Crop type discrimination and health assessment using hyperspectral imaging

Rahul Nigam[1],*, Rojalin Tripathy[1], Sujay Dutta[1], Nita Bhagia[1], Rohit Nagori[1], K. Chandrasekar[2], Rajsi Kot[3], Bimal K. Bhattacharya[1] and Susan Ustin[4]

[1]Agriculture and Land Eco-system Division, Earth, Ocean, Atmosphere, Planetary Sciences and Applications Area,
Space Applications Centre (ISRO), Ahmedabad 380 015, India
[2]National Remote Sensing Centre (ISRO), Hyderabad 500 037, India
[3]M.G. Science Institute, Ahmedabad 380 009, India
[4]Environmental and Resource Sciences, University of California, Davis, CA 95616, USA

Advancements in hyperspectral remote sensing technology have opened new avenues to explore innovative ways to map crops in terms of area and health. To study precise mapping of agriculture and horticulture crops along with biophysical and biochemical constituents at field scale, an airborne AVIRIS-NG hyperspectral imaging has been conducted in various agro-climatic regions representing diverse agricultural types of India. Crop classification with available and developed algorithms has been applied over homogeneous and heterogeneous agriculture and horticulture cropped areas. The spectral angle mapper and maximum likelihood algorithms showed classification accuracy of 77%–94% for AVIRI-NG and 42%–55% for LISS IV. The customized deep neural network and maximum noise function (MNF)-based classification schemes showed an accuracy of 93% and 86% for mapping of agriculture and horticulture crops respectively. The forward and inversion of canopy radiative transfer model protocol was developed for retrieval of crop parameters such as leaf area index (LAI) and chlorophyll content (C_{ab}) using AVIRIS-NG narrow bands. The retrieved LAI and C_{ab} showed 19%–27% and 23%–29% deviation from measured mean for homogeneous and heterogeneous agricultural areas respectively. Red edge position index-based empirical model and multivariate linear regression of multiple indices showed maximum correlation of 0.62 and 0.93 respectively, to map leaf nitrogen content. Water condition index was developed using vegetation and water indices to distinguish crop water-based abiotic stress. Wheat yellow rust disease has been identified at field scale using absorption band depth analysis at 662–702 and 2155–2175 nm, and further applied to AVIRIS-NG data to detect biotic stress at spatial scale. This study establishes that such missions have the potential to boost accurate mapping of economically valuable minor crops and generate health indicators to distinguish biotic and abiotic stresses at field scale.

Keywords: Assessment, biotic and abiotic stress, crop classification, health, hyperspectral imaging.

Introduction

THE growing burden of population over natural resources and economic cost of agricultural management limit the crop area and production in India. Thus, requirement of persistent and precise monitoring of agricultural growth and health is of paramount importance for judicious use of farm resources to manage potential crop yield. The technological advancements in the field of remote sensing proved their worth to characterize agricultural cropland from field to regional scale. Since the last three decades, traditional multispectral broadband sensors have been used for estimation of crop area and in-season monitoring. However, these sensors have known limitations in terms of spectral bandwidth and spatial resolution[1,2].

Moreover, they are unable to map biophysical and biochemical parameters of crops[3]. These factors lead to significant uncertainties in classification and health monitoring of crops. This needs specific narrow bands to study spectral properties with reference to molecular composition of the plant material. Hyperspectral remote sensing (or imaging spectroscopy) shows a great potential and improvements in classification of various crop types, retrieval of biophysical and biochemical contents, estimation of nutrient content, and detecting abiotic and biotic stresses compared to traditional broadband spectral information. Hyperspectral remote sensing technology provides the opportunity to map the response of different crop types in terms of morphological and physiological characteristics in continuous spectral bands[4,5]. Accuracy of classification will further increase by reduction in dimensionality and redundancy of hyperspectral data. Usage of different techniques such as principal component analysis (PCA), maximum noise fraction (MNF) transformation followed by pixel purity index can help in the reduction of data dimensionality and promise higher

