Intercontinental Geoinformation Days

igd.mersin.edu.tr

Extraction of high-precision built-up areas from SENTINEL-2B imagery via multi-index approach and fuzzy C-means algorithm

Bachri Imane*[1], Zonkouan Badjo Ruth Virginia [2], Benbouziane Abdelmajid [3], Raji Mohammed [4], Hakdaoui Mustapha [1], Beh Ibrahim Diomande [2]

[1]*Hassan II University of Casablanca, Faculty of Science Ben M'sik, Laboratory of Applied Geology, Geomatics and Environment, Casablanca, Morocco*
[2] *Alassane Ouattara University, Laboratory of Demography and Spatial Dynamics, Bouaké, Ivory Coast*
[3]*Hassan II University of Casablanca, Faculty of Science Ben M'sik, Laboratory of Sedimentary Basin Dynamics and Geological Correlations, Casablanca, Morocco*
[4]*Hassan II University of Casablanca, Faculty of Science Ben M'sik, Laboratory of Laboratory of Geodynamics of Ancient Chains, Casablanca, Morocco*

Keywords
Built up area
High spatial resolution
Sentinel-2B
Multi-index
FCM

ABSTRACT

Accurate Urban built-up area information is required for a wide range of applications, in particular risk prevention and city planning. However, the extraction of built-up urban areas using high spatial resolution multispectral image, such as Sentinel-2, remains a significant challenge due to spectral confusion and intra-urban variability with other types of land cover, especially between built-up areas and bare land. As a result, in this work, we aim to increase built up accuracy mapping for Tan-Tan city (Southern of Morocco) by using six spectral indices, including Normalized Difference Building Index (NDBI), New Built Up (NBI), and Normalized Difference Tillage Index (NDTI) for urban area, and Normalized Difference Vegetation Index: NDVI, linked to vegetation, as well as the Bare Soil Index (BSI) and Dry Bare-Soil Index (DBSI) for bare land by means of Fuzzy C Means (FCM) algorithm. The six spectral indices were extracted from Sentinel-2 during the dry season and were combined to generate six multi-index datasets. Herein empirical results show that DBSI index works with NDBI, while BSI works better with NDTI. Therefore, the two multi-index datasets DBSI/ NDVI / NDBI and BSI/ NDVI / NDTI were suitable for built-up extraction in dry season in preferring order. Their overall accuracies were 85.28%, and 83.99%, respectively.

1. INTRODUCTION

Although urban areas make up only a small percentage of the world's total land area, their population density and resource use rates are much higher than in surrounding areas, necessitating better resource management practices. Understanding the spatial distribution and development of urban areas is critical for urban planning and resource planning, and mapping the built-up areas is among the most basic tasks needed for this (Ettehadi Osgouei et al., 2019) . When carrying out using traditional methods like field surveying and aerial photography, such mapping practice necessitates a significant amount of resources.

Satellite imagery is essential for the study of urban development, and the evolution of sensors towards very high spatial resolution provides an extremely rich set of metrics and sub-metric data, creating new challenges in remote sensing. With the advent of such images, the content of the data to be exploited has become much denser over the last ten years. The transition from medium to high resolution has been accompanied by a redefinition of the stakes in civil and military applications. Within a metric resolution image, the

* **Corresponding Author**

*(bachri.imane@gmail.com) ORCID ID 0000-0001-6146-5213
ORCID ID 0000-0002-4901-9920
ORCID ID 0000-0002-2759-2080
ORCID ID 0000-0002-2866-0188
ORCID ID 0000-0002-3598-7672

Cite this study

Bachri I, Zonkouan Badjo R.V, Benbouziane A, Raji M, Hakdaoui M & Beh Ibrahim D (2021). Extraction of high-precision built-up areas from SENTINEL-2B imagery via multi-index approach and fuzzy C-means algorithm. 2nd Intercontinental Geoinformation Days (IGD), 249-252, Mersin, Turkey

notion of object has become accessible. However, the extraction of built-up urban areas using high spatial resolution multispectral image remains a significant challenge due to spectral confusion and intra-urban variability with other types of land cover, especially between built-up areas and bare land. These detection accuracy problems require additional data other than the satellite image itself, digital elevation models (DEMs), reference images, LIDAR data, or other (Zhang et al., 2013). The aim of this study is to extract and detect urban areas using a multispectral image of high spatial resolution (HR) Sentinel-2 for the city of Tan-Tan based on a multi-index approach including the Normalized Difference Building Index, the New Building Index, the Normalized Difference Tillage Index, and the Normalized Difference Vegetation Index and the index of bare soils and the index of dry bare soils.

