IOSR Journal of Applied Geology and Geophysics (IOSR-JAGG)
e-ISSN: 2321–0990, p-ISSN: 2321–0982.Volume 9, Issue 5 Ser. I (Sep. – Oct. 2021), PP 34-40
www.iosrjournals.org

GIS based approach for NDVI application to Landsat 7 ETM+ imagery and thematic accuracy evaluation

Francesco GiuseppeFigliomeni[1], Claudio Parente[1]
[1](Department of Science and Technology/ University of Naples "Parthenope", Italy)

Abstract: *GIS software allow functions which are suitable for processing remotely sensed images. Map Algebra tools and classification tools are just two examples of GIS functions usable to process remotely sensed data and obtain from them further information that can be represented in the form of thematic map.In a similar way, NDVI (Normalized Difference Vegetation Index), the useful measure of live green vegetation obtained with remotelysensed images, can be easily implemented using Map Algebra applied to the reflectance of the scene in red and infrared bands. Using a classification method, NDVI layer can be transformed in a land cover map distinguishingthree classes: waters (oceans, seas, lakes, rivers, ...), soils (plowed land, rocks,constructions, roads, ...), vegetation (woods, grasses, vineyards, ...).*
This article aims to demonstrate that Landsat 7 ETM+ images (band 3 and band 4) can be processed using the free and open source software named Quantum (Q) GIS to achieve an accurate land cover map. First of all, "topof atmosphere" (TOA) reflectance, i.e. the solar radiation coming from the earth surface and incident on the satellite sensor, is derived by digital numbers of the initial images using Map Algebra tools. Then NDVI is calculated and the resulting grid is classified using Maximum Likelihood (ML) algorithm based on training sites.The quality of the resulting map is evaluated in a rapid process carried out by using the same software. Particularly, a confusion matrix is realized and appropriate indices used to define the accuracy level of the thematic map, namely Producer Accuracy, User Accuracy and Overall Accuracy. The results demonstrate the correctness of the adopted approach and the high accuracy level of the land cover map.
***Key Word**: NDVI; Landsat 7 ETM+; GIS; TOA; Map Algebra.*

Date of Submission: 06-09-2021 Date of Acceptance: 20-09-2021

I. Introduction

Coined by Roger Tomlinson in 1963[1], the phrase "Geographic Information System" (GIS)refers broadly to a computer-based system of hardware and software capable of capturing, storing, analyzing, and displaying geographically referenced data[2]. From the very beginning, GIS have attracted interest due to their capability to analyze data over larger areas and with more variables than would be possible with traditional (non-computerized) manual methods [3].These systems have evolved considerably over time and today have very powerful tools capable of processing all types of data, therefore also remotely sensed images. One of the most effective tool for this purpose is certainly the Map Algebra, a set of algebraic operations, proposed by Dana Tomlin [4], for manipulating two or more raster layers ("maps") of similar dimensions and producing a new raster layer (map). Map Algebra includes, among others, arithmetic operations (addition, subtraction, multiplication, division), statistical operations (minimum, maximum, average, median), relational operations (greater than, smaller than, equal to), trigonometric operations (sine, cosine, tangent, arcsine), exponential and logarithmic operations (exponent, logarithm)[5].

Map Algebra tool can be successfully used for converting the brightness values of the remotely sensed images into reflectance at the top of the atmosphere (TOA) as well as for atmospheric corrections. In fact,TOA reflectance, the solar radiation incident on the satellite sensor[6] can be easily derived from BV using formulas requiring from time to time the single band and specific parameters comprised in the metadata file included with the dataset.Atmospheric corrections are carried out using model such as the dark-object subtraction (DOS) model [7]. Map Algebra tools enable the user to calculateTOA andexecute atmospheric corrections.

Map Algebra tools can be successfully used to implement vegetation indices, too: designed to maximize sensitivity to the vegetation characteristics while minimizing confounding factors such as soil background reflectance, directional, or atmospheric effects[8], these indices result from formulas relating different bands (usually red and infrared) that can be easily applied in GIS. However, the application of a vegetation index does not produce a cover land map: theresulting image is a graphical indicator that can support the analysis of the remote sensing measurements for evaluating whether the observed pixel contains vegetation or not. For consequence a classification process is necessary to produce vegetation/no-vegetation map. Several methods are available for classification of remotely sensed images and a lot of them are included in GIS software such as minimum distance [9], maximum likelihood [10], K-means [11], etc.