classification accuracy. Crop physical and biochemical parameters such as leaf area index (LAI), chlorophyll content (C_{ab}) and nitrogen content (N) will provide indicators to assess crop physiological state under varying environments[6]. LAI, C_{ab} and N have a direct role in photosynthetic processes of light harvesting and initiation of electron transport, and its responsiveness will change according to the severity of biotic and abiotic stresses[7]. These parameters can be retrieved using canopy radiative transfer models and different hyperspectral vegetation indices. The biophysical and biochemical constituents of crop canopies are directly expressed in the reflectance signatures that can be derived from imaging spectrometers measurements[8]. The spectral characteristics of vital expressions and controls of vegetation permit us for quantitative applications of imaging spectroscopy in order to address uncertainty of agro-ecosystem. Moreover, spectroscopic remote sensing can act as a bridge between field to regional scale and could also serve as a bridge to regions specific to global space-borne remote sensing missions, where coarse pixel size precludes direct comparison with fine scale measurements of important agro-ecosystem properties[9].

In India under joint collaboration of Indian Space Research Organisation (ISRO), National Aeronautics Space Administration (NASA), Jet Propulsion Lab (JPL), an airborne campaign was organized to perform spectroscopic imaging of selected agricultural sites of India. In this campaign, Airborne Visible/Infrared Imaging Spectrometer-Next Generation (AVIRIS-NG) sensor was flown aboard on ISRO B200 aircraft. AVIRIS-NG is an imaging spectrometer having around 425 contiguous narrow spectral bands in range of 380–2500 nm with high spectral resolution of about 5 nm and Instantaneous Field of View (IFOV) of 1 m rad (https://aviris-ng.jpl.nasa.gov/). In this study, data acquired over different agricultural sites in India from AVIRIS-NG have been used to classify crop types, and for the retrieval of biophysical and biochemical parameters, and generation of abiotic and biotic stress maps.

Study area

The homogeneous and heterogeneous agricultural sites of 20–550 sq. km were selected for AVIRIS-NG airborne flight. Here, Kota (Rajasthan), Maddur (Karnataka), Anand (Gujarat), Talala (Gujarat), Jhagdia (Gujarat), Roopnagar (Punjab) and Nagarjuna Sagar command area (Telangana) have been selected for agricultural studies. The study sites are selected on the basis of their unique agro-climatic settings, soil, crop (mono to mixed crops), rainfed and irrigated agricultural conditions. Kota site lies in central plateau and hill region, and represents homogeneous agricultural region dominated by wheat crop. Maddur site is located in Chamarajanagar district of

Karnataka, and lies in the southern plateau and hill region. The site represents heterogeneous agricultural area. Anand, Talala and Jhagdia sites lie in the Gujarat plains and hill region, and have multi crops, mango orchards under heterogeneous agricultural area. Nagarjuna Sagar command area is dominated by black cotton clayey soil of the southern plateau and hill region. Rupnagar site is homogeneous wheat area and lies in the trans-Gangetic plain region. Figure 1 gives the spatial distribution of sites over Indian land mass and Table 1 describes the area covered in AVIRIS-NG airborne flight.