2. MATERIALS AND METHODS

2.1. Study Area

Tan-Tan is a Moroccan city, located in the region of Guelmim Essmara, southwest Morocco, 330 km south of the city of Agadir, this city has a mild and temperate climate is located between two wadis that flow into the ocean: Oued Drâa and Oued Chebika bordered by sand dunes. The city of Tan-Tan had 60,698 inhabitants according to the 2004 census.

Sentinel-2 images were chosen based on the fact that Sentinel-2 is a free cost satellite imagery provides high accuracy for built-up areas due to its high spatial resolution compared to others multispectral images. Furthermore, several studies have shown that Sentinel-2 image outperforms Landsat OLI image for urban mapping utilizing a variety of indices.

Table 1. Sentinel-2 spectral bands showing the central wavelength, with the resolution

Spectral Bands	Central Wavelength	Resolution
B1- Coastal Aerosol	443 nm	60 m
B2- Blue	490 nm	10 m
B3- Green	560 nm	10 m
B4- Red	665 nm	10 m
B5- Vegetation Red Edge	705 nm	20 m
B6- Vegetation Red Edge	740 nm	20 m
B7- Vegetation Red Edge	783 nm	20 m
B8- NIR	842 nm	10 m
B8A- Vegetation Red Edge	865 nm	20 m
B9- Water Vapor	945 nm	60 m
B10- SWIR-Cirrus	1375 nm	60 m
B11- SWIR	1610 nm	20 m
B12- SWIR	2190 nm	20 m

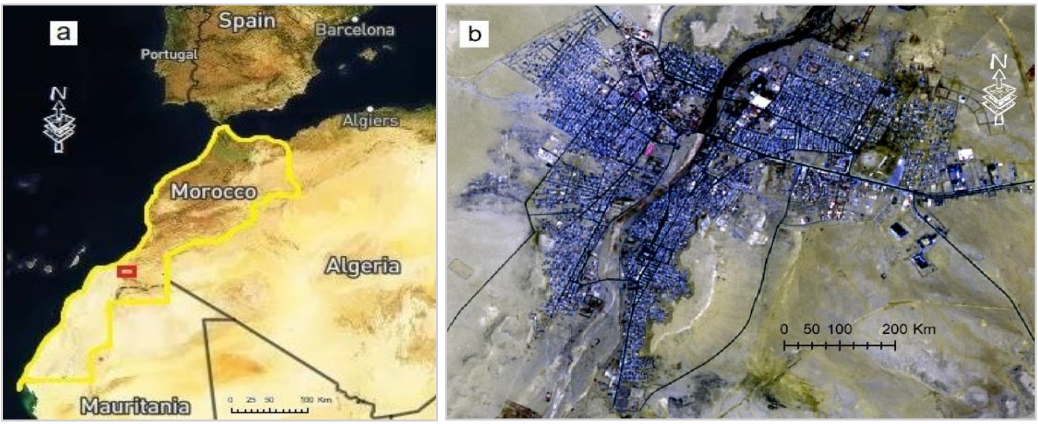


Figure 1. a. Location of study area Tan-Tan; **b.** Natural color composite (RGB 4-3-2) Sentinel-2B image.

2.2. Data Used

Sentinel-2 is a high-resolution multispectral imaging mission. It is based on a constellation of Sentinel-2A and Sentinel-2B, flying in the same orbit but 180° out of phase, and is designed to give a high revisit frequency of 5 days at the equator to observe the Earth's surface. The onboard sentinel-2 optical instrument has 13 spectral bands: four at 10 m, six at 20 m, and three at 60 m spatial resolution (Table 1). In the present work, the cloud-free Sentinel-2B images were downloaded for free from the Copernicus Open Access Hub website. The multispectral image acquired on June 30, 2020 level 1c was atmospherically corrected, resampled to 20 m using the nearest neighbor technique, clipped according to the study area, and exported in GeoTIFF format.

3. METHOD

The proposed method for extraction of built-up areas using sentinel-2B imagery consisted of three main steps: pre-processing of satellite imagery, processing including extraction of spectral indices and their combination as well as classification, and the last one is analysis of results.

