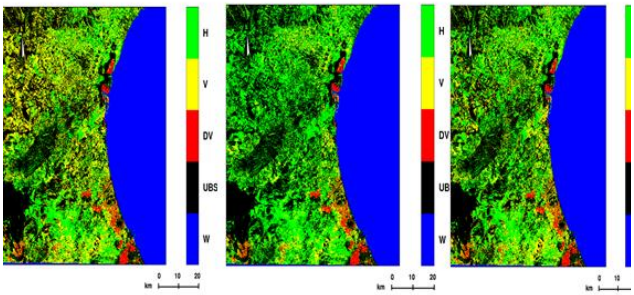

Figure 4: Halophyte cover map relative to SI selector: a SI1, b SI9, and c SI12. Land cover classes: W: water; UBS: urban and bare soil, DV: dense vegetation; V: vegetation; H: halophyte.

In this study, the sensitivity of SI features to halophyte vegetation cover is investigated using Sentinel-2 scenes collected over the territory of Ghannouch delegation, Tunisia. A simple decision tree was used to analyze the classification performance of each feature against a reference land cover map. The decision tree classification produced satisfactory accuracy results for all the classes considered.

4. CONCLUSION

This study is based on a multi-source methodology using remote sensing, spectral indices, physicochemical analysis, and a solid knowledge of the environment to verify the accuracy of the results of the classification of halophytes.

The Tunisian south is a favorable environment for the evolution of halophytes. Salinity indices play a key role in improving classification performance. The IS1 index provides the best results in terms of overall accuracy in the Ghannouch area. The implementation of appropriate decision tree procedures, based on environmental indices, is a method for the accurate monitoring of the distribution of vegetation.

5. REFERENCES

[1] M. Drusch, U. Del Bello, S. Carlier, O. Colin, V. Fernandez, F. Gascon, B. Hoersch, C. Isola, P. Laberinti, & P. Martimort, Sentinel-2: ESA's optical high-resolution mission for GMES operational services. Remote Sensing of Environment, 120, 25–36, 2012.

[2] C. Leprieur, Y. H, Kerr, S.Mastorchio, & J. C Meunier. Monitoring vegetation cover across semi-arid regions: comparison of remote observations from various scales. International Journal of Remote Sensing, 21, 281–300, 2000.

[3] S. K. McFeeters. The use of the normalized difference water index (NDWI) in the delineation of open water features. International Journal of Remote Sensing, 17, 1425– 1432, 1996.

[4] G. R. Hunt, J. W. Salisbury, C. Lenhoff, Visible and near-infrared spectra of minerals and rocks. V. Halides, phosphates, arsenates, vanadates, and borates. Modern Geology, 3, 121–132, 1972.

[5] G.Mancino, A. Nolè, F.Ripullone, , & A. Ferrara. Landsat TM imagery and NDVI differencing to detect vegetation change: assessing natural forest expansion in Basilicata, southern Italy. iForest, 7, 75–84, 2014.

[6] R. Jensen, Introductory digital image processing, Prentice Hall, Englewood Cliffs, New Jersey USA, 379, 1986.

[7] Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. Biometrics, 33(1), 159–174.

[8] Alaya, M. B., Saidi, S., Zemni, T., & Zargouni, F. (2014). Suitability assessment of deep groundwater for drinking and irrigation use in the Djeffara aquifers (Northern Gabes, south-eastern Tunisia). Environmental Earth Sciences, 71(8), 3387–3421.