Datasets used

In situ data

Crop and soil reflectance data were obtained with the ASD spectroradiometer across spectral regions of 350–2500 nm at 1 nm interval over aforementioned study sites in coherence with AVIRIS-NG flight. The instrument has been attached with standard fore-optic with 25° field of view (FOV) through a permanent fibre optic cable. All the spectral measurements were made between 1030 and 1420 h local standard time. LAI and chlorophyll index measurements for various crops were carried (LICOR-2000 Canopy analyser and Konica Minolta chlorophyll meter SPAD-502 Plus respectively) for all sites. At Anand, crop samples were collected from each site and leaf nitrogen content was estimated from them using an auto-analyser[10].
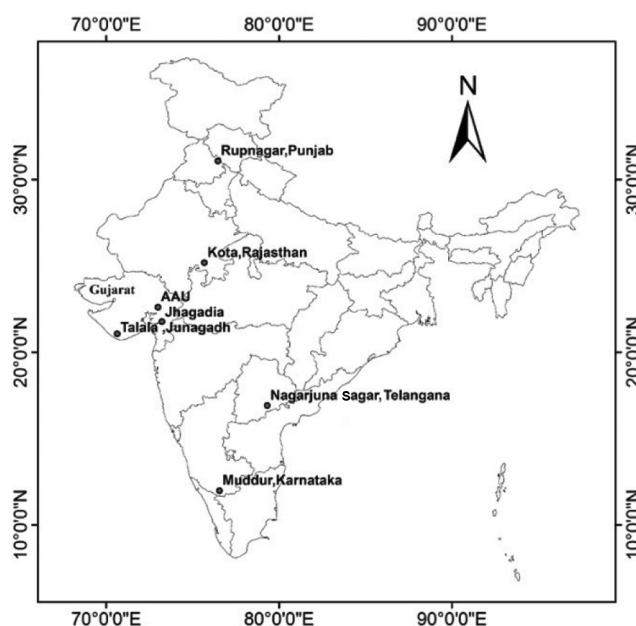

Figure 1. Location of the study area.

Table 1. Details of study sites

Site	Date of flight (2015–16)	Upper left (N) (degree)	Upper left (E) (degree)	Lower right (N) (degree)	Lower right (E) (degree)
Kota	5 February	25.18	75.67	25.08	75.68
Maddur	10 January	11.98	76.54	11.58	76.67
AAU, Anand	7 February	22.61	72.88	22.46	73.06
Talala	9 February	21.06	70.63	21.05	70.65
Jhagdia	8 February	21.77	72.95	21.45	73.29
Nagarjuna Sagar	20 December	16.90	79.31	16.93	79.35
Rupnagar	20 February	31.08	76.47	31.03	30.99
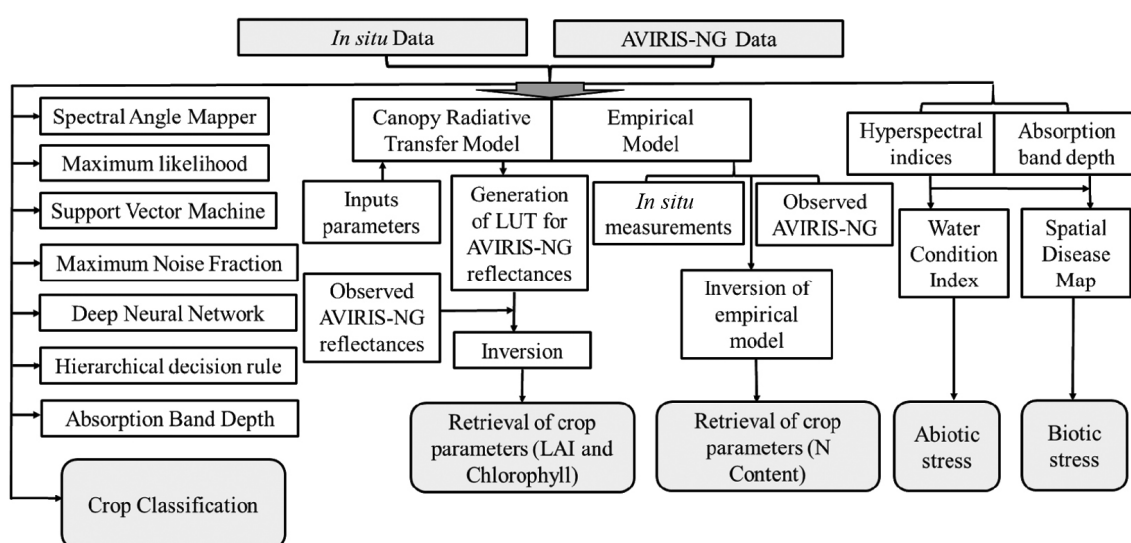
Figure 2. Flow chart of the overall methodology.

