

# **CHAPTER 3: EVALUATION OF SIMULATED AVIRIS-NG IMAGERY USING A SPECTRAL RECONSTRUCTION METHOD FOR THE RETRIEVAL OF LEAF CHLOROPHYLL CONTENT**

---

## **3.1 INTRODUCTION**

Remote sensing has become significant in earth science, observation, research, and solving environmental problems. Many satellites and UAV are available now to help us understand the Earth's environments. Sensors collect data ranging from visible to microwave frequencies in the electromagnetic spectrum. A wide range of sensors, including multispectral sensors (MSS) and hyperspectral sensors (HSS), give information in different spectral resolutions. However, MSS is limited in use because of its lower spectral resolution than HSS. The former gives information approximately in ten to fifteen spectral bands, whereas the later provides information in the order of hundreds of bands. Hyperspectral data has far more information than multispectral data and has dramatically expanded the variety of remote sensing applications [35]. HS imagery provides detailed information due to the narrow bandwidth but is scarcely available due to the high data acquisition cost. At present, there are only a few space-borne hyperspectral sensors available. On the other hand, widely available multispectral data can be used to simulate or construct hyperspectral data, particularly in situations where the latter are required but difficult to obtain.

There is extensive use of data simulation in the remote sensing field to generate pictures for virtual or novel sensors in the design stage. Simulated data can also be used to assess spatial properties important to a specific study and that image is

compared with the original image obtained from a specific sensor [36]. Data simulation is frequently used to assess the impact of various spectral or spatial resolutions in specific applications to determine the best resolution for a given problem [37, 38]. Some researchers assessed the applicability of linear spectral unmixing to different degrees of spatial resolution by simulating HS data with varied spatial resolution. In contrast, others assessed the impact of spatial and spectral resolution on vegetation categorization. Furthermore, this method has found its usage in assessing and testing new algorithms in HS remote sensing, such as target detection and identification [39].

In the current study, the interest of vegetation parameter is LCC, a biochemical parameter that gives information about the pigment level and individual vegetation's nutrition. Thus, creating a thorough and accurate picture of crop development status requires a quantitative estimation of crop LCC. With the change in stages of the vegetation, whenever exposed to stresses, the LCC value also changes. Change in pigment level determines the vegetation health. As a result, determining its level is required to compute plant stress and nutritional status to maximize crop yield [8, 40]. Traditional approaches for estimating LCC are costly, time-consuming, and damaging. Hyperspectral remote sensing provides a non-destructive and scalable monitoring strategy. Thus, investigating vegetation's spectra using hyperspectral remote sensing yields promising findings for retrieving the biochemical parameter [41].

Various approaches are used to estimate the biochemical composition of the vegetation. The most straightforward method is to directly build a statistical relation between in situ parameter's value, and canopy reflectance obtained using aerial or satellite sensors [42-44]. The most extensively used semi-empirical methods for

estimating LCC are vegetation indices-based and multivariate statistical approaches [45]. The vegetation index (VI) is a linear or nonlinear combination of reflectance at different wave-lengths used to indicate information about vegetation. VIs exploit the fact that a ratio of a band sensitive to a parameter of interest to a band insensitive to that parameter of the electro-magnetic spectrum gives ample information about it. The benefit of utilizing VIs is that they improve spectral information by identifying variations caused by physiological and morphological factors [46]. Physical modelling for estimating leaf biochemical contents has made significant progress. Numerical inversion of Radiative transfer (RT) models like PROSPECT, LEAFMOD, etc., have been used to predict LCC [47-50]. Several researchers have also worked with enhanced learning algorithms combined with VIs for the retrieval of vegetation parameters in order to avoid the mathematical complexity of physical models [24, 26, 51] .

UPDM is a spectral reconstruction approach based on the spectral unmixing method; it is a modified version of the pattern decomposition method [52]. UPDM is a sensor-independent technique in which each satellite pixel is expressed as the linear sum of fixed, standard spectral patterns for significant classes, and the same normalized spectral patterns can be used for different solar-reflected spectral satellite sensors. There have been studies where UPDM has been used to generate different HSS data using MS data. Zhang et al. suggested a spectral response technique for hyperspectral data simulation from Landsat 7 Enhanced Thematic Mapper Plus (ETM+) and Moderate Resolution Imaging Spectroradiometer (MODIS) data that employed the UPDM. Liu et al. used the same method to recreate 106 hyperspectral bands from the EO-1 Advance Land Imager (ALI) multispectral bands using normal ground spectra of water, vegetation, and soil. Badola et al. used UPDM to reconstruct Airborne

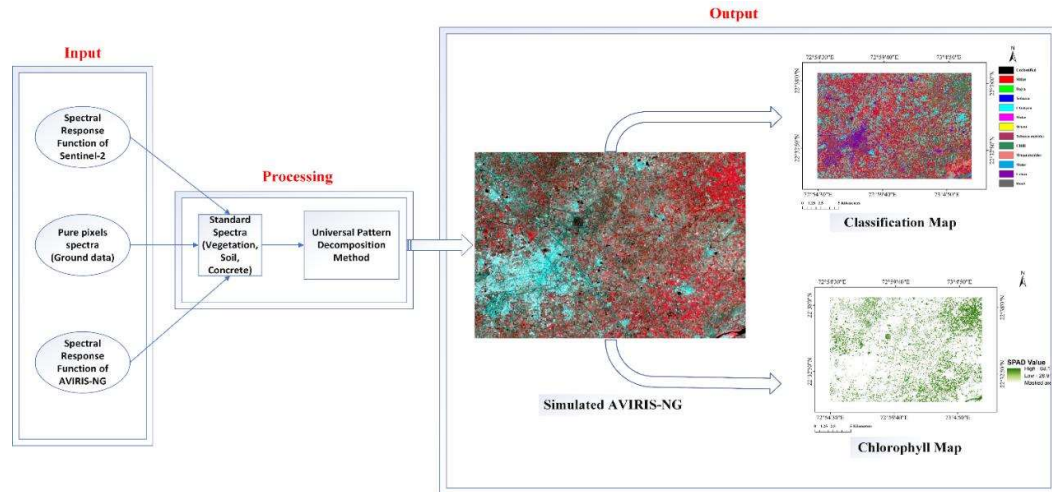
Visible-Infrared Imaging Spectrometer - Next Generation (AVIRIS-NG) imagery from Sentinel-2 for improved wildfire fuel mapping in Boreal, Alaska. The spectral reconstruction technique has also found its application in the mineral mapping [39, 53, 54]. Muramatsu et al. [52] used UPDM for vegetation type mapping from ALOS/AVNIR-2 AND PRISM DATA. Till now, UPDM has proved its usefulness in various fields for mapping and classification purposes; however, the efficiency of reconstructed HS imagery using UPDM for retrieving vegetation parameters with application in agriculture has yet to be explored.