This paper aims to demonstrate that it is possible to process Landsat 7 ETM+ images (band 3 and band 4) using the free and open source software named Quantum (Q) GIS to achieve an accurate thematic map distinguishing three land cover classes: water, bare soils and vegetation. In particular, solar radiation incident onthe satellite sensor, currently known as “topof atmosphere” (TOA) reflectance, is derived by the digital numbers(DN) of the initial images. Then, the most studied vegetation index called Normalized Difference Vegetation Index (NDVI) [12] is calculated and classified using Maximum Likelihood(ML) algorithm based on training sites. Finally, the accuracy of the resulting thematic map is valuated: starting from test sites, confusion matrix[13] is assembled and quality indices (QI), such as Producer Accuracy, User Accuracy and Overall Accuracy[14], are calculated to evaluate the results.

The article is organized as follows. Section 2 presents data and methods: firstly, the main characteristics of Landsat 7 ETM+ imagery are described; then the methods applied in our experiments (NDVI, ML, QI) are presented. Section 3 provides the results of the adopted methods and the discussion in relation to them. Section 4 reports an overview of the outcomes and further considerations.

II. Material And Methods

This study is carried out on Landsat 7 ETM+ imagery acquired on 2000-08-02and concerning a largearea around the Gulf of Naples (Italy) including Naples city and the Vesuvius Vulcan. All experiments are carried out in Laboratory of Geomatics, GIS and Remote Sensing of the University of Naples “Parthenope” by using QGIS, a free and open source GIS software, version 3.16.

Landsat 7 ETM+ imagery: The Landsat-7 satellite is part of NASA's ESE (Earth Science Enterprise) program, a joint venture of NASA and USGS (United States Geological Survey); the mission aims to extend and improve upon the long-term collection of medium-resolution multispectral imagery of the Earth's surfaces provided by the earlier Landsat satellites[15].

Landsat 7 satellite was launched on April 15,1999 and remains operative. It is a sun synchronous satellite and orbits the Earth at an altitude of 705 kilometers (438 miles) moving from north to south (descending orbit) over the sunlit side of the Earth. The satellite makes a complete orbit every 99 minutes, completes about 14 full orbits each day, covering the whole planet in 16 days with a swath equal to 185 kilometers[16].

The sensor for detecting earth reflectance on board of Landsat 7 is called Enhanced Thematic Mapper Plus(ETM+), with a 15-meter panchromatic band;30-meter visible, near-IR and SWIR (shortwaveinfrared)bands; and a 60-meter thermal band;thermal data are resampled to 30 meters[17]. Details on sensor characteristics are shown in Table 1.

Table no 1: Landsat 7 ETM+ sensor characteristics.

Spectral Band	Wavelengths (micrometers)	Geometric resolution (meters)
Band 1	0.45-0.52	30
Band 2	0.52-0.61	30
Band 3	0.63-0.69	30
Band 4	0.76-0.90	30
Band 5	1.55-1.75	30
Band 6	10.40-12.50	60
Band 7	2.08-2.35	30
Band 8	0.52-0.90	15

For this study Landsat ETM+ images of Campania Regionacquired on 2000-08-02 are downloaded fromUSGS web site [18]. A clip of this dataset is considered for the further applications extending between the following UTM-WGS84 (zone 33 N) plane coordinates: E_1 = 414,495 m; E_2 = 474,495 m; N_1 = 4,505,235 m; N_2 = 4, 550,235 m. The territorial framework of the initial dataset is shown in Figure no 1; the study area is shown in Figure no 2: in both case RGB composition (true color) of the Landsat 7 ETM+ imagery is given.

