Praticò S., Di Fazio S., Modica G. (2021) Multi Temporal Analysis of Sentinel-2 Imagery for Mapping Forestry Vegetation Types: A Google Earth Engine Approach. In: Bevilacqua C., Calabrò F., Della Spina L. (eds) New Metropolitan Perspectives. NMP 2020. Smart Innovation, Systems and Technologies, vol 178. Springer, Cham. https://doi.org/10.1007/978-3-030-48279-4_155

Multi Temporal Analysis of Sentinel-2 Imagery for Mapping Forestry Vegetation Types: a Google Earth Engine Approach

Abstract. The use of satellite imagery and Remote Sensing (RS) techniques for vegetation monitoring is increasingly recognized and used in several research fields. Multi-temporal analysis of multispectral images and derived vegetation indices are becoming a useful tool to detect forestry vegetation types. However, the use of multi-temporal images is a time-consuming activity for searching and downloading data and need a huge amount of storage space and workstations with high computing performance. To solve the problem, in this paper, it is proposed the use of Google Earth Engine (GEE) as cloud platform for geospatial analysis, which allows visualization and processing of RS. The paper is articulated in three sections and presents the first results of an ongoing research, on the use of GEE to analyze Sentinel-2 images and derived vegetation indices time-series to compute a supervised classification in order to map forestry vegetation types. In the first section an introduction about RS and supervised and unsupervised classification methods to map forestry vegetation is given. In the second section the proposed method is explained and tested in a square plot of 100 km^2 inside the protected area of Aspromonte National Park, in the South of Italy. In the last one, the main results of classification process are analyzed.

Keywords: Google Earth Engine (GEE), Sentinel-2, Vegetation Indices (VIs), Satellite Time Series, Remote Sensing.

1 Introduction

In the last decade, the correct management and protection of Natura 2000 sites, and more in general of the natural protected areas, have acquired ever-growing importance for the European Union (EU) member States. The Natura 2000 is a network of sites across Europe characterized by habitats recognized of a high natural value due to the presence of both animal and vegetation biodiversity to be protected. To achieve the obligations of Natura 2000 network, Member States have to report, every 6 years, the conservation status of the habitats under protection. Forest monitoring, also in terms of spatial arrangement, can represent an important issue to implement medium- and long-term sustainable landscape-planning strategies, also with the aim to direct future forest management so as to strengthen those change trends which will have shown positive effects and contrast those with negative ones [1].
Moreover, forest monitoring should take into account how to correctly assess the forest degradation that affects many Mediterranean forests and represents a major issue in sustainable forest management (SFM) [2] .

Remote Sensing (RS) techniques and Geographic Information Systems (GIS) are widely used for landscape and habitat monitoring and planning, ecosystem services management and, more in general, for vegetation mapping and Land Use/Land Cover (LU/LC) analysis due to the growing and easy availability of multi-sources data and free satellite imagery from sources like Landsat (https://www.usgs.gov/land-resources/nli/landsat) and Sentinel (https://sentinel.esa.int/web/sentinel/home) [3 - 20]. As far as vegetation mapping is concerned the use of RS methods in place of conventional mapping techniques offer some advantages. First of all RS allows a faster maps productions specially in those areas not easily accessible [21]. Weiers et al. [22] used RS facilities to monitor natural habitats at regional scale in Germany. Chetan et al.[23] used Landsat time series to detect the changes occurred in selected protected areas in Romania. Bock et al. [24] used satellite RS to compare object-based and pixel-based classifications in some Natura 2000 areas. Diaz Varela et al. [25] used Landsat images to map vegetation in Natura 2000 sites of Spain. Wang et al. [26] developed a multi-scale method to monitor changes of Land Use/Land Cover (LU/LC) in some northeastern USA natural parks. Pôças et al. [27] applied landscape metrics on LU/LC maps obtained by RS to monitor changes in some Natura 2000 sites of Portugal. In the framework of the BIO_SOS research project (http://www.biosos.eu), Mairota et al. [28] used RS techniques to quantify the human impact on natural areas in South Italy, while Lucas et al. [29] combined pixel-based and object-based techniques to classify Land Use/Land Cover (LU/LC) and different habitat types.