In the presented work, we used UPDM to simulate AVIRIS-NG imagery using Sentinel-2 for a region in the Gujarat district that contains mainly agricultural areas to expand its use in the retrieval of vegetation parameters. The simulated data is evaluated visually, statistically, and further verified for the LCC retrieval of various crops in different vegetative stages.

## **3.2. MATERIALS AND METHODOLOGY**

### **3.2.1 Study area and data acquisition**

The detail about the study site for this work is mentioned in Chapter 2, under Section 2.3.1. Figure 2.8 shows the map of the district and strip of area on which ground data sampling was held to overlap with the aerial image acquisition by AVIRIS-NG, from 26-28 March, 2018. Figure 3.1 shows the workflow of the work.



**Figure 3.1.** Workflow of the study

### 3.2.2 Remotely sensed and ground data

#### 3.2.2.1 Hyperspectral data

The second AVIRIS-NG Indian campaign was held from 26–28 March 2018. The collected level-2 surface reflectance AVIRIS-NG scene was downloaded from the VEDAS-SAC (source: <https://vedas.sac.gov.in/aviris/>) geoportal assessed on 8 September 2018. The downloaded scene includes the agricultural area of the Anand district, Gujarat. AVIRIS-NG mission is a collaboration of NASA and ISRO. Data can be downloaded from the ‘VEDAS’ site in both radiance and reflectance format. Due to the presence of various types of atmosphere aerosols, it makes the data vulnerable to distortions and need to be pre-processed using various atmospheric correction methods. Presented AVIRIS-NG data were pre-processed, and current L1 products were used (radiance to reflectance).

#### 3.2.2.1 Multispectral data

Sentinel-2 images acquired on 28 March 2018 for the study site were downloaded from the Copernicus open access hub of ESA (source: <https://scihub.copernicus.eu/dhus/#/home>) and downloaded on 16 October 2021. The original L1C (top-of-atmosphere) product was converted to L2A (bottom-of-atmosphere) reflectance using the Sen2cor processor. Sentinel-2 gives information in thirteen bands, as mentioned in Chapter 2, Section 2.1.2 and bands specification are shown in Table 2.1. For this study, bands B1 and B10 were removed, and the remaining bands were resampled to a pixel size of 10 m; all bands were stacked, and the study area was clipped. The analysis of multispectral imagery was performed in SNAP 8.0.

### **3.2.2.3 Ground data**

The pure spectra for various vegetation were used from the spectral library generated during the field campaign i.e., 26–28 March 2018. In the sampling, spectra of different species were collected at different vegetative stages of the agricultural species present in the area such as tobacco, chickpea, maize, wheat, millet, etc. The field spectra were taken with a high-resolution ASD FieldSpec 4 spectroradiometer. The leaf chlorophyll content was collected from SPAD during the field campaign. Extensive sampling was performed to collect data such as stage, soil condition, chlorophyll content, and plant height, coinciding with the AVIRIS-NG flight overpass dates.

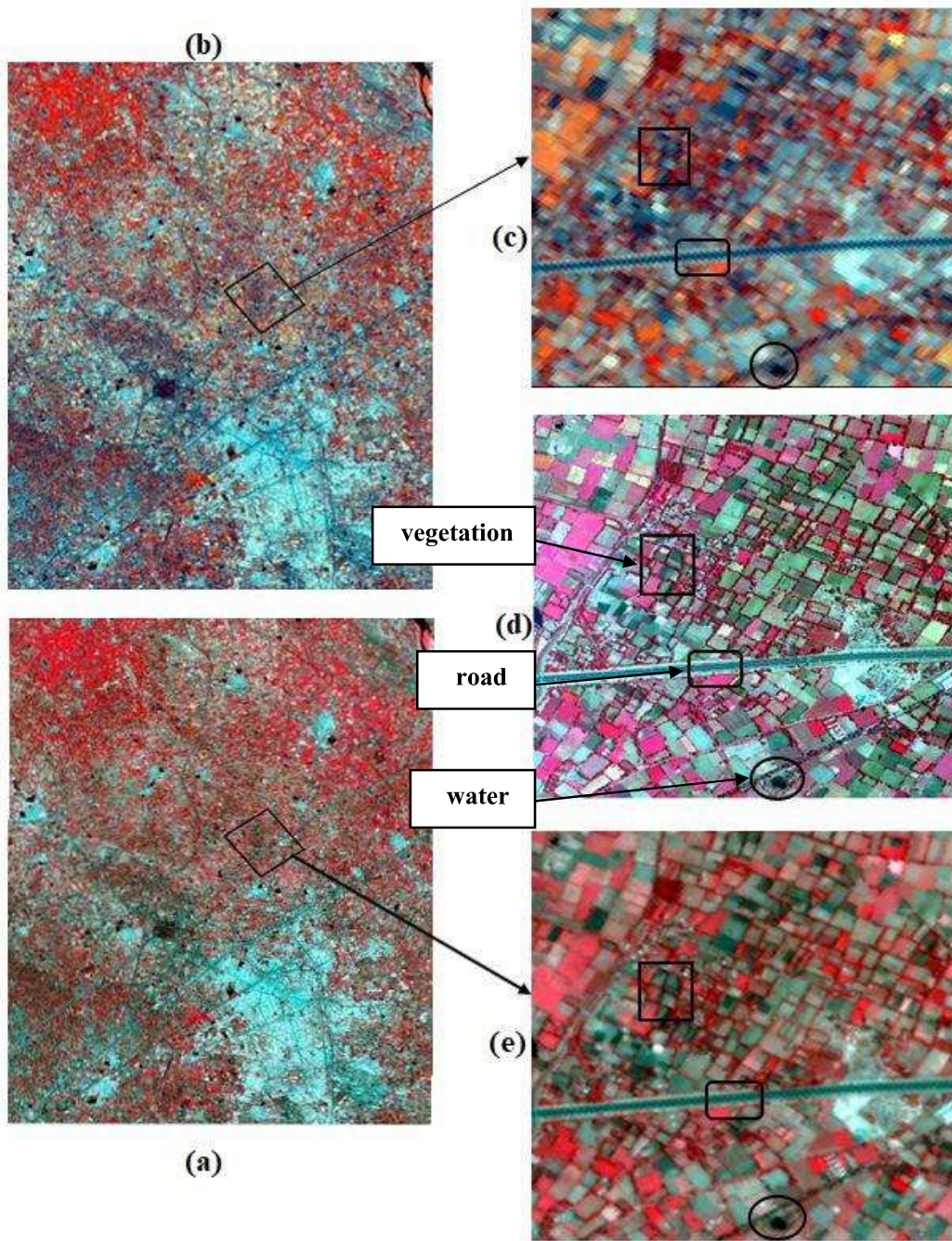
### **3.2.3 Simulating AVIRIS-NG from Sentinel-2 using UPDM**