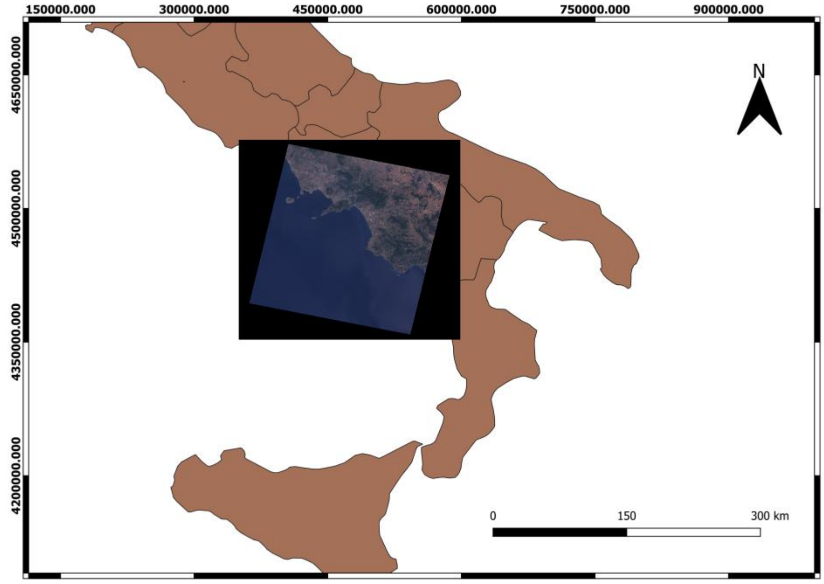

Figure no 1: Territorial framework of the considered dataset Landsat ETM+ imagery in true color RGB composition.

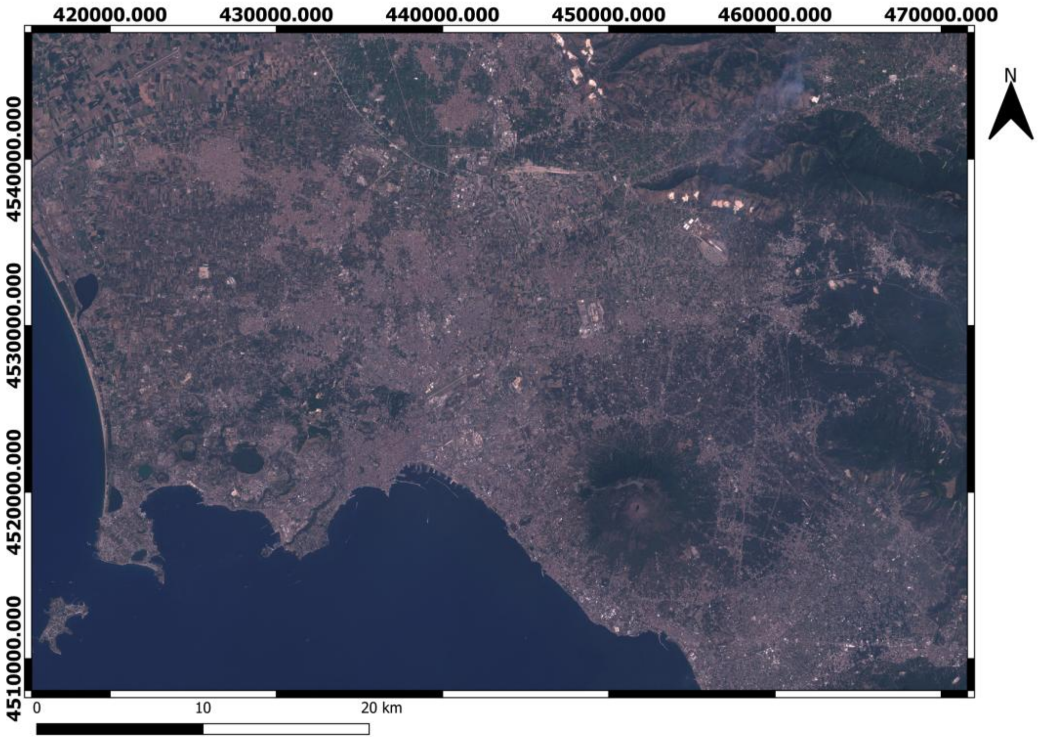

Figure no 2: The study area: Landsat ETM+ imageryin true color RGB composition.