Airborne data

AVIRIS-NG is an imaging spectrometer having around 425 contiguous narrow spectral bands in the spectral range 380–2500 nm at 5 nm interval with high signal-to-noise ratio (SNR) (>2000 @ 600 nm and >1000 @ 2200 nm) and accuracy of 95% having FOV of 34° and IFOV of 1 m rad (https://aviris-ng.jpl.nasa.gov/). Ground sampling distance (GSD) vis-à-vis pixel resolution varies from 4 to 8 m for flight altitude of 4–8 km for a swath of 4–6 km.

Satellite data

Resourcesat-2 (RS-2) Linear Self Scanning Sensor (LISS) IV provides three broad spectral bands, viz. green (520–590 nm), red (620–680 nm) and NIR (770–860 nm). LISS IV data have 5.8 m spatial and 10 bit radiometric resolution. The LISS IV data over Talala region of February 2016 has been used in this study. AVIRIS-NG data have also been used to generate RS-2, LISS IV bands at parent spatial resolution and quantization of AVIRIS-NG using spectral response function of LISS IV over Maddur region.

Methodology

Figure 2 shows a flow chart of the overall methodology.

Data pre-processing

The level-2 AVIRIS-NG surface reflectance data have been used in this study. From the spectral data, Fraunhofer lines were removed for further analysis. The laboratory-computed spectral response functions of three spectral bands of RS-2 LISS IV sensor were applied over AVIRIS-NG data to generate LISS IV equivalent spectral bands.

Dimensionality reduction

To reduce data dimensionality in the present study, PCA, MNF transformation and deep neural network (DNN) based spectral band reduction methods were applied over AVIRIS-NG data at different sites.

Classification techniques

The classification techniques such as spectral angle mapper (SAM), maximum likelihood classifier (MLC),

support vector machine (SVM), MNF, hierarchical decision, classification-based on absorption band depth (ABD) and DNN have been applied using *in situ* and AVIRIS-NG data to classify different agricultural and horticultural crops. Kappa coefficient and overall accuracy coefficients were used for classification accuracy assessments[11].

Spectral angle mapper: SAM is a spectral classification that uses an *n*-dimensional angle to match pixels to reference spectra. The algorithm determines similarity between two spectra by calculating the angle between them, treating the spectra as vectors in a space with dimensionality equal to the number of bands. SAM compares the angle between the endmember spectrum vector and each pixel vector in *n*-dimensional space. Smaller angles represent closer matches to the reference spectrum. The class with which a pixel records the lowest angle is the one in which it is classified[12].

Maximum likelihood classification: MLC helps in the case of overlapping classes and calculates the likelihood of a pixel belonging to certain class based on its posterior probability. If Nc is the number of classes and likelihood of a pixel *p* belonging to certain class Wi ($i = 1$, 2, ... , Nc) can be defined in terms of posterior probability $P(Wi/p)$ and the class with which the pixel will have maximum likelihood, then it is assigned to that class[13].

Support vector machine: SVM is a binary classifier based on statistical learning theory for generating a linear separating hyper-plane that maximizes the margin between two targeted classes, i.e. maximizes the distances between the closest vectors (also known as support vectors) of the two classes. However, when noisy data lead to intermixing of classes, introduction of slack parameters or regularization parameters or penalty parameters which create a soft margin to allow some amount of training samples of one class to lie on another side of the margin, makes the concept more robust and efficient in handling noisy data[14]. Multiclass SVM classifiers are modified versions of binary SVM classifiers where pairwise strategy is mostly used in which binary classifiers for each possible pair of classes are formed[15]. The class labels that appear the most are assigned to that pixel[16].

Maximum noise fraction: MNF transformation utilizes the most common measures of image quality[17], i.e. SNR and chooses newer components such that SNR is maximized in contrast to PCA[18]. As the set of eigenvectors set obtained after maximizing noise fraction is the same as maximizing SNR (just in reverse order), MNF maximizes the variance of noise with respect to variance of whole data. When original data are transformed alongwith these new components, the MNF will show better image quality.