#### **3.2.3.1 Computing standard spectra**

The reflectance from a pixel's FOV gives the combined spectra of classes in the pixel. The study site contained mostly agricultural fields, harvested land, bare soil, and some

urban areas, as shown in Figure 3.2. We chose vegetation, bare soil, and urban area as three primary classes/endmembers to be given to the algorithm. The spectra of vegetation and soil were taken from the spectral library generated during field sampling. For urban area, a spectrum of concrete was downloaded from the USGS library and used in the process of simulation. The spectral resolution and wavelength of the concrete spectrum, was same as that of the spectral resolution and wavelength range of vegetation and soil spectra

The standard spectra for each class, i.e., vegetation, soil, and concrete, were calculated by convolving the pure spectra with the spectral response function (SRF) of the sensors to normalize them. The SRF for Sentinel-2 was downloaded from the ESA document library, whereas the SRF for AVIRIS-NG was calculated. In UPDM, since there are three endmembers to represent the whole image, there is some loss of information. Because in terms of spectral resolution as high as 5nm there is definitely going to be some information loss. For example, in different crops, the identification of species is primarily done from red edge inflection point, which vary minutely. Whole vegetation species is represented by one vegetation spectra, although the reflection from multispectral image's pixel adjusts the reflection of newly generated image's pixel, so we get much better image compared to MSI in terms of spectral resolution. Other than spectral resolution, the spatial resolution of newly generated image is similar to the resolution of MSI. So, the idea is to generate an aerial image with larger coverage similar to satellite image for practical purpose. For example, the coverage of sentinel-2 is around 300km whereas for the AVIRIS-NG it is 10-100km depending on the flight height. In the work, an attempt has been made to combine the best of both images (large coverage of MSI and spectral resolution of HSI) to generate an image for parameters estimation purpose.



**Figure 3.2.** (a) Simulated AVIRIS-NG (bands 97, 56, and 36); (b) Sentinel-2 imagery (bands 8, 4, and 3); (c) clipped area for Sentinel-2 imagery; (d) AVIRIS-NG imagery (bands 97, 56, and 36); (e) clipped area for simulated AVIRIS-NG imagery.

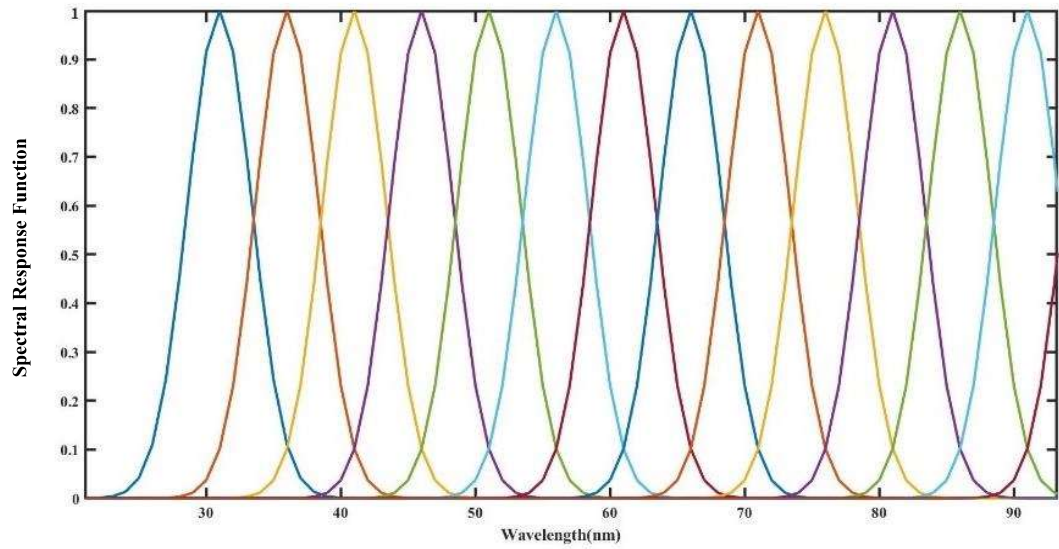
The full width at half maximum (FWHM) values for each band are available; in order to calculate SRF for AVIRIS-NG sensors, a Gaussian function ( $g$ ) for the  $i^{\text{th}}$  band was calculated as follows:

$$g(\bar{\lambda}_i, \sigma_i) = \exp^{-\frac{(\bar{\lambda}_i - \lambda)^2}{2\sigma_i^2}} \quad (3.1)$$

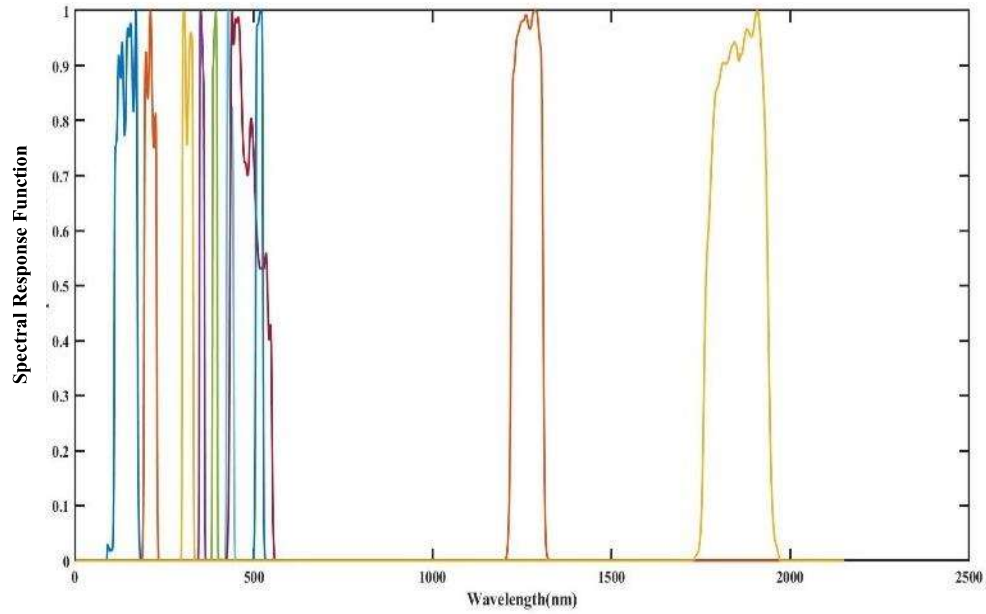
The FWHM of each band was used to get the bandwidth.

$$\sigma_i = \frac{\text{FWHM}}{2\sqrt{2\ln 2}} \quad (3.2)$$

where  $\bar{\lambda}$  is the central wavelength,  $\sigma$  is the bandwidth, and  $\lambda$  is the wavelength. Figure 3.3 (a, b) show the SRFs of AVIRIS-NG and Sentinel-2 sensors.