TOA calculation: Landsat 7 data from the USGS are fitted in 8 bit files (rangesfrom 0 to 255). In order to calculate NDVI,reflectance values, i.e. the physical measurements of the part of the solar energy reflected by earth features, are required.At first DNs are transformed to radiance using theformula [19] .

$$L_\lambda = (gain * DN) + bias_\lambda \ (1)$$

Where:
Lλ is the calculated radiance associated to the ground areaenclosed in the pixel and referred to the λ wavelength range ofthe specific band;
DN is the digital number of the pixel of the Landsat 7ETM+ band image;

gain and bias$_\lambda$ are sensor-specific calibration parametersdetermined before the launch.
Table 2 shows gain and bias values for bands used in thisapplication and available in literature [7].

Table no 2: Gain and Bias values for Red and NIR bands.

SpectralBand	Gain	Bias
Band 3	0.621654	-5.62
Band 4	0.639764	-5.74

At-sensor radiance values are converted to TOA reflectance values using the following formula [7, 19]:

$$R_\lambda = \frac{\pi * L_\lambda * d^2}{E_{sun,\lambda} * sin(\vartheta_{SE})} \quad (2)$$

Where:
$R_\lambda$is the reflectance referred to the λwavelength range of the considered band;
L_λ is the spectral radiance (from earlier step);
d is the Earth-Sun distance in astronmoical units;
$E_{SUN,\lambda}$is the mean solar exoatmospheric irradiance referred to the λ wavelength range of the considered band;
$\theta_{SE}$is theSunelevation angle.

The value of d and $E_{SUN,\lambda}$are available in literature [7] while $\theta_{SE}$is reported in the metadata file included with the dataset.

Normalized Difference Vegetation Index (NDVI): The vegetation is characterized by a unique spectral signature that easily distinguishes it from other objects present on the Earth surface.Plants absorb light in blue and red and reflect in green and near infrared; particularly, the reflection is considerable in the near infrared creating a marked step with red (R). This step is the main basis for calculating a "Vegetation Index" to quantify the concentrations of green leaf vegetation around the globe; then by combining the daily Vegetation Indices into 8-, 16-, or 30-day composites, scientists create detailed maps of the Earth's green vegetation density that identify where plants are thriving and where they are under stress (i.e., due to lack of water) [20].
The Normalized Difference Vegetation Index (NDVI), the most studied vegetation index, evaluates the presence of photosynthetic activity as it relates the spectrum of red, in which there is absorption by chlorophyll, and that of the near infrared in which the leaves reflect or transmit light to avoid overheating. The values range from-1 to 1: since their high level of reflectance in near-infrared spectrum and low level of reflectance in red spectrum, vegetated areas generally return NDVI high values; bare soiland buildingareas have similar reflectance in the two spectrum regions and produce values near zero; water, clouds, and snow have larger reflectance in Red than in NIR, so these features present negative values [21].For calculating NDVI two images of the same scene, simultaneously acquired, one concerning the reflectance in Red, the other in NIR, are necessary.In the case of Landsat 7 ETM+ imagery, the formula for NDVI calculation is:

$$NDVI = \frac{Band\ 4 - Band3}{Band\ 3 + Band4} \qquad (3)$$

NDVI classification and land cover map production: NDVI can be classified using supervised approach based on Maximum likelihood algorithm. The supervised classification is based on a relationship between a known class assignment and characteristics of the entity to be classified[22]. Maximum likelihood classification assumes that the statistics for each class in each band are normally distributed, and calculates the probability that a given pixel belongs to a specific class; unless a probability threshold is selected, all pixels are classified; each pixel is assigned to the class that has the highest probability[23].

Thematic Accuracy evaluation: The accuracy of classification is evaluated using test sites representative of the classeswhich are investigated. In fact, samples of testing pixels are selected on the classified image, then their class identity is compared with the reference data which represent the ground truth as recognizable through visual analysis and photointerpretation of true panchromatic image, true color and false color compositions of the multispectral bands. These areas are chosen taking care to avoid thatoverlapped with those used as training sites.The pixels of agreement and disagreement are shown in a matrix, called confusion matrix,which allows to quantify the accuracy level through the following indices [13-14]:

Producer Accuracy (PA), which defines the percentageof pixels correctly classified among the test pixels ofgiven class;
User Accuracy (UA), which defines the percentage ofpixels correctly classified among the pixels assigned togiven class;
Overall Accuracy (OA), which defines the percentage ofpixels correctly classified on the entire test set.

