Analysis of Dynamics of Land Cover Types Using Land Sat8: A Case Study Anantapur

V. Vandana[1], Mule Abhi Roop[2], Dr. D. Gowri Sankar Reddy[3]
[1]*M. Tech scholar, Department of ECE, SVU College of Engineering, S. V. University, Tirupati, AP, India*
[2]*Research scholar, Department of ECE, SVU College of Engineering, S. V. University, Tirupati, AP, India*
[3]*Associate Professor, Department of ECE, SVU College of Engineering, S. V. University, Tirupati, AP, India*
doi.org/10.64643/IJIRTV12I7-188499-459

***Abstract*-One of the grave issues in dry areas such as Anantapur is drought which impacts crops, water and livelihood of people. In this study, the author employs the Landsat 8 satellite image of 2016 and 2024 to compare the drought conditions via the remote sensing techniques. Four indices were computed; NDWI of water, NDVI of vegetation, NDSI of soil, and LST of land surface temperature. These indices assist in the realization of the alterations of water, vegetation, soil as well as heat within the years. Classification of land features was performed with the help of Support Vector Machine (SVM) and Principal Component Analysis (PCA) was applied together with SVM in order to enhance the accuracy. The findings revealed that SVM was a good performer and PCA-SVM was even a better performer, thus indicating that PCA-SVM is more dependable. Relating the years 2016 and 2024, the research noted variations in water supply, plant health, the status of the soils, and surface temperature that could assist in determining the regions under drought over the years. Such an approach would help in planning, drought control and management of resources in the region of Anantapur better.**

***Index Terms*- Landsat 8, Anantapur Region, Remote Sensing, GIS, Digital Image Processing, QGIS Software, Drought Analysis, Water Index, Vegetation Index, Soil Index, Land Surface Temperature, Machine Learning, Support Vector Machine (SVM), Principal Component Analysis (PCA), PCA–SVM, Land Cover Classification.**

I. INTRODUCTION

Drought is among the worst natural catastrophes, which attacks crops, water supply, and lives of people, particularly in semi-arid regions. Drought is very likely to occur in the Anantapur area of Andhra Pradesh in India due to low rainfall, high temperature and overreliance on farming. In order to understand and cope with drought, water, vegetation, soil, and land surface temperature need to be monitored on a regular basis [2,4,6].

Remote sensing, Digital Image Processing, and GIS are the effective tools which have been applied to study drought due to their ability to cover wide areas, ability to repeat observations and give precise information on the condition of the land. Various indices could be determined with the help of satellite images, including the Normalized Difference Water Index (NDWI) of water bodies [9], the Normalized Difference Vegetation Index (NDVI) of vegetation health [7], the Normalized Difference Soil Index (NDSI) of the soil condition, and Land Surface Temperature (LST) of the surface heat [1,8]. Such indices come in handy to understand drought conditions in a systematic manner. QGIS software is usually applied to the processing of satellite images, to the digital image processing tasks, to some index calculation as well as to the land cover classification [10], and to the generation of thematic maps. Image classification is also significant in the use of Machine Learning techniques. The most common supervised learning algorithm of land cover classification is the Support Vector Machine (SVM) among them [5]. Principal Component Analysis (PCA) is a statistical method of eliminating redundant information and enhancing performance. PCA is used together with SVM to create PCA SVM, which is more accurate in classification and even results in better results than SVM by itself.

The question of the study utilizes the Landsat 8 satellite images of the year 2016 and 2024 in the analysis of the drought in the Anantapur region. Its key tasks include computation of drought-related indices (water index, vegetation index, soil index, and land surface temperature) with the help of QGIS and Digital Image Processing and land cover features classification with the help of Machine Learning methods (SVM and PCA-SVM) and the comparison of the results by year. The results of this paper will be useful in the drought trends and will aid in the improved water management, land use planning and drought mitigation policies [1, 4].

The processing of Landsat images to be used in this study was done with the help of QGIS version 3.40 which is a free GIS software. Semi-Automatic Classification Plugin (SCP) was used in the analysis of all remote sensing and image processing operations, which included satellite image preprocessing, ROI extraction, raster operations, as well as classification. OpenStreetMap shapefiles were included to map ROI accurately. The QGIS coupled with SCP offered a complete method of clipping, classifying, and creating derivative datasets to enable the analysis of the study area in detail.