(a)



(b)

**Figure 3.3. (a)** Spectral response function for the sensors of AVIRIS-NG (for the first 90 bands); **(b)** spectral response function for the sensors of Sentinel-2.

### 3.2.3.2 Spectral unmixing

The method used for the study was UPDM, which is a sensor-independent spectral reconstruction method. UPDM is a multidimensional analysis, whose working is similar to the spectral unmixing method. The reflectance of each pixel is decomposed into the summation of standard spectra of various significant classes. The three major classes present in the study site are vegetation, urban, and soil. Under the vegetation class, different agricultural crops are present, which was identified and noted during data sampling. Buildings and road come under the concrete class, and harvested agricultural, ploughed field, and bare soil area come under the soil class. These three

endmembers were used in UPDM to generate the AVIRIS-NG image for the entire area.

$$R_i = P_{iv}C_v + P_{is}C_s + P_{ic}C_c \quad (3.3)$$

$R_i$  is the reflectance of a particular pixel in the band  $i$  measured by any sensor.  $C_v$ ,  $C_s$ , and  $C_c$  are decomposition coefficients for vegetation (v), soil (s), and concrete (c), respectively;  $P_{iv}$ ,  $P_{is}$ ,  $P_{ic}$  are respective normalized standard spectra for the mentioned classes.

Mathematically, UPDM can be expressed in the form of matrices as shown below.

$$R = P \times C \quad (3.4)$$

$$\begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_n \end{pmatrix} = \begin{pmatrix} P_{1v} & P_{1s} & P_{1c} \\ P_{2v} & P_{2s} & P_{2c} \\ \vdots & \vdots & \vdots \\ P_{nv} & P_{ns} & P_{nc} \end{pmatrix} \begin{pmatrix} C_v \\ C_s \\ C_c \end{pmatrix} \quad (3.5)$$

Here,  $n$  is the number of bands, and  $R_1, R_2, \dots, R_n$  is the reflectance value of a pixel in each band.  $P$  is the standard pattern matrix of order  $n \times 3$ , and it is unique for every sensor.  $C$  is the decomposition coefficient matrix.

For a multispectral sensor, Equation (3.2) becomes

$$R_m = P_m C_m \quad (3.6)$$

Decomposition coefficients matrix ( $C_m$ ) were calculated from Sentinel-2 using UPDM.  $C_m$  calculates the fraction of class in a pixel. This matrix would remain same for simulated hyperspectral data due to its sensor independent property.  $R_m$  is reflectance value matrix for Sentinel-2 image and  $P_m$  is the standard spectra matrix for major classes (v, s, c).  $C_m$  is obtained by minimizing sum of squared error method:

$$C_m = (P_m^T \cdot P_m)^{-1} \cdot P_m^T \cdot R_m \quad (3.7)$$

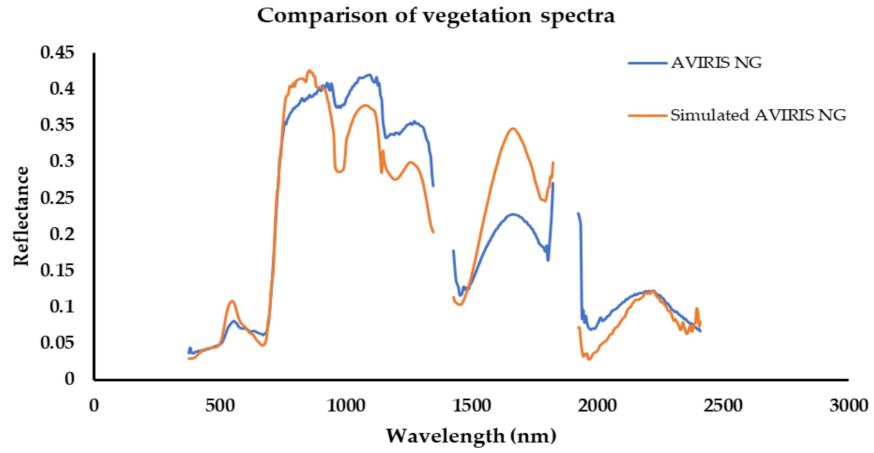
Furthermore, for the simulation of HS spectra,

$$R_h = P_h C_h \quad (3.8)$$

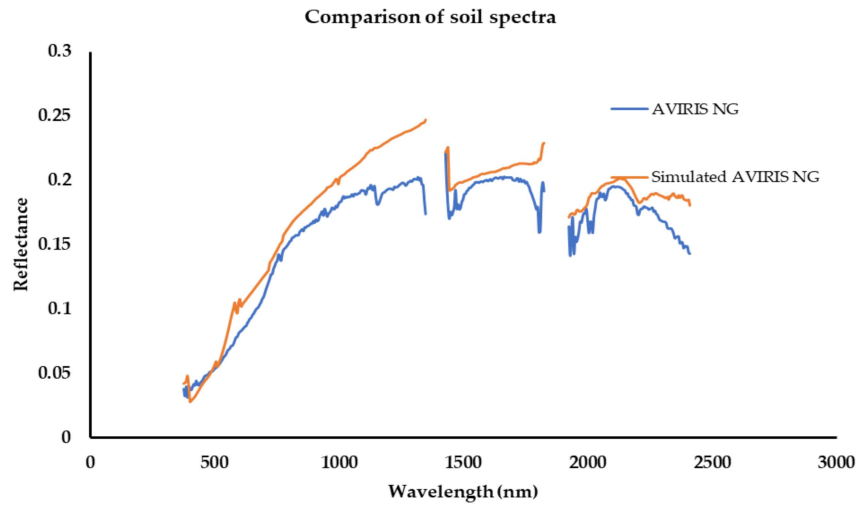
$C_m$  from Equation 3.7 replaces  $C_h$  in the Equation 3.8. The spatial resolution of simulated HS data would have the same value as that of MS data, this is why the coefficient matrix does not change for both the sensors.

$$R_h = P_h \cdot (P_m^T \cdot P_m)^{-1} \cdot P_m^T \cdot R_m \quad (3.9)$$

$R_h$  is the resultant simulated HS reflectance matrix, and it contains 425 simulated bands of AVIRIS-NG. The simulation part was done in MatLab. Figure 3.4 & 3.5 show the comparison original and simulated spectra of a vegetation and soil pixel. Some of the noisy bands were removed from both the spectra due to one or other kind of scattering or absorption e.g., bands 196-210 and 290-309 were removed due to water vapor absorption bands, whereas bands 408-425 were removed due to strong water vapor and methane absorption.