3.1. Pre-processing of Sentinel-2B imagery

The Sentinel-2B level 1c image bands were converted from Top-of-Atmosphere (TOA) reflectance values to Bottom-of-Atmosphere (BOA) reflectance via Sen2Cor plugin. The corresponding spectral bands were considered in this process: blue, green, red, near infrared (NIR), shortwave infrared 1 (SWIR-1) and shortwave infrared 2 (SWIR-2). The SWIR-2 band has a resolution of 20 m; hence, it was resampled to 10 m resolution utilizing the nearest neighbor method. Then, the image was clipped according to the study area These steps were performed using the SNAP 8.0.0 software provided by European spatial agency (ESA).

3.2. Sentinel-2B processing

3.2.1. Spectral indices used

The spectral indices part of multispectral transformations that consist in transforming the luminance measured at the satellite sensor into meaningful quantities in the environmental domain. Due to the type of multispectral satellite data, it can describe the state of a phenomenon. An index is a numerical variable describing the intensity or extent of a phenomenon that is too complex to be broken down into a manageable number of parameters. Table 2 contains the six spectral indices used in this research. The SAGA GIS software was employed to calculate the six spectral indices used. The results of the spectral indices presented in Figure 2.

Table 2. Different spectral indices used in this research on Sentinel-2B image.

Index Name	Formula on sentinel-2 image	References
Bare Soil Index (BSI)	$\frac{(B11+B4)-(B8+B2)}{B11+B4+B8+B2}$	(Roy et al., 2002)
Dry Bare Soil Index (DBSI)	$\frac{(B11-B3)}{(B11+B3)}-NDVI$	(Rasul et al., 2018)
New Built Up (NB)	$\frac{\boldsymbol{B11*B4}}{\boldsymbol{B8}}$	(Chen et al., 2010)
Normalized Difference Built-up Index (NDBI)	$\frac{(B11-B8)}{(B11+B8)}$	(Zha et al., 2003)
Normalized Difference Vegetation Index (NDVI)	$\frac{(B8-B4)}{(B8+B4)}$	(Rouse et al., 1974)
Normalized Difference Tillage Index (NDTI)	$\frac{(B11-B12)}{(B11+B12)}$	(Van Deventer et al., 1997)