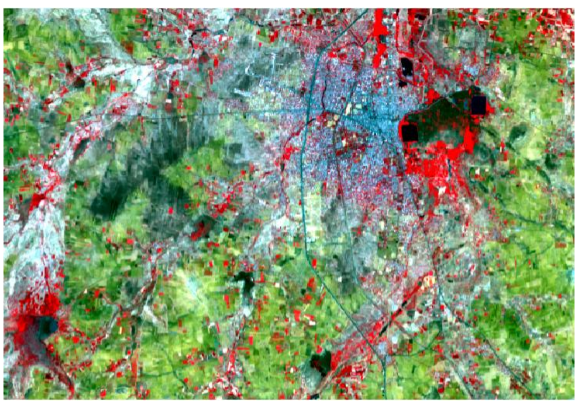

Fig.1. Input Image

1.1 Materials

The current research is based on a particular region in the city of Anantapur, the Rayalaseema district, Andhra Pradesh, India, and that area is between the rivers Bukkarayi Samudram and Prasanayapalle as well as between Brahmanapalle and Brahmanayaleru. The region has a semi-arid climatic region with mostly agricultural land, sparse vegetation and scattered settlements. All satellite data was cut to this ROI with a spatial resolution of 30 meters per pixel and thus it gave a calculated 2.78 million pixels to analyze. Official shapefiles acquired on open street map were used to delineate the ROI.

II. METHODS

2.1 Land Cover Classification Using SVM and PCA–SVM

In this work, two methods were used in order to perform the land cover of the Anantapur region; the Support Vector Machine (SVM) classifier and the Principal Component Analysis-Support Vector Machine (PCA-SVM) model. The changes in the land cover patterns between 2016 and 2024 were analyzed and compared using Landsat 8 Operational Land Imager (OLI) images of land cover patterns in the years 2016 and 2024 [5]. In the case of the classification based on SVM, training samples of the key land cover, i.e., vegetation, water bodies, soil, and barren land, were selected with care in the satellite imagery. SVM classifier was subsequently used in order to differentiate these classes by finding the best hyperplane that would maximize the distance between the various land cover types.

The spectral bands were firstly subjected to Principal Component Analysis in the PCA-SVM method to reduce dimensionality and remove redundant information, yet maintain the largest amount of variability in the data set. The obtained principal components were then taken as input features into the SVM classifier. Such a combination of PCA allows increased computational efficiency, eliminates noise and classification separability between classes, allowing more robust and accurate classification. Using both PCA-SVM and SVM methods on datasets with two years of data, the research delivers a thorough evaluation of variation in land cover over time and proves that dimensionality reduction is useful in enhancing the capability of classifications.

2.2 Support Vector Machine (SVM)

A very strong algorithm in supervised machine learning that has been heavily applied in remote sensing is Support Vector Machine (SVM) [5] used to classify land cover in a very accurate manner. In this research, the sample representative training samples of each category of land covers including vegetation, water bodies, soil, and barren land were carefully cut off the satellite images to give credible discrimination of classes. With these training data, SVM classifier offered by the QGIS was used to generate detailed and accurate classified maps of the study area.

SVM works by finding the best hyperplane which best separates various categories in high-dimensional feature space. This hyperplane is selected to make maximum margin--the distance between the decision boundary and the closest data points of each of the two classes--this will enhance the strength of the classifier and its generalization capability. Moreover, SVM uses kernel functions (linear, polynomial, or RBF kernels) in order to address non-linear and complicated boundaries of classes that are prevalent in remote sensing data. Consequently, the algorithm has high classification accuracy and is especially useful in multispectral image processing in which spectral signature overlap exists.

2.3 Principal Component Analysis (PCA)

A common statistical technique of dimensionality reduction and feature extraction is Principal Component Analysis (PCA), which is typically used in image processing as well as remote sensing [6]. Satellite data can have several spectral bands which have overlapping or redundant data. PCA is a solution to this problem, which converts a set of the original correlated spectral variables into a different set of uncorrelated variables termed as principal components. These elements are ranked in the manner that the initial few hold the largest amount of variance that can exist in the data that they are able to capture the largest amount of patterns and variability that can exist in the imagery.