Hierarchical decision rule: Vegetation indices (VIs) have been combined with Hierarchical decision rule-based classification. Vegetation indices were computed and taken as input for hierarchical decision rule-based classification[19]. Different vegetation indices used were normalized difference vegetation index (NDVI), water band index (WBI) and normalized difference infrared index (NDII).

Absorption band depth: The continuum-removed reflectance was obtained by dividing the original reflectance values (R) by the corresponding values of the continuum line (RL) for all the channels in the wavelength region between the endpoints of the absorption feature: CR = (R/RL)[20]. The depth (D) of the absorption feature was calculated as the difference between the continuum line and minimum value in the continuum-removed spectral feature BD = $(1 - CR)$.

Deep neural network: After dimensionality reduction PCA technique was applied on six parts of the dataset and principal components explaining highest variability were selected for Anand site. Eigenvectors of the first two principal components were used to ascertain significant bands. Three sets of top 10, 25 and 50 significant bands with high frequency of being selected in six subsets were considered. Training sites were marked in the image using ground-truth sites of crops, vegetation classes and fallow lands. DNN with 391 bands, three hidden layers of 300, 150 and 50 nodes, rectified linear unit (Relu) activation function and output layer of training classes was used for band selection[21]. Bands which often activated the nodes of the output layer were selected as the most significant using a back traversal of neural network. The most optimal band set from both the methods was selected based on overall accuracy and average class accuracy. Classification was carried out using DNN classifier with selected optimum bands and training dataset[22]. Three hidden layers and Relu activation function were used in the neural network. Spectral profiles were generated using GT sites for different crops, vegetation classes and fallow lands. Spectral profile of predicted pixel was matched with GT-based reference profiles of selected land features with Euclidian distance of 0.075. Augmented dataset was generated with pixels whose spectral profiles matched with the reference profiles. Seperability analysis between training classes was carried out for optimum set of bands selected using PCA and DNN-based methods. Jeffries–Matusita (JM) distance and average JM distance between classes were computed and compared for both methods. On the basis of accuracy measures and seperability analysis, the final set of optimum bands was selected for classification.

Retrieval of crop parameters

One-dimensional canopy radiative transfer simulation model PROSAIL, the combined form of PROSPECT and SAIL (scattering by arbitrary inclined leaves) has been used in this study for retrieval of canopy parameters. PROSPECT simulates reflectances at leaf level and SAIL addresses the directionality[23]. The model simulates reflectance using leaf biophysical–chemical constituents such as leaf structure parameter (N), chlorophyll ($a + b$) content (C_{ab}), leaf equivalent water thickness (Cw), leaf dry matter content (Cm), LAI, leaf inclination angle (LIA), hot spot parameter (SL), horizontal visibility (vis), sun zenith angle (θ_s), view zenith angle (θ_v), relative azimuth angle (ϕ_{sv}) and soil albedo (ρ_s). The CRT model was customized for AVIRIS-NG spectral bands. The model will simulate AVIRIS-NG bands in forward simulation according to the given inputs. Input parameters of models were divided into different intervals within their theoretical lower and upper limits to cover whole dynamics of crops according to *in situ* observations and the reported literature. Considering their limits and intervals, combinations of different inputs resulted in various scenarios for the respective crop types. A cost function (*S*) was used for inversion that represents the sum of square differences between AVIRIS pixel band reflectances and model-simulated band reflectances. Minimum of the cost function was obtained using least square approach which gives unique value of LAI and C_{ab} for a given set of observed reflectances using generated Look Up Table (LUT) through forward simulations. This approach is similar to the variational method in which difference of error is minimized, but differs in observation error covariance matrices. This may be the scope of future research under that variational approach[24]. In the
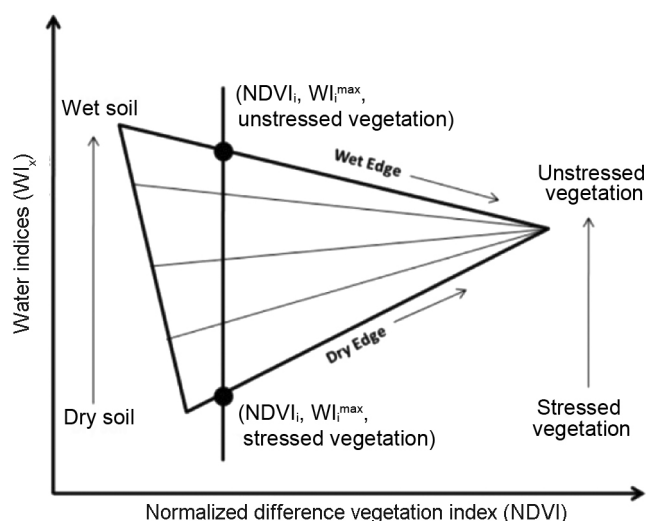
variational method, cost function, which is a function of total variance is minimized.