**Figure 3.4.** Comparison of vegetation spectra for AVIRIS-NG and simulated AVIRIS-NG



**Figure 3.5.** Comparison of soil spectra for AVIRIS-NG and simulated AVIRIS-NG

### 3.2.4 Classification using Spectral Angle Mapper (SAM)

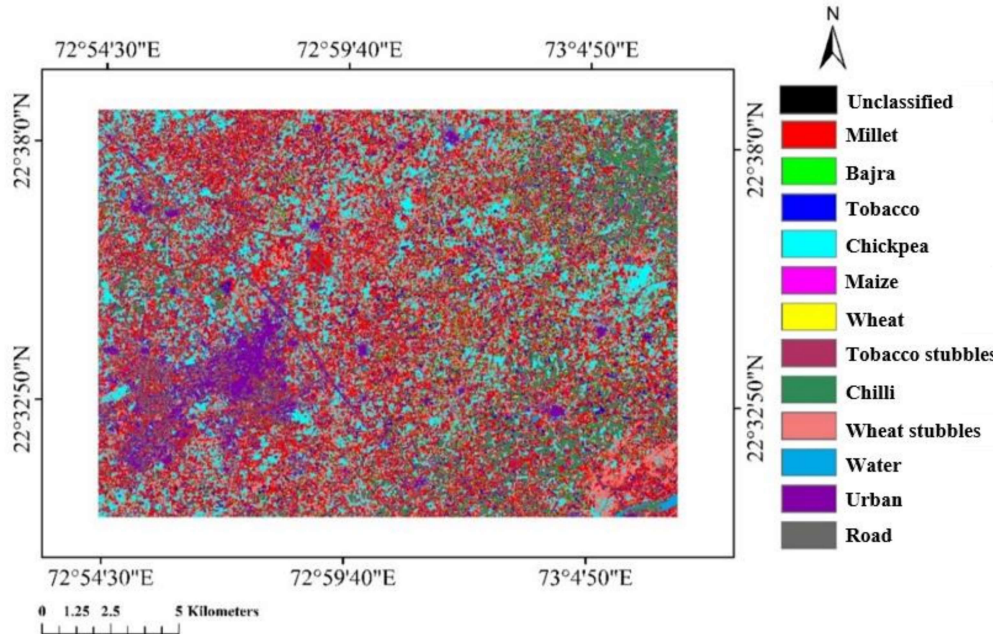
Different classes in the generated HS image were identified using SAM, which identifies unknown pixels on the basis of the degree of similarity between their spectral

profiles and those of the references. SAM considers the reference classes and spectra as vectors in a space that has dimensions equal to the number of bands.

$$\alpha = \cos^{-1} \frac{\sum X.Y}{\sqrt{\sum(X)^2 \sum(Y)^2}} \quad (3.10)$$

where  $\alpha$  is the angle between vectors (references and original spectra), X is the original spectrum, and Y is the reference spectrum. Different values of  $\alpha$  were given to run the classification; for  $\alpha = 0.5$  radian, the classifier gave the best results.

Classification distinguished 12 different classes in the simulated imagery including different agricultural species, urban, water, and road (Figure 3.6). The ground-truth image i.e. original AVIRIS-NG image was used to classify the image. The classified image was further evaluated using indices shown in Table 3.1



**Figure 3.6** Classification on the simulated image

### 3.2.5 Calculation of vegetation index for the retrieval of LCC

Singh et al. [55] generated several models for the estimation of vegetation LCC; amongst 12 models, the best-fitted modelled normalized VI, chlorophyll content index 10 (CCI10), was chosen to estimate the LCC for the given area.

$$CCI10 = \frac{R_{1954} - R_{416}}{R_{1954} + R_{416}} \quad (3.11)$$

where  $R_i$  denotes the reflectance values for the  $i^{th}$  band which is 1954<sup>th</sup> and 416<sup>th</sup> bands.

This vegetation index was further used to create the model for the retrieval of chlorophyll content (LCC), equated as

$$LCC = 93.45*(CCI10) + 27.36 \quad (3.12)$$

## 3.4 RESULTS AND DISCUSSION

### 3.4.1 Simulation of AVIRIS-NG imagery

In the present work, synthetic AVIRIS-NG imagery was generated using Sentinel-2 (multispectral imagery) and UPDM (a spectral unmixing method). Some of the noisy bands for the simulated and AVIRIS-NG were removed, i.e., bands from 1353.55–1423.67nm, 1829.36–1919.53nm, and 2415.38–2500.54nm. Figure 3.2 shows the Sentinel-2 imagery and its corresponding simulated HS image, where a small portion from both the images is clipped out and compared with original AVIRIS-NG imagery. The red, green, and blue (RGB) band combination shown in the figure was bands 8, 4, and 3 for Sentinel-2 and bands 97, 56, and 36 for the simulated and AVIRIS-NG imagery. It is clear from the figure that the simulated image has more detailed information in addition to the richness expected in spectral resolution compared to the Sentinel-2 image, and it was quite close to the original AVIRIS-NG imagery. In order

to further investigate the simulated imagery, Figures 3.4 and 3.5 show the spectral comparison for vegetation, where it is evident that the simulated image showed a similar pattern from the blue to red edge (396.89–752.51nm) region and from the SWIR to far-IR (1929.55–2405.37nm) region. In the case of soil, regions 376.86–807.60nm and 1924.54–2235.08nm showed the best similarity with the original image. It is clear that the simulated imagery and AVIRIS-NG imagery have very good spectral signature similarity, and this property of simulated imagery could be further scrutinized for the vegetation parameter retrieval. Under close inspection, it can be seen that the simulated imagery has more clarity, and classes could be differentiated clearly as compared to Sentinel-2 image. To justify this, a few of the regions are marked in Figure 3.2. The squared region shows the ‘vegetation’ area, the rounded square area covers the ‘road’, and the circled portion shows the ‘water bodies’ in the clipped images of Sentinel-2, AVIRIS-NG, and simulated AVIRIS-NG. It is evident from the figure that the simulated image had clarity in spectral resolution as compared to the Sentinel-2 image. However, it is obvious that these classes were most easily identifiable in AVIRIS-NG. Thus, it is clear that the simulated image has a high degree of resemblance in characteristics compared to the original HS image.