With high-dimensional data, PCA eliminates redundancy and compresses it into fewer significant components, and reduces the noise, and improves the interpretability of spectral characteristics. This does not only enhance the visualization and analysis of data, but also saves a lot of computation time in classification. The major features obtained using PCA are quality inputs of the machine learning algorithms. With the addition of the classifiers like Support Vector Machines (SVM), the resultant PCA-SVM model is more efficient and has the ability to give correct land cover maps. Therefore, PCA is clearly important in optimizing remote sensing processes as it automatically facilitates more vigorous, discriminative and computationally efficient categorization of sophisticated satellite data.

2.4 Block Diagram

The block diagram of the proposed methodology is shown in figure 3. The process is initiated with the Data Acquisition then the necessary Pre-Processing to prepare the satellite imagery takes place. Selection of ROI is then done to isolate the study area. Then Feature Extraction is implemented to obtain significant spectral and spatial data. Lastly, the data is categorised under the chosen Classification algorithm to produce the final output thematically.

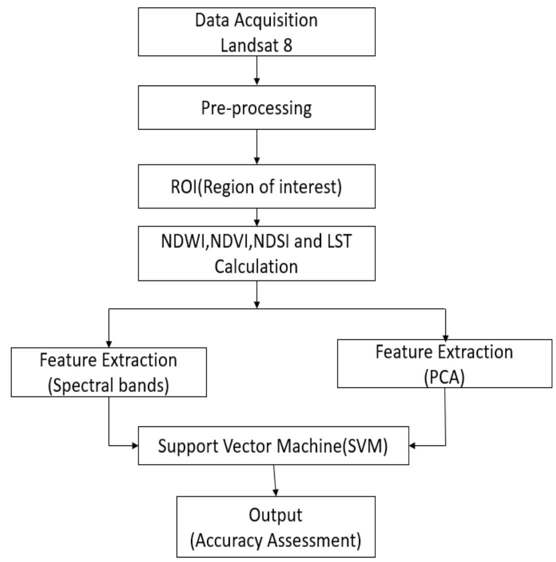

Fig.2. Block diagram of the proposed method2.4.1 Input Image (Landsat 8)

The data used was the Landsat 8 satellite views which were obtained off the Earth Explorer site of the United States Geological Survey (USGS). This is the main input data that was taken to perform drought analysis in the Anantapur district.

2.4.2 Pre-processing

Landsat 8 images were pre-processed to undergo radiometric correction and transform raw DN into surface reflectance and geometric correction to orient the imagery to ground coordinates to be sure of quality data to further analyze.

2.4.2.1 Radiometric Correction:

- Dark Object Subtraction (DOS): Eliminates the effect of the atmosphere by exclusion of the lowest DN values of dark objects.

- Top-of-Atmosphere (TOA) Reflectance: Normalise the solar energy by converting DN values to reflectance.

2.4.2.2 Geometric Correction

- Orthorectification: Corrects geometric distortions using a Digital Elevation Model (DEM) and satellite geometry to produce georeferenced images.

2.4.3 Region of Interest (ROI):

The specific study area, such as the Anantapur region, was extracted from the full satellite image to focus the analysis on the area of interest.

2.4.3.1 Water Calculation: Using indices like NDWI, the water content or wetness is calculated for the selected area.

2.4.3.2 Vegetation Calculation: Using indices like NDVI, the vegetation health and coverage are calculated for the selected area.

2.4.3.3 Soil Calculation: Using indices like NDSI, the soil condition is calculated for the selected area.

2.4.3.4 Land Surface Temperature (LST) Calculation: Using thermal bands, the land surface temperature is calculated for the selected area. the Ground Surface Temperature (GST), also referred to as Land Surface Temperature (LST), was calculated from Landsat 8 Band 10 (Thermal Infrared band) using the Raster Calculator in QGIS. The process was carried out in the following steps: The raw satellite image values (DN) were first converted into spectral radiance using:

Step1: Convert Radiance to Brightness Temperature (BT)

Radiance values were converted into Brightness Temperature (BT) in Kelvin:

$$BT = \frac{\mathrm{K}_2}{\ln\left(\frac{\mathrm{K}_1}{L\lambda} + 1\right)}$$

Step2: Estimate Surface Emissivity (ε):

$$PV = \left(\frac{NDVI - \mathrm{NDVI}_{\min}}{\mathrm{NDVI}_{\max} - \mathrm{NDVI}_{\min}}\right)^2$$

$$\varepsilon = 0.004 \times \mathrm{P}\upsilon + 0.986$$

different land covers (soil, vegetation, water) emit heat differently, surface emissivity (ε) was calculated using NDVI through the Proportion of Vegetation