All nitrogen-sensitive Vegetation Indices (VIs) in blue, green, red, NIR and SWIR-1 spectral band regions were computed from ground spectra and AVIRIS-NG to estimate the existing bias between them. Individual VIs from ground spectra and the plant nitrogen content were computed to develop multivariate linear regression models with significant correlated VIs. The model was then applied to AVIRIS-NG spatial data to generate distributed plant nitrogen map. The developed models were validated with independent *in situ* data.

Abiotic stress

To integrate the moisture status and surface reflectance, water condition index (WCI) based on WIx–VI triangle space has been defined. WCI is related to the surface soil moisture status/vegetation water content, where higher values of WCI indicate wet conditions and vice versa (Figure 3). WCI is defined as

$$ \text{WCI} = \frac{\text{WI}_{xi} - \text{WI}_{\min}}{\text{WI}_{\max} - \text{WI}_{\min}}, \tag{1} $$

where

$$ \text{WI}_{\min} = a + b\text{NDVI}_i \text{ and } \text{WI}_{\max} = a_0 + b_0 * \text{NDVI}_i. $$

where NDVI_i is the normalized difference vegetation index of the *i*th pixel, a and a_0 are the intercept while b and b_0 are the slope of the dry and wet edge respectively. Other hyperspectral indices such as NDVI[25], WBI[26], NDII[27], normalized difference water index (NDWI)[28], land surface water index (LSWI)[29] have been computed according to the literature.
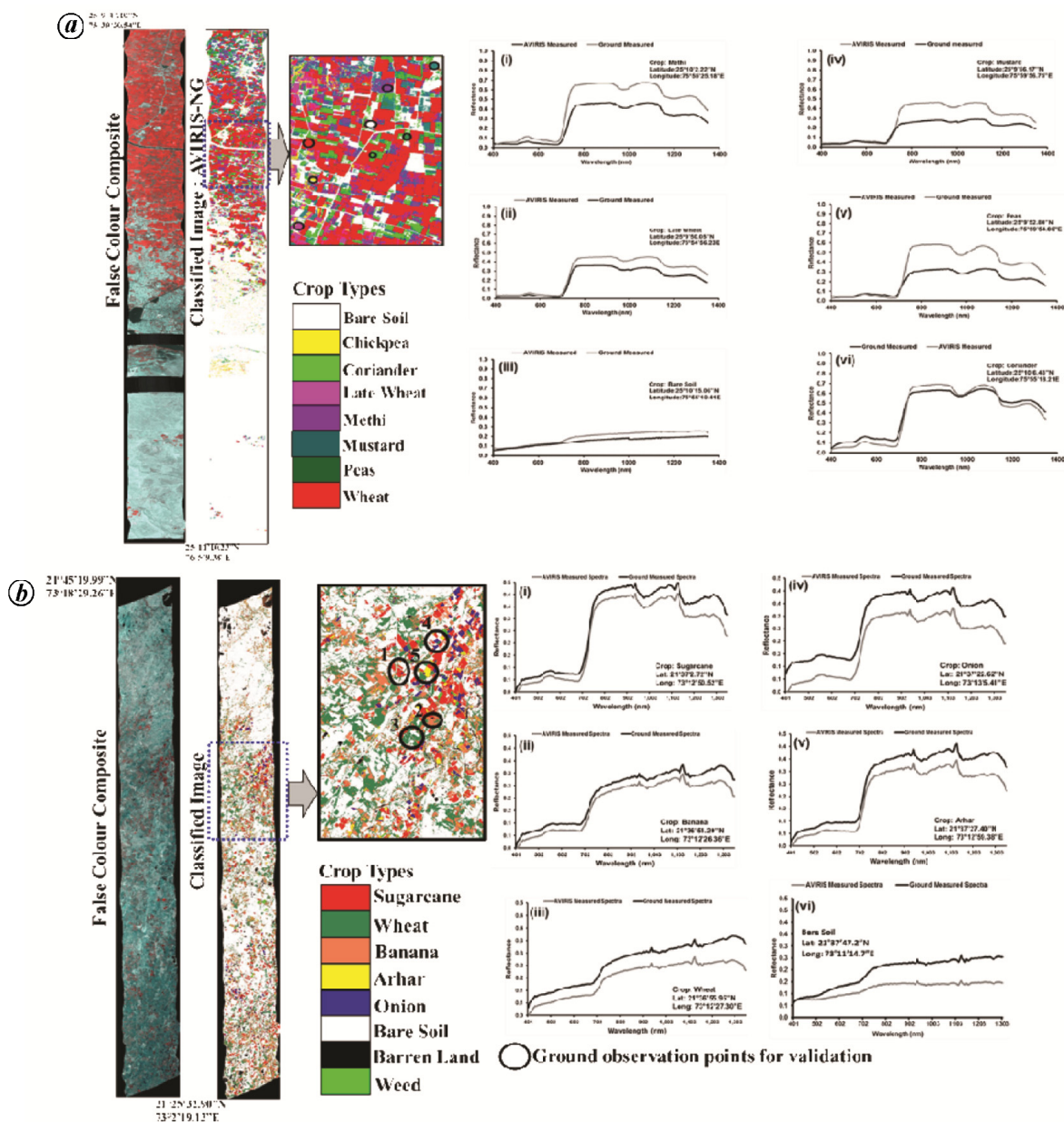
Biotic stress

Two different approaches, viz. disease index and absorption depth-based classification were used at Rupnagar site to discriminate healthy and rust-infested wheat crop. The indices such as leaf rust disease severity index (LRDSI) 1 and 2 has been used for classification[30].