### **3.4.2 Classification over simulated image**

During sampling, for the agricultural area, the region identified, i.e., the name of crops, was also noted. Furthermore, since the study area included urban, water bodies, and soil in addition to crops, a total of 12 classes were identified and given to the classifier. In total, 12 classes were identified using the SAM classifier in the simulated region, i.e., fennel (*Foeniculum vulgare*), millet (*Pennisetum glaucum*), tobacco (*Nicotiana tabacum*), chickpea (*Cicer arietinum*), maize (*Zea mays*), wheat (*Triticum*), tobacco

stubbles, chili (*Capsicum frutescens*), and wheat stubbles, as well as urban, water, and road areas. Pure pixels were utilized to classify the image, and an independent set of ground datasets were used to validate the classified image. As endmembers to categorize different species and classes for the categorization, the spectral profiles from the sample sites were used. The performance of the classifier was evaluated using the overall accuracy (OA), kappa coefficient, precision, recall, and F1-score, as presented in Table 3.1. The overall accuracy came out to 87.40% with a kappa coefficient of 0.85.

**Table 3.1.** Performance assessment of the classifier

| <b>Overall accuracy</b> | <b>Precision</b> | <b>Recall</b> | <b>F1 score</b> | <b>Kappsa coefficient</b> |
|-------------------------|------------------|---------------|-----------------|---------------------------|
| 87.40 %                 | 85.63 %          | 83.29%        | 84.47%          | 0.85                      |

Precision and recall are two critical model assessment measures. While precision is the proportion of relevant outcomes, recall is the percentage of total relevant results properly categorized by the algorithm. There is a simpler measure that considers both precision and recall, called the F1-score, which is just the harmonic mean of accuracy and recall. A higher value of this score indicates a better model, with a maximum attainable score of 1 or 100%.

Mathematically, the F1-score is calculated as follows

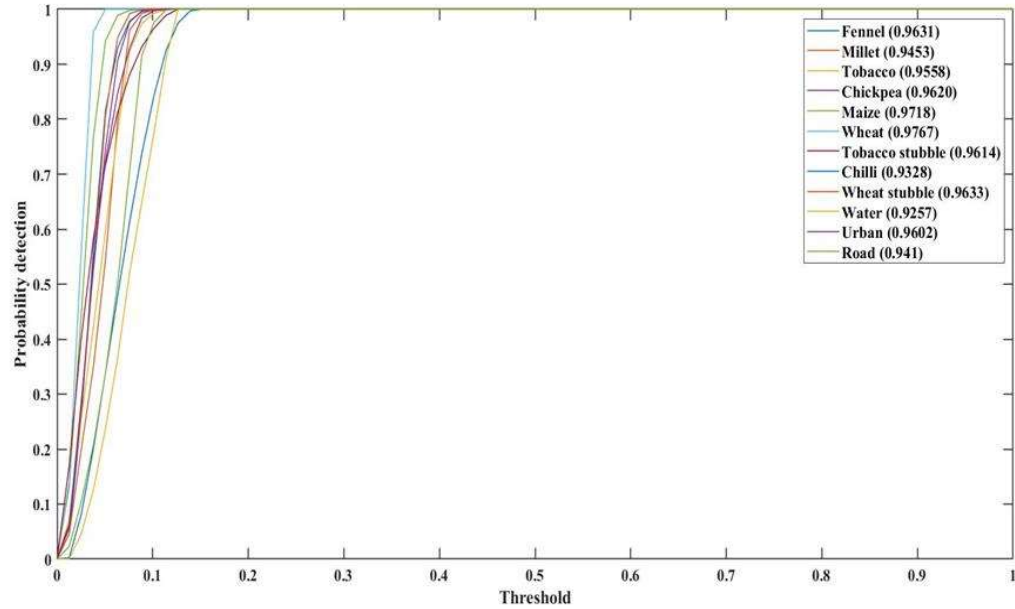
$$F1 = 2 * \frac{\text{Precision*Recall}}{\text{Precision+Recall}} \quad (3.13)$$

The user and producer accuracies for each class are shown in Table 3.2. It is evident that the ‘urban’ class has the highest user accuracy, whereas ‘fennel’ and ‘water’ have the highest producer accuracy.

**Table 3.2.** Producer and user accuracy of each class for the classified image

| Classes          | Producer accuracy (%) | User accuracy (%) |
|------------------|-----------------------|-------------------|
| Fennel           | 100                   | 78.4              |
| Millet           | 74.22                 | 89.95             |
| Tobacco          | 72.51                 | 95.63             |
| Chickpea         | 98.85                 | 94.44             |
| Maize            | 66.41                 | 90.23             |
| Wheat            | 42.68                 | 98                |
| Tobacco stubbles | 97.72                 | 73.94             |
| Chilli           | 99.84                 | 84.86             |
| Wheat stubbles   | 86.5                  | 92.06             |
| Water            | 100                   | 42.8              |
| Urban            | 88.12                 | 98.24             |
| Road             | 72.65                 | 89.17             |
| Unclassified     | 0.07                  | 100               |