Step 3: Calculate Final Land Surface Temperature (LST):

Finally, the corrected Land Surface Temperature (LST) in Celsius was obtained using:

$$LST = \frac{BT}{1 + \left(\frac{\lambda . BT}{\rho}\right) . \ ln\ \varepsilon} - 273.15$$

2.4.4 Support Vector Machine (SVM): Land cover of the study area was classified using the Support Vector Machine (SVM) technique with the Radial Basis Function (RBF) kernel. This method separates different land cover types by mapping data into a higher-dimensional space and finding the optimal hyperplane for classification.

2.4.5 Principal Component Analysis (PCA-SVM):

The first step is the Principal Component Analysis (PCA) which helps to reduce the size of the data and to identify the most important features. Support Vector machine (SVM) with Radial Basis Function (RBF) kernel is then used to classify the transformed features to distinguish the various land cover features within the study area.

III. RESULTS

The study of Landsat 8 images of 2016 and 2024 showed significant changes in the study area. NDWI showed that there were changes in the water availability, NDVI reflected changes in the health and the vegetation coverage, NDSI reflected changes in the soil condition, and LST images reflected changes in space in the surface temperature over time. The classification of land cover by SVM and PCA-SVM revealed that integrating PCA with SVM was better in regard to the classification results than the result of using SVM. The results can be used to locate the regions where the drought occurs and aid in proper planning and management of resources in Anantapur region.