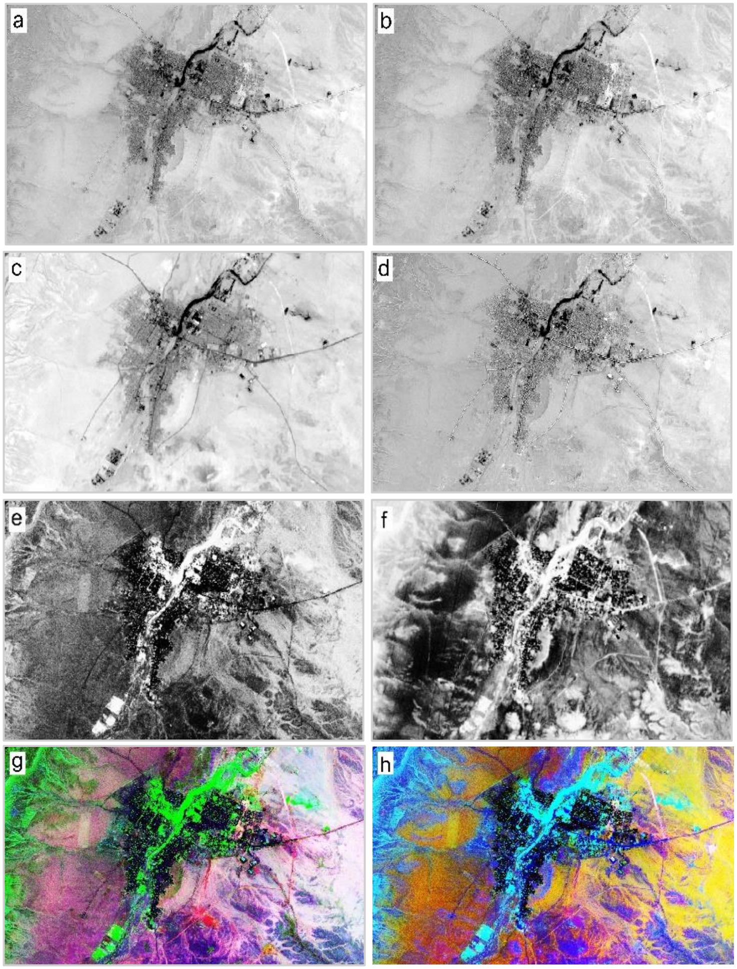
3.2.2. Multi-index combinations

An informal analysis of different combinations of spectral indices as components of the multi-index dataset was used to determine the most appropriate index combination. Table 3 illustrated the correlation between different indices. They discovered a strong association between the bare soil indices BSI and DBSI. That is, the two bare soil indexes generate nearly identical information over Tan-Tan region. The NDTI , NBI, NDBI indices, on the other hand, has poor associations with both the BSI and DBSI indices. The NDVI and NDTI, NDBI indices have a mean correlation. Based on the correlation between the six spectral indices, the spectral difference between the three main land cover classes (bare land, vegetation, urban area) could be expected to increase by overlaying two distinct combinations DBSI/NDVI/NDBI and BSI/NDVI/NDTI. The Quantum GIS 3.18 software was utilized to generate the required spectral layer stacking proposed.

Table 3. Values of correlation between spectral indices used in the research.

	BSI	DBSI	NBI	NDBI	NDTI	NDVI
BSI	1.00000	0.98416	0.23987	0.13283	0.44765	-0.26974
DBSI	0.98416	1.00000	0.32804	0.30715	0.43765	-0.26421
NBI	0.23987	0.32804	1.00000	0.21163	-0.48695	-0.47264
NDBI	0.13283	0.30715	0.21163	1.00000	-0.42525	0.56775
NDTI	0.44765	0.43765	-0.48695	-0.42525	1.00000	0.57795
NDVI	-0.26974	-0.26421	-0.47264	0.56775	0.57795	1.00000

Figure 2. Spectral indices result used (a) BSI (b) DBSI (c) NBI (d) NDBI (e) NDVI (f) NDTI (g) DBSI/NDVI/NDBI (h) BSI/NDVI/NDTI



4. RESULTS AND DISCUSSIONS

The results of the built-up area extracted utilized Sentinel-2B image and multi-index datasets created via Fuzzy C-Means Algorithm can be seen from the treated images in Figure 3.

The statistics that is, produced accuracy (PA), user's accuracy (UA), overall accuracy, and kappa coefficient produced from the result of the classified images are shown in Table 4 et 5. Based on the Accuracy of mapping of each types of data in Table 5, the multi-index datasets DBSI/NDVI/NDBI generated an overall accuracy of 85.28% and a coefficient kappa of 0.89, the multi index datasets BSI/NDVI/NDTI generated an overall accuracy of 83.99% and a coefficient kappa of 0.8 and the sentinel-2B image generated overall accuracy of 53.71% and a coefficient kappa of 0.67.

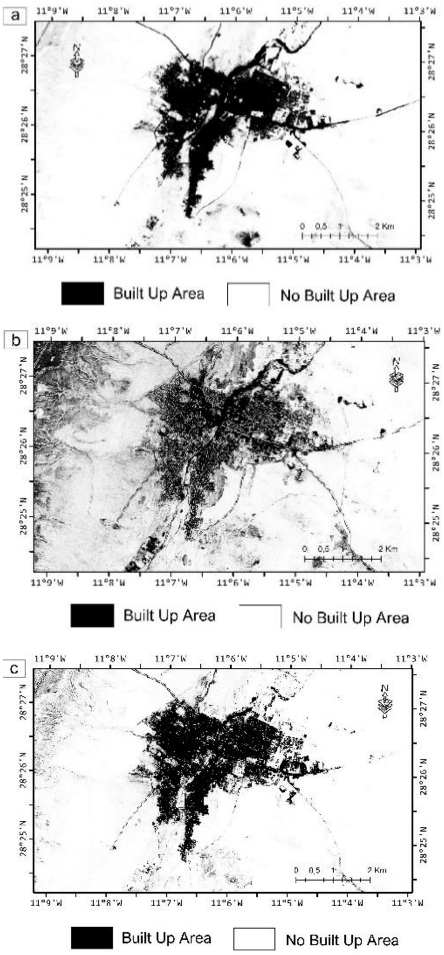


Figure 3. Results of the built-up area extracted of Sentinel-2B image and multi-index datasets created via Fuzzy C-Means Algorithm (a) Sentinel-2B image (b) dataset DBSI/NDVI/NDBI (c) dataset BSI/NDVI/NDTI