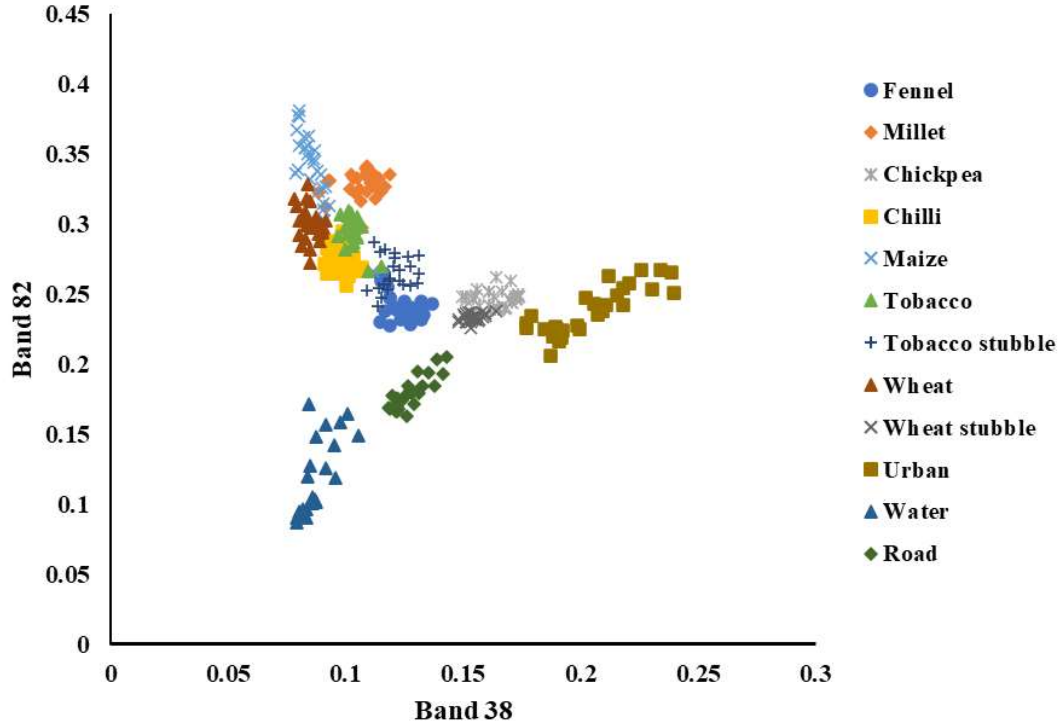
A Receiver Operating Characteristics (ROC) curve is a binary classification problem assessment measure. It is a probability curve that separates ‘signal’ from ‘noise’. The Area Under Curve (AUC) is a measure of a classifier’s ability to discriminate between classes and is used to summarize the ROC curve. A ROC curve was plotted for the classifier to show its performance, and the AUC was further calculated. A greater AUC indicates a better ability of the model in differentiating between positive and negative classes. When  $AUC = 1$ , the classifier works best to distinguish amongst classes; a value  $> 0.5$  shows that the model performs well in classification. For the SAM classifier applied for each class, we obtained values  $> 0.9$ , as shown in Figure 3.7.



**Figure 3.7.** ROC curve for the classification with AUC values

### 3.4.3 Class Separability (CS) Analysis

It is critical from the standpoint of computing cost and complexity that items of distinct classes be distinguished using specific metrics. The idea of CS is one such statistic. The class separability plot was created utilizing pixel values between the two bands to evaluate any mixed pixel effects, as shown in Figure 3.8. CS was plotted for two bands, out of few band combinations chosen randomly, band 38 vs. band 82, gave the most appropriate separation. The plot shows that the ‘urban’, ‘water’, and ‘road’ classes were well separated. Moreover, classes of different vegetation were fairly separated with a few overlapping pixels amongst them.

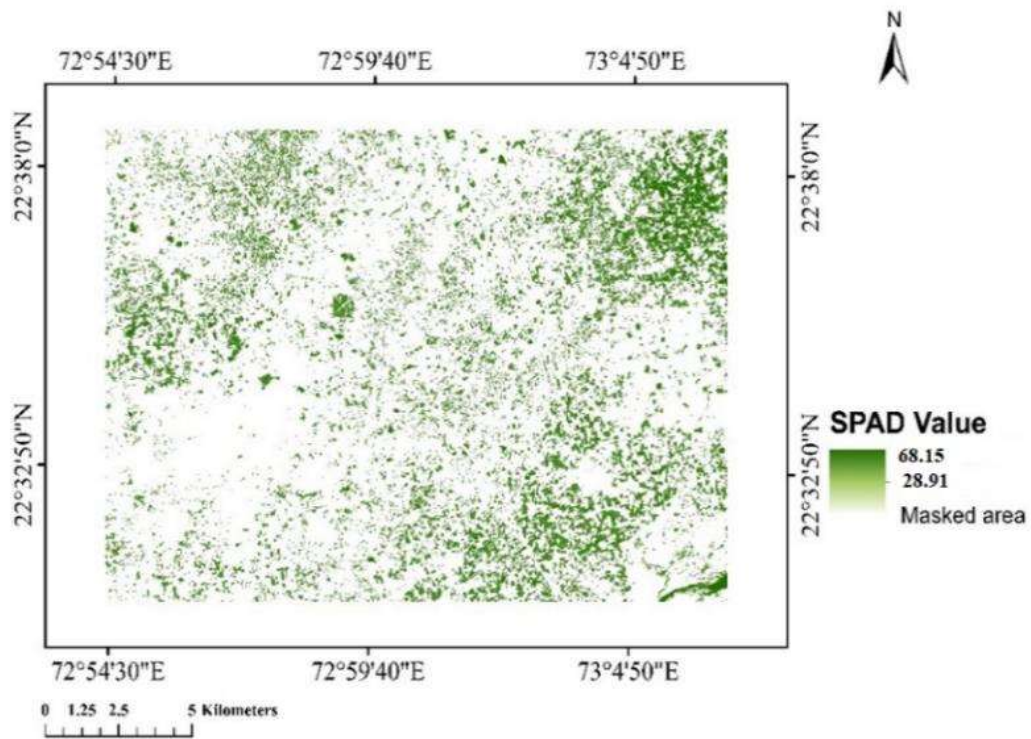


**Figure 3.8.** Class separability chart

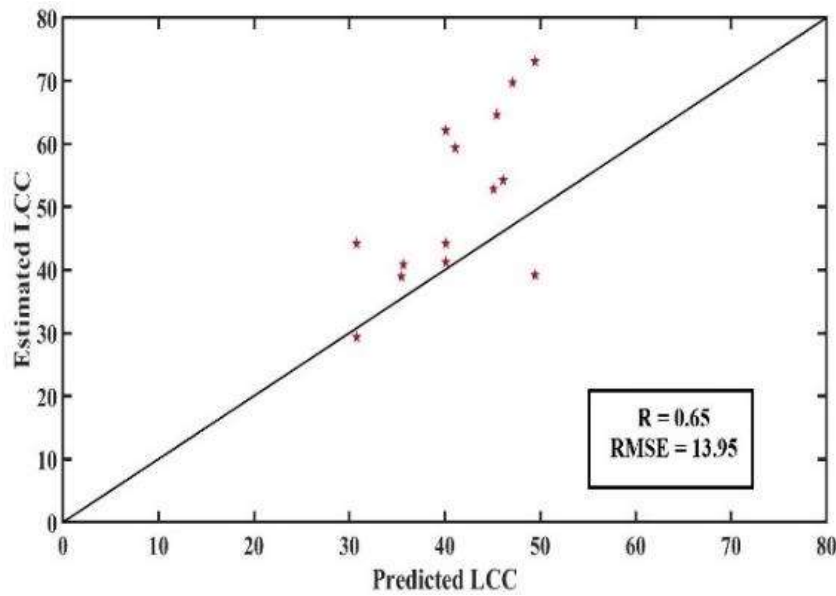
#### 3.4.4 LCC Retrieval

In order to estimate the LCC from the simulated HS imagery, the vegetation index CCI10, developed by Singh et al. [55], was used for AVIRIS-NG imagery. CCI10 calculates the normalized ratio of bands 416nm in the blue region and 1934nm in the NIR region, and it was found to be the best model for the estimation of LCC in the vegetation area for the site. The estimated values for the LCC values for the entire area for simulated imagery were found in the range of 28.91 to 68.15 of SPAD values. These values were validated with the in-situ chlorophyll data, yielding  $R = 0.65$ , with RMSE value = 13.95. Figure 3.9 shows the chlorophyll map, and Figure 3.10 is the scatter plot for the comparison between estimated LCC and in situ LCC values. A chlorophyll map was generated for the simulated image of the entire area, using VI