Results and discussion

Optimum band selection

To reduce data dimensionality for AVIRIS-NG (400–2500 nm) and LISS IV equivalent multispectral bands generated from AVIRIS-NG, PCA was applied at Maddur site representing heterogeneous agricultural area. Typically, the first few PCs explained maximum proportion of

Figure 3. Conceptual plot of the Water Indices and Normalized Difference Vegetation Index triangle to determine soil and crop wetness status.

Figure 4. Crop-type discrimination over (*a*) homogeneous and (*b*) heterogeneous agricultural areas using supervised SAM classification.

variability in terms of eigenvalues in the data. Adjacent hyperspectral wavebands showed noise, saturation and redundancy of the data. Therefore, based on the analysis and variability of the data, it is conferred that higher the eigenvector, higher the importance of the band. For AVIRIS-NG and LISS IV five and two bands showing high eigenvalues respectively, have been used to classify crop types.

Selection of optimum bands was carried out using PCA and DNN-based band selection methods at Anand site. Three sets of top 10, 25 and 50 significant bands resulted in the selection of 20, 40 and 80 optimum bands respectively, using both the methods. Bands selected by PCA

method lie in BLUE, GREEN, RED and NIR regions, whereas bands selected by DNN method also have SWIR region.

Discrimination of mango and sapota was done with 18 high SNR bands in MNF transformed space for AVIRIS-NG image over Talala region.