III. Resultsand discussion

The formulas(2), (3) and (4) are implemented in QGIS using Raster Calculator starting from Red and NIR images previously processed (DN converted in TOA reflectance and atmospheric corrections done through DOS model). The resulting NDVI image is shown in Figure no 3.

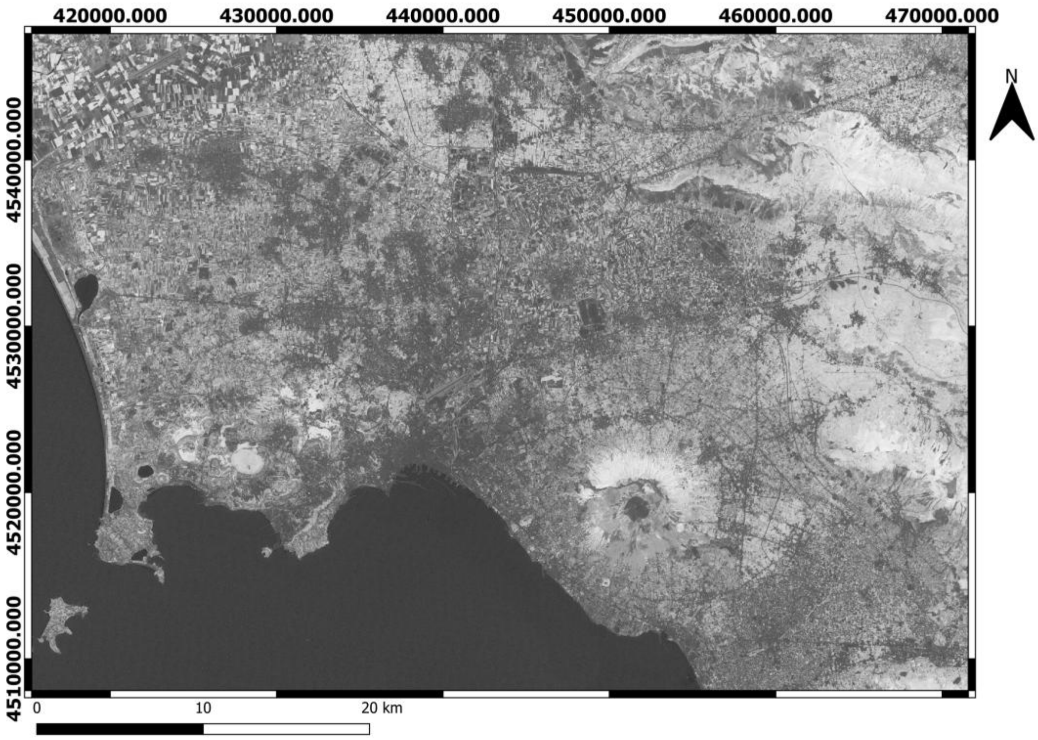

Figure no 3: NDVI image resulting from Red and NIR bands of Landsat ETM+ imagery.

The supervised classification of NDVI image is carried outusing SAGA-GIS tools available in QGIS (SAGA, Module Supervised Classification for Grids); particularly, the application of the Maximum Likelihood algorithm defines the following thresholds: -0.390 for water/soil and 0.287 for soil/vegetation. The resulting land cover map is shown in Figure no 4.