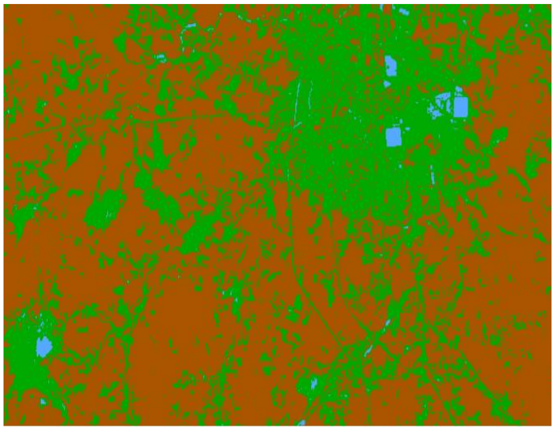

Fig.3. SVM-2016

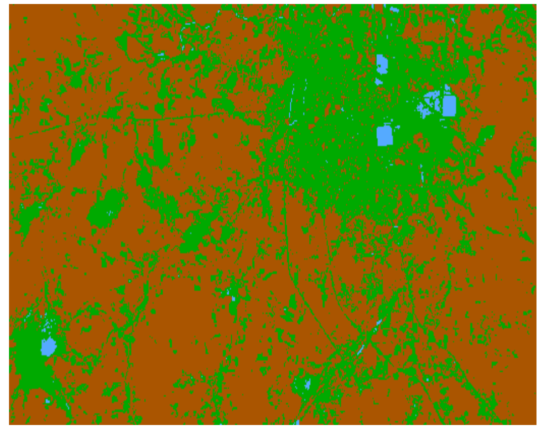

Fig.4. PCA-SVM-2016

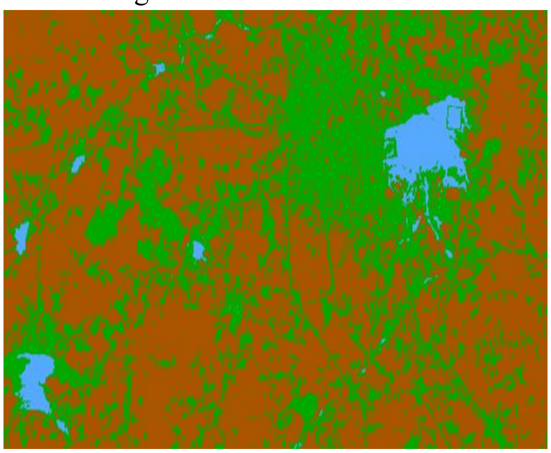

Fig.5. SVM-2024

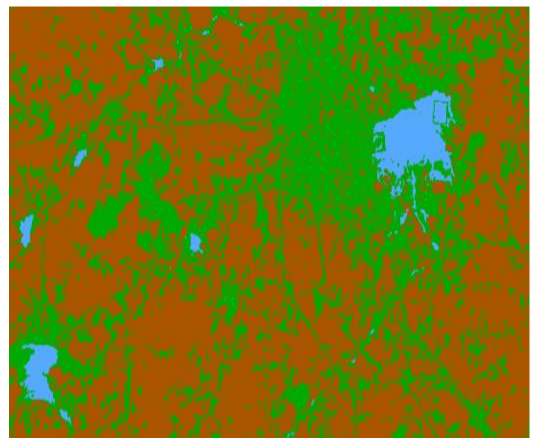

Fig.6. PCA-SVM-2024

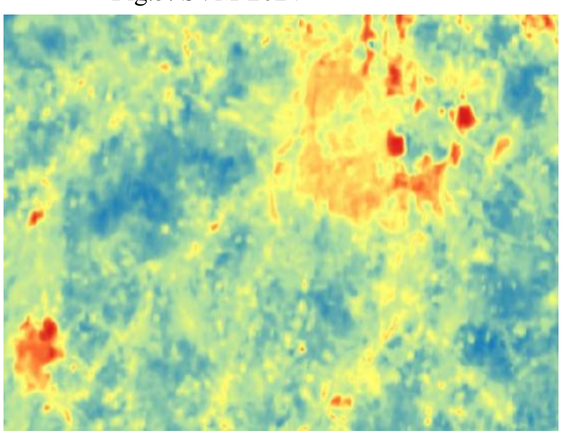

Fig.7. LST-2016

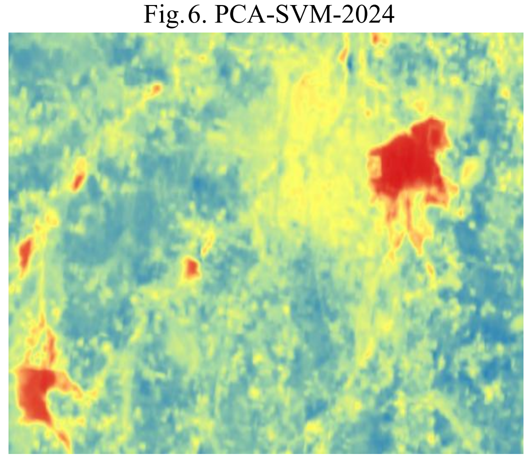

Fig.8. LST-2024

Algorithm	Overall Accuracy	UA (User's Accuracy)	PA (Producer's Accuracy)	Kappa Coefficient
SVM	97.93	94.28	100	0.95
SVM-PCA	98.79	98.08	98.60	0.97

Table.1. Accuracy Assessment (2016)

Algorithm	Overall Accuracy	UA (User's Accuracy)	PA (Producer's Accuracy)	Kappa Coefficient
SVM	98.19	97.32	98.31	0.95
SVM-PCA	98.93	99.35	98.21	0.98

Table.2. Accuracy Assessment (2024)

Year	LST (Land Surface Temperature) in Degree Celsius	
	Max (Temp)	Min (Temp)
2016	47.440155	30.5104637
2024	49.5316963	28.6913738

Table.3. Land Surface Temperature

Indices	SVM-2016	PCA-SVM-2016
Water	1.868	1.866
Vegetation	74.29	72.27
Soil	145.96	147.98

Table.4. Land Cover Types Area (Km^2)- 2016

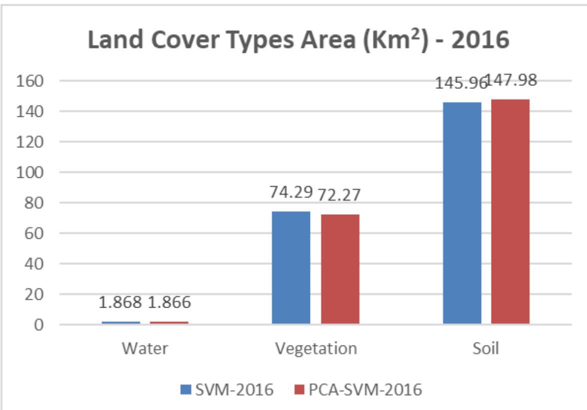

Fig.9. Land Cover Types Area (Km^2)- 2016

Indices	SVM-2024	PCA-SVM-2024
Water	7.453	6.874
Vegetation	106.31	75.632
Soil	108.35	139.61