Extraction of forestry vegetation types with a traditional spectral approach is a not easy challenge because of the spectra similarity between various species. In order to achieve a correct classification, especially in forest areas, it is necessary to compute deep multi-temporal and multi-scale analysis to detect different vegetation types using their phenological characteristics. Thus, it is necessary an analysis of multi-date composite images to characterize the vegetation seasonal variations due to growth/senescence alternation [30–33].

It has been widely demonstrated that, when working on forest environments, time-series composite imagery improve the classification results [34], but present some specific problems. In particular, data searching and downloading can result as time-consuming activities: moreover, they are needed data-storage devices of great capacity (to store petabyte-scale of RS image archives) and workstations with high performance systems (to compute and run all the necessary algorithms).

These problems can be solved using cloud-based computation platforms like Google Earth Engine (GEE - https://earthengine.google.com). GEE is a cloud-based platform that allows, through a JavaScript-based code editor, an easy access to huge set of free RS data and computing resources to process geospatial big datasets [35,36].

In this paper, making reference to application in a representative protected area of South Italy, they are shown the advantages brought by the use of GEE for classifying forestry types starting from Sentinel-2 time-series and derived vegetation indices.

The method proposed was developed in GEE code editor environment (Fig.1) and written using JavaScript-based language. This represents just a first step of an ongoing research. The proposed method, anyway, already shows high potential for further improvement, so as to support effectively future research and application.

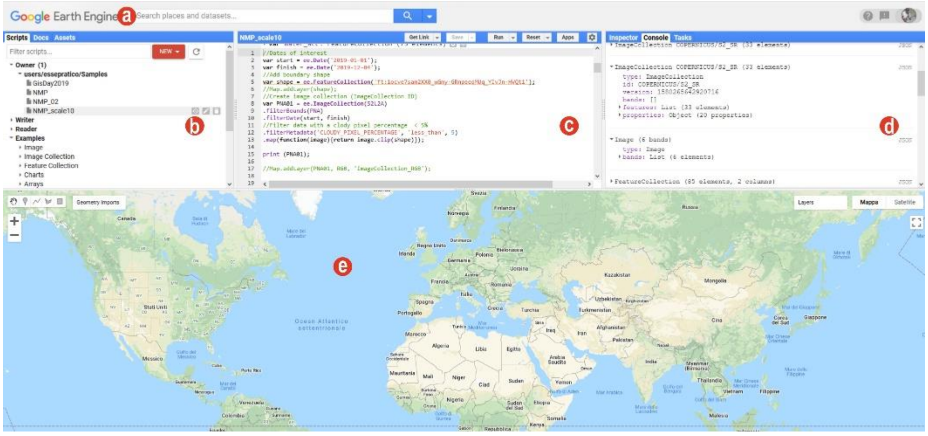

Fig. 1. Google Earth Engine (GEE) code editor environment. (a) is the searching area for datasets and place, (b) is the scripts management area, (c) is the core of code editor, the area where all scripts are written, (d) is the output management area and (e) is the area for map and input visualization.

2 Materials and Methods

2.1 Study area

The study area falls within the territory of the Aspromonte National Park in Calabria Region (South Italy). This is a natural protected area characterized by high biodiversity in vegetation and animals and represents a mountainous landscape typical of Southern Italy. With the aim to test the proposed method, inside the boundary of the Park, a squared plot of 100 km^2 was chosen as study area (Fig.2).

2.2 Images acquisition and processing

This study is based on the use of Sentinel-2 (A and B satellites) images. Starting from the Sentinel-2 archives hosted in GEE repository, the 2019 time-series was created by filtering all the images available for the study area falling in the time interval 2019/01/01– 2019/12/31. An additional filter was then applied in order to discard those images affected by a cloud coverage >5%. The time-series obtained was composed of 33 images.

In order to obtain phenology information useful to improve the classification accuracy, for each image the Normalized Difference Vegetation Index (NDVI), representing the ratio between difference and sum of Near InfraRed (NIR) band and Red band [37], was calculated by using the following formula and added as a band to the images

$$\frac{(Nir_{833} - Red_{665})}{(Nir_{833} + Red_{665})} \tag{1}$$

Fig. 2. Study area.

For each date of the time interval chosen, only visible (Blu – Green – Red), and NIR bands were extracted from the original image datasets and then stacked with the NDVI. All the input Sentinel-2 images constitute a 3D data-cube consisting of the union of the 2D spatial information given by the coordinates of each pixel and the 1D spectral information of each pixel given by the reflectance value for each considered band. The data fusion process allowed to obtain one image (i.e., a five-band orthomosaic in our case), starting from a series of images of the same area. For the adopted methodology, data fusion was conducted on the basis of the mean pixel value for each band. Thus, each pixel of the obtained composite orthomosaic contains, for each selected band, the mean value calculated considering the entire time-series.