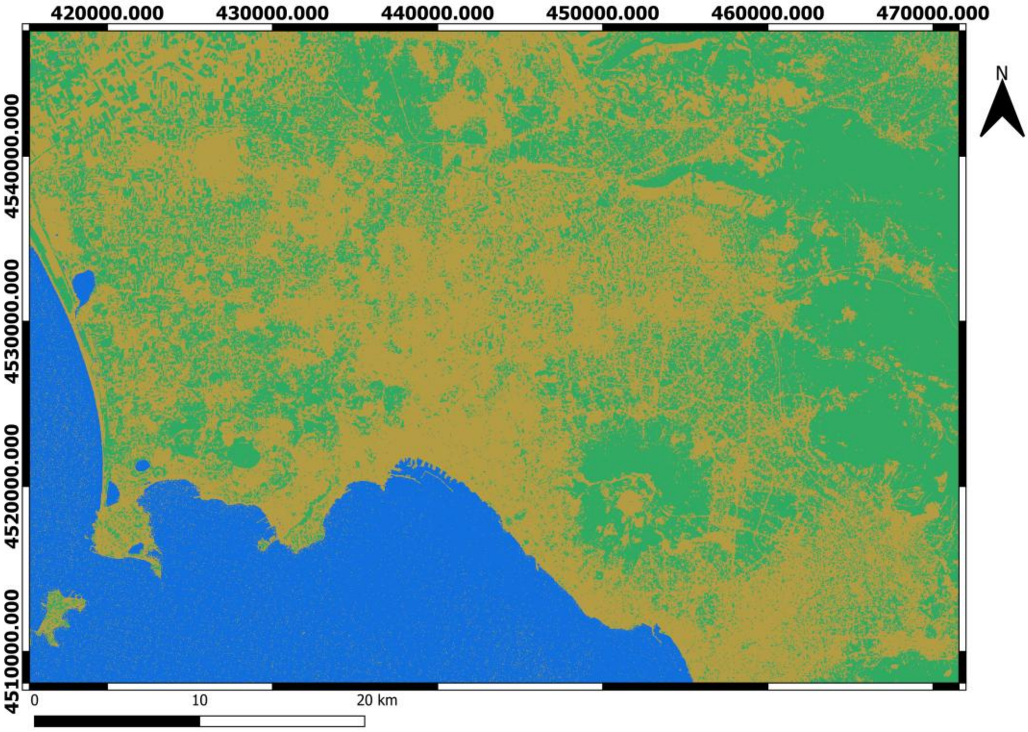

Figure no 4: The land cover map resulting fromNDVI classification in three classes: water (blu), soil (brown) and vegetation (green).

Table no3 shows the confusion matrix related to the accuracy of the land cover map.In this confusion matrix, the system recognizes all the pixels (1528) of water; of the 1293 pixels of soil, it judged that 1is water, 1284 are soil and 8 are vegetation; of the 1281 pixels of vegetation, it predicted that 4are soil and 1277 are vegetation. All correct predictions are located in the diagonal of the table, so it is easy to visually inspect the table for prediction errors, as values outside the diagonal will represent them.

Table no 3: The confusion matrix related to the accuracy of the land cover map.

	Water	Soil	Vegetation	
Water	1528	0	0	1528
Soil	1	1284	8	1293
Vegetation	0	4	1277	1281
	1529	1288	1285	4102

Table no4 shows the quality indices calculated for thematic accuracy evaluation. Good results are achieved for all indices. In fact, PA and UA present high values for all considered classes: PA is equal 1 for water, 0.993 for soil and 0.997 for vegetation; UA is equal 0.999 for water, 0.997 for soil and 0.9994 for vegetation. For consequence, the overall accuracy is near to 1 (0.997), confirming the high performance of the NDVI classification.

Table no 4: Quality indices for thematic accuracy evaluation..

	Water	Soil	Vegetation	
Producer accuracy	1	0.993	0.997	
User accuracy	0.999	0.997	0.994	
Overal accuracy				0.997

IV. Conclusion

This study demonstrates that Map Algebra tools usually available in GIS software allow to easily convert the DNs of remotely sensed satellite images, such as Landsat 7 ETM+, in TOA reflectance as well as to apply NDVI. In fact, the first operation only requires algebraic functions involving the specific band and related parameters which are: available in literature (i.e. gain and bias), comprised in the metadata file included with the dataset (i.e. local sun elevation angle), or derivable by specific tables reported in literature (i.e. the Earth-Sun distance and the mean solar exoatmospheric irradiance referred to the λ wavelength range of the considered band)in dependence of the imagery acquisition date. The second operation requires algebraic functions mutually involving two specific bands (red and infrared).

The algorithms for image classification, e.g. Maximum likelihood, are useful to derive land cover map from NDVI distinguishing at least three different classes: water, soil and vegetation. The operation requires the selectionof samples of testing pixels on the classified image and the comparison of their class identity with the reference data (ground truth). This process is implementable in GIS, too.

The confusion matrix and related indices, such as PA, UA and OA, testify the high level of thematic accuracy. In fact, water pixel are all correctly classified, but also for soil as well as vegetation pixels the percentages of correctly classified ones is very high.