Table 4. Classification accuracies of the different approaches for the Tan-Tan region

Data	Sentinel-2B		DBSI/NDVI/NDBI		BSI/NDVI/NDTI	
Accuracy types	PA (%)	UA (%)	PA (%)	UA (%)	PA (%)	UA (%)
Built-Up	67.91	9.75	90.12	96.21	83.73	82.98
Bare land	47.49	71.20	97.52	75.69	87.52	97.65

Table 5. Tan-Tan region's total accuracy and Kappa statistics

Data Type	Overall Accuracy	Kappa
Sentinel-2B	53.71%	0.67
DBSI/NDVI/NDBI	85.28%	0.89
BSI/NDVI/NDTI	83.99%	0.80

According to the visual analysis of the results shown in Figure 3, it was apparent that the Sentinel-2B image and the multi-index created, distinguished built-up land from other land cover types differently. Both multi-index datasets created DBSI/NDVI/NDBI and BSI/NDVI/NDTI, which were derived from Sentinel-2B images and checked for built-up using mapping fuzzy c means classifier in Tan-Tan (arid region), showed high precision, rapid, and automatic classification of built-up class. Nevertheless, based on the statistical results (the overall accuracy and the Kappa coefficient) the best result was obtained by using the DBSI/NDVI/NDBI multi-index.

5. CONCLUSION

Extraction of built up areas using high resolution image has been the biggest issue in mapping urbanized areas due to spectral confusion and intra-urban variability with other types of land cover. In this study a multi-index approach for extracting of built up areas from Sentinel-2B satellite images in dry region was suggested. The Fuzzy C Means algorithm was used to classify the multi-index images generated with various index combinations. The DBSI/NDVI/NDBI combination greatly improved the misclassification of bare lands as built-up areas, according to the findings of this study (Overall accuracy 85.29% and kappa coefficient 0.89). but we must not neglect the fact that the combination BSI/NDVI/NDTI also generated good results(Overall accuracy 83.99% and kappa coefficient 0.8). These results indicate the effectiveness of the suggested multi-index. More research is expected to incorporate multi-polarization SAR Sentinel 1 data in order to use SAR data to distinguish between bare fields and built-up areas.

REFERENCES

Chen, J., Li, M., Liu, Y., Shen, C. & Hu, W. (2010). Extract residential areas automatically by new built-up index. *2010 18th International Conference on Geoinformatics, Geoinformatics 2010*. https://doi.org/10.1109/GEOINFORMATICS.2010.5567823

Ettehadi Osgouei, P., Kaya, S., Sertel, E. & Alganci, U. (2019). Separating Built-Up Areas from Bare Land in Mediterranean Cities Using Sentinel-2A Imagery. *Remote Sensing*, *11*(3). https://doi.org/10.3390/rs11030345

Rasul, A., Balzter, H., Ibrahim, G. R. F., Hameed, H. M., Wheeler, J., Adamu, B., Ibrahim, S. & Najmaddin, P. M. (2018). Applying Built-Up and Bare-Soil Indices from Landsat 8 to Cities in Dry Climates. *Land*, *7*(3). https://doi.org/10.3390/land7030081

Rouse, J. W. . J., Haas, R. H., Schell, J. A. & Deering, D. W. (1974). Monitoring Vegetation Systems in the Great Plains with Erts. *NASSP*, *351*, 309. https://ui.adsabs.harvard.edu/abs/1974NASSP.351..309R/abstract

Roy, P., Rikimaru, A. & Miyatake, S. (2002). Tropical forest cover density mapping. *Tropical Ecology*, *Vol 43*, 39–47.

Van Deventer, A. P., Ward, A. D., Gowda, P. H. & Lyon, J. G. (1997). *Using Thematic Mapper Data to Identify Contrasting Soil Plains and Tillage Practices*.

Zha, Y., Gao, J. & Ni, S. (2003). Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *International Journal of Remote Sensing*, *24*(3), 583–594. https://doi.org/10.1080/01431160304987

Zhang, J., Li, P. & Xu, H. (2013). Urban built-up area extraction using combined spectral information and multivariate texture. *International Geoscience and Remote Sensing Symposium (IGARSS)*.