2.3 Classification process

Making reference to the main characteristics of the study area, and considering that this represents just a starting step to be implemented for future elaborations, for the classification process only four Land Cover (LC) classes were considered: water (due to the presence of an artificial basin in the scene), bare soil, deciduous trees and coniferous trees. Once obtained the resulting composite image, a set of training and validation points was selected in GEE with a visual approach. A total number of 400 points, 100 for each of the defined LC classes, were randomly sampled inside the study area boundaries. This set of points was split to be used either as trainer (25 points – ¼ of the total) and validation (75 points – ¾ of the total) points.

In order to perform a supervised classification process in GEE environment, among the available algorithms the Classification And Regression Trees (CART) algorithm [38,39] was chosen. The accuracy of the obtained classification was evaluated directly in GEE using the selected validation points in order to obtain a confusion matrix and the overall classification accuracy. The proposed method is synthetized in Fig.3.

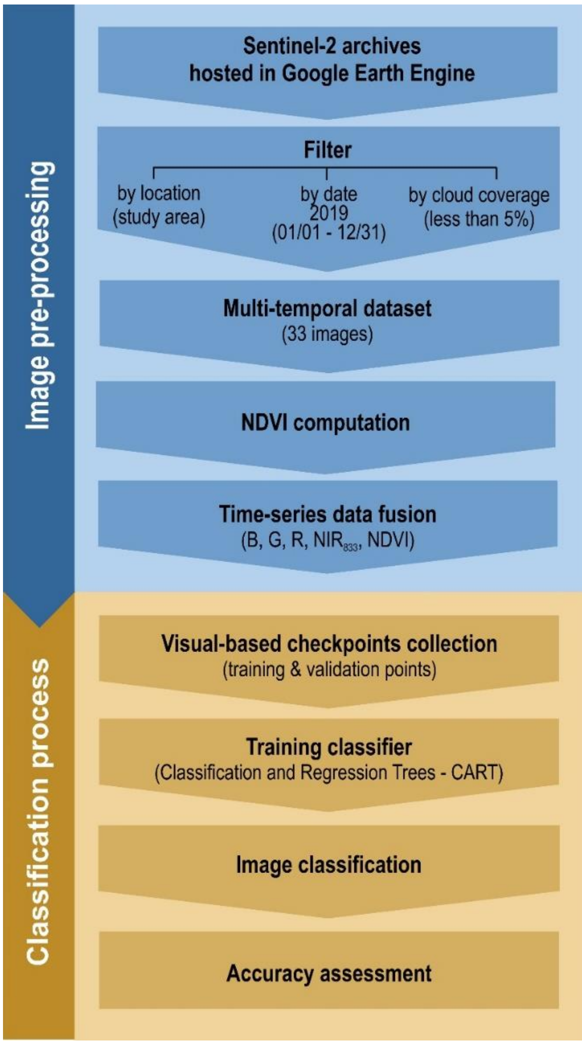

Fig. 3. Workflow of the presented method. The first part shows all the image processing steps used to obtain the composite image starting from a time-series of Sentinel-2 images. The second part shows the process steps to obtain a supervised classification and assess its accuracy by using Classification and regression Trees (CART) classifier.

3 Main results and conclusions

The proposed method allowed to obtain a multi-temporal data-cube of Sentinel-2 satellite images with the NDVI value. Of the over 100 images available for the study area and the analyzed year, only 33, all those presenting cloud coverage less than 5%, were used. These images were stacked becoming the data input for the classification process. The map of the classification obtained is presented in Fig.4, and shows the territorial distribution of the four classes considered in the analyzed scene.