Definitely, the experiments demonstratethat all operations for processing Landsat 7 ETM+ up to the production of a thematic map that distinguishes three land cover classes (water, soil and vegetation) can be carried out using a free and open source software such as QGIS, version 3.16.

References

[1]. Tomlinson RF. A geographic information system for regional planning. Journal of Geography (ChigakuZasshi), 1969; 78(1): 45-48.
[2]. Montana L, Geographic Information System, International Encyclopedia of Public Health, Elsevier. 2008.
[3]. Shaw DM, Atkinson SF. An introduction to the use of geographic information systems for ornithological research. The Condor, 1990; 92(3):564-570.
[4]. Tomlin CD. Map algebra: one perspective. Landscape and Urban Planning, 1994; 30(1-2): 3-12.
[5]. GISGeography Team. What is Map Algebra? [Raster Math], https://gisgeography.com/map-algebra-global-zonal-focal-local/
[6]. D'Allestro P., Parente C. GIS application for NDVI calculation using Landsat 8 OLI images. International Journal of Applied Engineering Research, 2015; 10(21): 42099-42102.
[7]. Chander G, Markham BL, Helder DL. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. Remote Sensing of Environment, 2009; 113: 893-903.
[8]. Fang H, Liang S. Leaf Area Index Models in: Reference Module in Earth Systems and Environmental Sciences, Elsevier, 2014.
[9]. Wacker AG, Landgrebe DA. Minimum distance classification in remote sensing. LARS Technical Reports, Laboratory for Application of Remote Sensing, Purdue University,1972.
[10]. Sisodia PS, Tiwari V, Kumar A. Analysis of supervised maximum likelihood classification for remote sensing image. In: International conference on recent advances and innovations in engineering (ICRAIE-2014), IEEE, 2014.
[11]. Na S, Xumin L, Yong G. Research on k-means clustering algorithm: An improved k-means clustering algorithm. In: Third International Symposium on intelligent information technology and security informatics, IEEE, 2010.
[12]. Rouse JW, Haas RH, Schell JA, Deering DW. Monitoring vegetation systems in the Great Plains with ERTS.In Third ERTS Symposium, NASA SP-351 I, 1973: 309-317.
[13]. Congalton RG, Green K. Assessing the accuracy of Remotely Sensed Data – Principles and Practices. CRC Press – Tailor & Francis Group, FL (USA), 2010.
[14]. Liu C, Frazier P, Kumar L. Comparative assessment of the measures of thematic classification accuracy. Remote Sensing of Environment, 2007; 107: 606-616.
[15]. ESA Earth Observation Portal. Landsat-7, https://directory.eoportal.org/web/eoportal/satellite-missions/l/landsat-7
[16]. USGS. What are the acquisition schedules for the Landsat satellites?, Mapping, Remote Sensing, and Geospatial Data, https://www.usgs.gov/faqs/what-are-acquisition-schedules-landsat-satellites?qt-news_science_products=0#qt-news_science_products
[17]. USGS. Landsat—Earth Observation Satellites, Fact Sheet: 2015-3081, https://doi.org/10.3133/fs20153081
[18]. USGS website: http://earthexplorer.usgs.gov
[19]. Parente C. TOA reflectance and NDVI calculation for Landsat 7 ETM+ images of Sicily. In The 2nd Electronic International Interdisciplinary Conference, 2013.
[20]. Weier J, Herring D. Measuring vegetation (NDVI & EVI), 2000. http://earthobservatory.nasa.gov/Features/MeasuringVegetation/
[21]. Lillesand T, Kiefer RW, Chipman J. "Remote sensing and image interpretation", John Wiley & Sons, 2014.
[22]. Nisbet R, Miner G, Yale K. Classification. In: Handbook of Statistical Analysis and Data Mining Applications, 2018; 9: 169-186.
[23]. Park B. Quality Inspection of Poultry Carcasses, in: Computer Vision Technology for Food Quality Evaluation, Elsevier, 2008.

Francesco GiuseppeFigliomeni, et. al. "GIS based approach for NDVI application to Landsat 7 ETM+ imagery and thematic accuracy evaluation." *IOSR Journal of Applied Geology and Geophysics (IOSR-JAGG)*, 9(5), (2021): pp 34-40.