(Equation (3.11)) and the model (Equation (3.12)). In the chlorophyll map, areas with certain values in the lower and upper bounds were masked such that it showed values for the vegetation part only. The model was generated for the AVIRIS-NG campaign conducted at the same site for the year 2016 (Figure 3.9 and Figure 3.10) to check if the model generated would work in another agricultural site as well. The AVIRIS-NG flight was planned and was flown at one day only that is why the mapping was done on one day's image.



**Figure 3.9.** Chlorophyll map for the simulated AVIRIS-NG



**Figure 3.10.** Comparison of LCC estimated with in-situ data

In previous studies performed on the classification of simulated images via UPDM, Badola et al. [56] found that the classifier performed best on the AVIRIS-NG image with an accuracy of 94.6%, followed by simulated imagery (89%) and Sentinel-2 imagery (77.8%) for forestry area. They also assessed the effects of different reflectance values on classification accuracy, finding that, for up to a 25% difference in reflectance values, there was very low impact on classification accuracy. Liu et al. [39] also attempted classifying a simulated image (ALI multispectral to EO-1 Hyperion data), and they found overall accuracies of 87.6% and 86.8% for simulated data and ALI data, respectively. Jyoti et al. [57] performed mangrove classification on AVIRIS-NG imagery using a support vector machine (SVM), and they found an accuracy of 87.61%, whereas Shahbaz et al. [58] performed vegetation species identification using SVM and ANN for an AVIRIS-NG image and found an accuracy of 86.75% and 86.45%. In view of the studies shown, the AVIRIS-NG simulated in this study showed very comparable results for classification purposes. The classification accuracy obtained for simulated AVIRIS-NG was very similar to the

classification accuracy of the original AVIRIS-NG image; therefore, UPDM works well for the simulation of hyperspectral image.

### **3.5 CONCLUSION**

The present work showed the efficiency of simulated hyperspectral imagery (AVIRIS-NG) for the retrieval of LCC. For the simulation of imagery, many endmember combinations were investigated to get the best possible results: vegetation–soil (VS), vegetation–soil–concrete (VSC), vegetation–soil–asphalt (VSA), and vegetation–soil–concrete–asphalt (VSCA). For VS and VSCA, we found poor results, whereas, for VSA and VSC, we got fair results, among which VSC gave the best estimation for the study site. Accordingly, it can be inferred that this model worked best with only three classes; thus, when using UPDM to generate imagery, only three distinct classes should be taken for calculation. From the simulated results, there was higher similarity in the visible to red-edge region and far-SWIR region, whereas the model usually overestimated in the SWIR region for vegetation. However, in the case of soil, the model worked acceptably for almost the entire spectrum. Therefore, it can be deduced that this method may also work well for soil science studies. According to the findings of Zhang et al. [52] with respect to differences between simulated and original AVIRIS-NG spectra, they were much less for the visible to mid-infrared (<1000 nm) region, while some differences were seen beyond that. These differences could be due to the spatial resolution, since Sentinel-2 has 20m spatial resolution in the SWIR band that we resampled to 10m. The simulated imagery was more detailed than the Sentinel-2 imagery, and the generated hyperspectral data looked like the AVIRIS-NG data, with fine spatial features retained. Overall, the simulated hyperspectral imaging outperformed Sentinel-2 images in terms of spectral resolution. Classification was also

performed on the simulated image, and it showed good accuracy, with classes tobacco, chickpea, maize, wheat, wheat stubbles, and urban having a user accuracy above 90%, whereas water had the lowest user accuracy value. For UPDM, three endmembers were used to simulate the image. Since there were 12 classes including different crops, urban areas, water bodies, roads, and buildings present in the study area, classification was run on the simulated image to check if the UPDM-generated image could effectively identify different classes.

In the CS analysis, performed on the classified image, it was found that almost every vegetation species had a few overlapping pixels, and, albeit fairly separated, all of the vegetation classes were near to each other. This could be explained by the fact that, while providing endmembers to the UPDM algorithm, one type of vegetation spectrum was used in the calculation; hence, the base spectrum for each species was the same. Although in order to separate vegetation classes more clearly bands from red and NIR region could be chosen.

The LCC retrieval was performed using linear regression from the CCI10 model, and the accuracy of the model can be further increased if nonlinear regression is implemented. Furthermore, use of the radiative transfer model can also be explored. Until now, no study has investigated the usefulness of simulated AVIRIS-NG imagery for the estimation of vegetation parameters. Therefore, in order to extend the application of the data simulation from UPDM to precision agriculture, a chlorophyll map was generated for the entire area. LCC estimated using the CCI10 model from the spectra generated was compared with the in-situ data, revealing acceptable results. It was found that spectra in visible, red-edge, SWIR, and far-IR regions showed good correlation with the original AVIRIS-NG bands. The regions near 1900nm and 416nm

showed similar trends to the original image, indicating a good correlation with LCC in terms of the CCI10 VI model.

So, in conclusion, in the presented study, a hyperspectral imagery for AVIRIS-NG was simulated by utilizing a spectral reconstruction approach (UPDM) and multispectral imagery (Sentinel-2). The generated image was further evaluated with statistical analysis and visually interpreted. In order to explore the usability of the generated model in terms of vegetation biochemical parameter retrieval, it was applied to calculate the leaf chlorophyll content for different vegetation species. A chlorophyll map for the region was generated to show the variation in LCC of different vegetation under different stages, and this was validated for some points by LCC taken during data sampling. Singh et al.,2021 generated different models for the estimation of LCC for the same region using AVIRIS-NG imagery, and the best model ( $R = 0.706$ ) for the site was chosen to retrieve LCC for simulated HS imagery. It was found that the simulated imagery gave quite agreeable ( $R = 0.65$ ) LCC values with respect to in situ data. In conclusion, this free-of-cost HS imagery simulated using UPDM and MS imagery can be used further in the agriculture community for the retrieval of vegetation parameters.