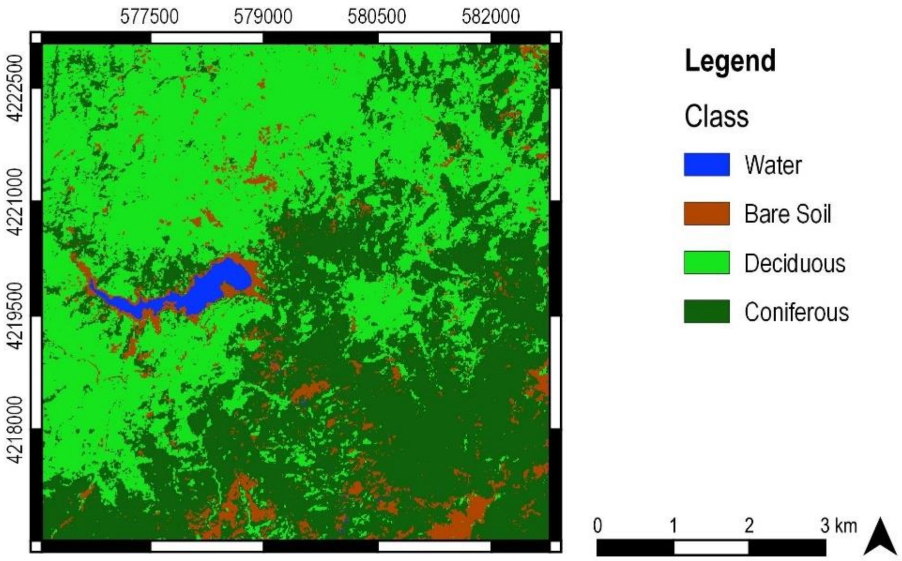

Fig. 4. Landscape classification map of the study area according to the four classes in Legend, derived from a Sentinel-2 time-series and multi-temporal NDVI analysis in Google Earth Engine environment

The confusion matrix generated for the 300 points (75 points for each class) used as validation points is reported in table1. Of the total points selected as validation for the water class, 69 were rightly assigned and 6 were wrongly assigned to bare soil class.

Table 1. Confusion matrix of the validation points.

	Water	Bare soil	Deciduous trees	Coniferous trees
Water	**69**	6	0	0
Bare soil	0	**67**	1	7
Deciduous trees	0	0	**75**	0
Coniferous trees	0	0	1	**74**

For the bare soil class, 67 points were correctly assigned and 8 points were misassigned to different classes (1 to deciduous trees and 7 to coniferous trees). The 75 points chosen for deciduous trees were all assigned correctly. For the coniferous trees validation points only one point was assigned to a wrong class (deciduous trees). The overall accuracy obtained is 0.95, meaning that only 5% of validation points (15 points) resulted wrongly assigned to a different class. First aim of this still ongoing research was to show the potential of joining spectral parameters with phenology information given by NDVI analysis. The high accuracy obtained at the end of the classification process highlights the good performance of the algorithms used in forestry vegetation mapping, using the NDVI value to deep distinguish deciduous and coniferous types. Moreover, this paper also highlights how the cloud computing platform GEE can work as a useful tool for multi-temporal vegetation analysis, thanks to its capability of using large satellite datasets and derived information. Moreover, GEE allows to share the written scripts with other researchers and stakeholders in order to improve code development and knowledge about the evolutionary dynamics in the studied area.

References

1. Di Fazio S, Modica G, Zoccali P. Evolution Trends of Land Use/Land Cover in a Mediterranean Forest Landscape in Italy. ICCSA 2011, Part I, Lecture Notes in Computer Science. 2011; 6782/2011:284–299.
2. Modica G, Merlino A, Solano F, Mercurio R. An index for the assessment of degraded Mediterranean forest ecosystems. Forest Systems. 2015; 24(3):e037.
3. Saganeiti L, Pilogallo A, Faruolo G, Scorza F. Territorial Fragmentation and Renewable Energy Source Plants : Which Relationship ? 2020;
4. Scorza F, Pilogallo A, Saganeiti L, Murgante B, Pontrandolfi P. Comparing the territorial performances of renewable energy sources' plants with an integrated ecosystem services loss assessment: A case study from the Basilicata region (Italy). Sustainable Cities and Society. 2020; 56(June 2019):102082. Available from: https://doi.org/10.1016/j.scs.2020.102082
5. Modica G, Pollino M, Solano F. Sentinel-2 Imagery for Mapping Cork Oak (Quercus suber L.) Distribution in Calabria (Italy): Capabilities and Quantitative Estimation. In 2019. p. 60–7.
6. Las Casas G, Scorza F, Murgante B. New Urban Agenda and Open Challenges for Urban and Regional Planning. In 2019. p. 282–8. Available from: http://link.springer.com/10.1007/978-3-319-92099-3_33
7. Lanucara S, Praticò S, Modica G. Harmonization and interoperable sharing of multi-temporal geospatial data of rural landscapes. Vol. 100, Smart Innovation, Systems and Technologies. 2019.
8. Giuliani C, Veisz AC, Piccinno M, Recanatesi F. Estimating vulnerability of water body using Sentinel-2 images and environmental modelling: the study case of Bracciano Lake (Italy). European Journal of Remote Sensing. 2019; 52(4):64–73. Available from: https://doi.org/10.1080/22797254.2019.1689796
9. De Luca G, Silva JMN, Cerasoli S, Araújo J, Campos J, Di Fazio S, et al. Object-based land cover classification of cork oak woodlands using UAV imagery and Orfeo Toolbox. Remote Sensing. 2019; 11(10).