Table.5. Land Cover Types Area (Km^2)- 2024

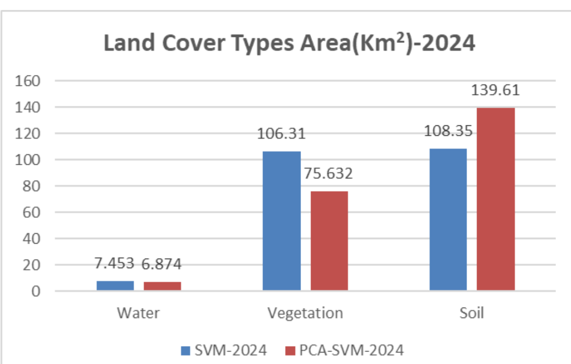

Fig.10. Land Cover Types Area (Km^2)- 2024IV.

IV. CONCLUSION

In this research, multi-temporal Landsat 8 images were used to evaluate the drought changes in the area of Anantapur. The calculated indices (NDWI, NDVI, NDSI, and LST) have shown significant changes in water availability, the state of vegetation, the exposure of soil and surface temperature in different years of the research. The classification of land cover with the use of SVM and PCA–SVM revealed that PCA–SVM had more accuracy which is why it is effective in working with the features of dimensionality reduction. Comprehensively, remote sensing data and machine learning methods were effective in drought monitoring and management of resources in drought-prone areas.

V. FUTURE SCOPE

The next generation of work can be dedicated to the research of satellite imaging of higher quality, which includes Sentinel-2 or commercial data, to enhance the spatial resolution of the drought assessment in the Anantapur area. Further investigation of advanced machine learning and deep learning algorithms, such as a Random Forest, CNNs, and hybrid classifiers, could be carried out to increase the accuracy of classification. Drought can be analyzed by including other datasets, like soil moisture sensors, rainfall records, ground-based observations, etc. The early warning systems of the sensitive regions can be supported with time-series modeling and predictive drought forecasting. The incorporation of socio-economic information and validation of the information at the field will also aid in comprehending the overall effects of drought in the district of Anantapur.

REFERENCES

[1] M. Mansourmoghaddam, M. Yaghoobian, and H. Zolfaghari, “Modeling and estimating land surface temperature using machine learning and Landsat-8 data,” *Remote Sensing*, vol. 16, no. 3, art. 454, 2024.

[2] C. B. Pande, S. Singh, and B. Shukla, “Characterizing land use/land cover change dynamics using remote sensing and GIS,” *Environmental Sciences Europe*, vol. 35, no. 102, 2023.

[3] F. Marcello et al., “Assessment of urban heat island using Landsat-8 thermal bands and NDVI-based emissivity,” *Sustainable Cities and Society*, vol. 84, 2022.

[4] S. Kafy, M. R. Hasan, and Z. Sultana, “Spatiotemporal drought assessment using NDVI

and LST from Landsat imagery," *Environmental Monitoring and Assessment*, vol. 193, no. 11, 2021.

[5] W. Guo, Q. Wang, and F. Li, "Land cover change detection based on Landsat time-series using random forest classifier," *International Journal of Remote Sensing*, vol. 41, no. 7, 2020.

[6] H. Xu, "Analysis of vegetation dynamics using NDVI time-series and Landsat data," *Journal of Applied Remote Sensing*, vol. 13, no. 1, 2019.

[7] A. Huete and K. Didan, "Vegetation monitoring with NDVI and EVI: Improved methods for Landsat and MODIS," *Remote Sensing of Environment*, vol. 204, pp. 292–305, 2018.

[8] E. Sekertekin and N. Bonafoni, "Land surface temperature retrieval from Landsat 8 thermal infrared sensor using mono-window algorithm," *Remote Sensing*, vol. 9, no. 7, 2017.

[9] R. Yadav and S. Chouhan, "Monitoring water bodies using NDWI from Landsat OLI imagery," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 7, 2016.

[10] M. A. Wulder et al., "Landsat 8 operational land imager data: Surface reflectance products for environmental monitoring," *Remote Sensing of Environment*, vol. 170, pp. 113–127, 2015.