10. Las Casas G, Scorza F, Murgante B. Razionalità a-priori: una proposta verso una pianificazione antifragile. Scienze Regionali. 2019; 2:9. Available from: http://www.rivisteweb.it/doi/10.14650/93656
11. Di Fazio S, Modica G. Historic rural landscapes: Sustainable planning strategies and action criteria. The Italian experience in the Global and European Context. Sustainability (Switzerland). 2018; 10(11):1–27.
12. Saganeiti L, Favale A, Pilogallo A, Scorza F, Murgante B. Assessing urban fragmentation at regional scale using sprinkling indexes. Sustainability (Switzerland). 2018; 10(9):1–23.
13. Modica G, Praticò S, Di Fazio S. Abandonment of traditional terraced landscape: A change detection approach (a case study in Costa Viola, Calabria, Italy). Land Degradation and Development. 2017; 28(8).
14. Modica G, Solano F, Merlino A, Di Fazio S, Barreca F, Laudari L, et al. Using Landsat 8 imagery in detecting cork oak (Quercus suber L.) woodlands: a case study in Calabria (Italy). Journal of Agricultural Engineering. 2016; 47(4):205.
15. Pontrandolfi P, Scorza F. Sustainable Urban Regeneration Policy Making: Inclusive Participation Practice. In: Gervasi O et al, editor. Computational Science and Its Applications -- ICCSA 2016 ICCSA 2016 Lecture Notes in Computer Science. Springer, Cham;2016.
16. Modica G, Pollino M, Lanucara S, La Porta L, Pellicone G, Di Fazio S, et al. Land Suitability Evaluation for Agro-forestry: Definition of a Web-Based Multi-Criteria Spatial Decision Support System (MC-SDSS): Preliminary Results. In 2016. p. 399–413. Available from: http://link.springer.com/10.1007/978-3-319-42111-7_31
17. Modica G, Praticò S, Pollino M, Di Fazio S. Geomatics in analysing the evolution of agricultural terraced landscapes. Vol. 8582 LNCS, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2014.
18. Vanden Borre J, Paelinckx D, Mücher CA, Kooistra L, Haest B, De Blust G, et al. Integrating remote sensing in Natura 2000 habitat monitoring: Prospects on the way forward. Journal for Nature Conservation. 2011; 19(2):116–25. Available from: http://dx.doi.org/10.1016/j.jnc.2010.07.003
19. Fraser RH, Olthof I, Pouliot D. Monitoring land cover change and ecological integrity in Canada's national parks. Remote Sensing of Environment. 2009; 113(7):1397–409. Available from: http://dx.doi.org/10.1016/j.rse.2008.06.019
20. Recanatesi F, Tolli M, Ripa MN, Pelorosso R, Gobattoni F, Leone A. Detection of Landscape patterns in airborne LIDAR data in the Nature Reserve of Castelporziano (Rome). Journal of Agricultural Engineering. 2013; 44(2s):472–7.
21. NOSSIN JJ. A Review of: " Remote Sensing, theorie en toepassingen van landobservatie (Remoie Sensing theory and applications of land observation "). Edited by H, J. BUITEN and J. G. P. W. CLEVERS. Series 'Dynamiek, indenting and bcheer van landelijke gebieden', part. International Journal of Remote Sensing. 1991; 12(10):2173–2173. Available from: https://www.tandfonline.com/doi/full/10.1080/01431169108955248
22. Weiers S, Bock M, Wissen M, Rossner G. Mapping and indicator approaches for the assessment of habitats at different scales using remote sensing and GIS methods.

Landscape and Urban Planning. 2004; 67(1–4):43–65.

23. Cheţan MA, Dornik A, Urdea P. Analysis of recent changes in natural habitat types in the Apuseni Mountains (Romania), using multi-temporal Landsat satellite imagery (1986–2015). Applied Geography. 2018; 97(June):161–75.
24. Bock M, Xofis P, Mitchley J, Rossner G, Wissen M. Object-oriented methods for habitat mapping at multiple scales - Case studies from Northern Germany and Wye Downs, UK. Journal for Nature Conservation. 2005; 13(2–3):75–89.
25. Díaz Varela RA, Ramil Rego P, Calvo Iglesias S, Muñoz Sobrino C. Automatic habitat classification methods based on satellite images: A practical assessment in the NW Iberia coastal mountains. Environmental Monitoring and Assessment. 2008; 144(1–3):229–50.
26. Wang Y, Mitchell BR, Nugranad-Marzilli J, Bonynge G, Zhou Y, Shriver G. Remote sensing of land-cover change and landscape context of the National Parks: A case study of the Northeast Temperate Network. Remote Sensing of Environment. 2009; 113(7):1453–61. Available from: http://dx.doi.org/10.1016/j.rse.2008.09.017
27. Pôças I, Cunha M, Pereira LS. Remote sensing based indicators of changes in a mountain rural landscape of Northeast Portugal. Applied Geography. 2011; 31(3):871–80. Available from: http://dx.doi.org/10.1016/j.apgeog.2011.01.014
28. Mairota P, Cafarelli B, Boccaccio L, Leronni V, Labadessa R, Kosmidou V, et al. Using landscape structure to develop quantitative baselines for protected area monitoring. Ecological Indicators. 2013; 33:82–95. Available from: http://dx.doi.org/10.1016/j.ecolind.2012.08.017
29. Lucas R, Blonda P, Bunting P, Jones G, Inglada J, Arias M, et al. The earth observation data for habitat monitoring (EODHaM) system. International Journal of Applied Earth Observation and Geoinformation. 2015; 37:17–28.
30. Vogelmann JE, Xian G, Homer C, Tolk B. Monitoring gradual ecosystem change using Landsat time series analyses: Case studies in selected forest and rangeland ecosystems. Remote Sensing of Environment. 2012; 122:92–105. Available from: http://dx.doi.org/10.1016/j.rse.2011.06.027
31. Simonetti D, Simonetti E, Szantoi Z, Lupi A, Eva HD. First Results from the Phenology-Based Synthesis Classifier Using Landsat 8 Imagery. IEEE Geoscience and Remote Sensing Letters. 2015; 12(7):1496–500.
32. Recanatesi F, Giuliani C, Ripa MN. Monitoring mediterranean oak decline in a peri-urban protected area using the NDVI and sentinel-2 images: The Case Study of Castelporziano state natural reserve. Sustainability (Switzerland). 2018; 10(9).
33. Zhang M, Gong P, Qi S, Liu C, Xiong T. Mapping bamboo with regional phenological characteristics derived from dense Landsat time series using Google Earth Engine. International Journal of Remote Sensing. 2019; 00(00):1–15. Available from: https://doi.org/10.1080/01431161.2019.1633702
34. Lu D, Weng Q. A survey of image classification methods and techniques for improving classification performance. International Journal of Remote Sensing. 2007; 28(5):823–70.
35. Gorelick N, Hancher M, Dixon M, Ilyushchenko S, Thau D, Moore R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment. 2017; 202:18–27. Available from:

https://doi.org/10.1016/j.rse.2017.06.031

36. Crowley MA, Cardille JA, White JC, Wulder MA. Generating intra-year metrics of wildfire progression using multiple open-access satellite data streams. Remote Sensing of Environment. 2019; 232(July):111295. Available from: https://doi.org/10.1016/j.rse.2019.111295
37. Rouse JW, Hass RH, Schell JA, Deering DW. Monitoring vegetation systems in the great plains with ERTS. Third Earth Resources Technology Satellite (ERTS) symposium. 1973; 1:309–17. Available from: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19740022614.pdf
38. Loh W-Y. Classification and regression trees. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. 2011; 1(1):14–23. Available from: http://doi.wiley.com/10.1002/widm.8
39. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning [Internet]. New York, NY: Springer New York;2009. (Springer Series in Statistics). Available from: http://link.springer.com/10.1007/978-0-387-84858-7