At Jhagdia site, continuum removal (CR) and further normalization was done to identify spectral bandwidth 661–702, 947–998 nm to discriminate fresh and ratoon sugarcane. Moreover, these bandwidths showed maximum different in absorption band depth in whole vegetative spectrum and represent crop pigment and leaf structure properties.
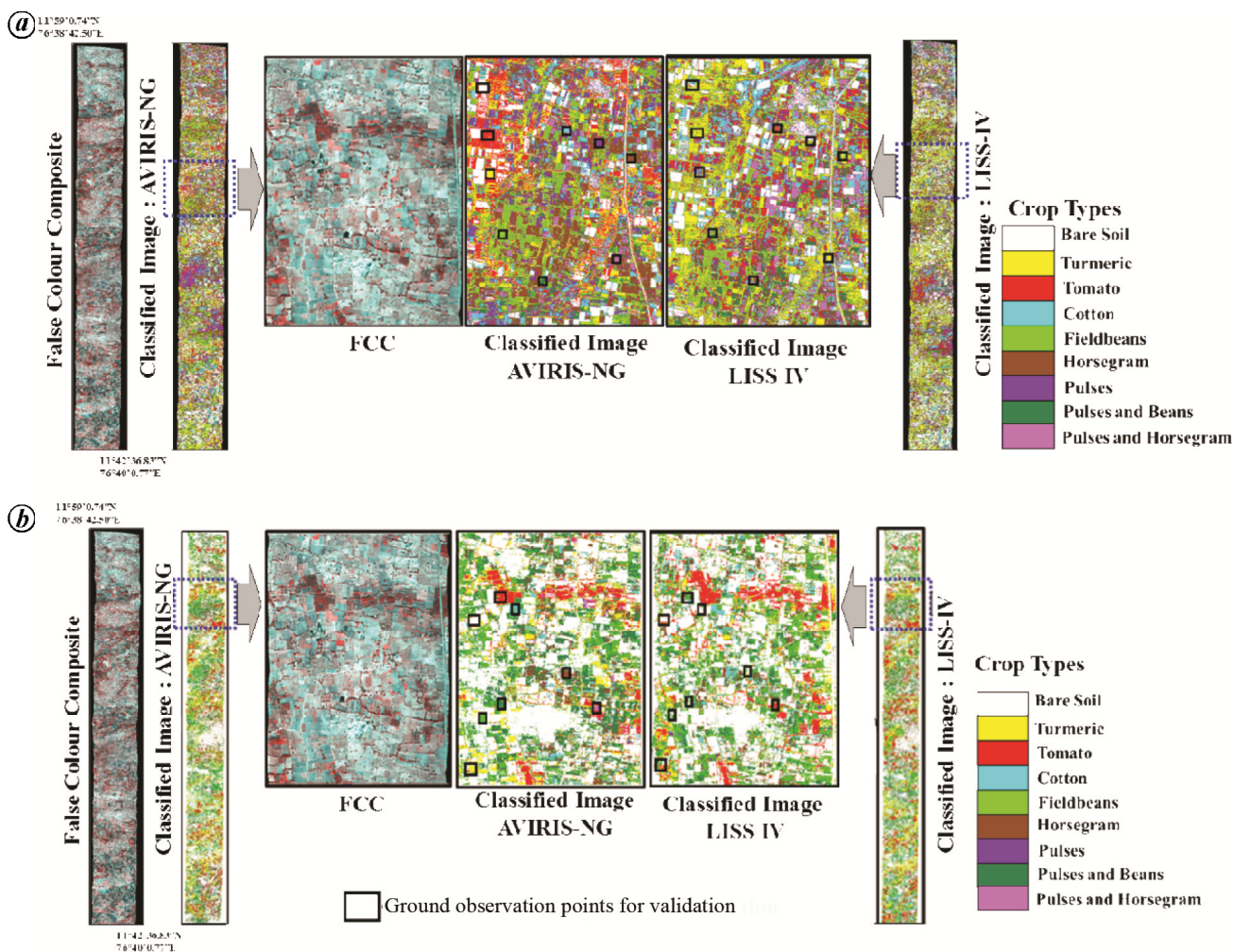
Figure 5. Crop-type discrimination over heterogeneous agricultural site of Maddur, Karnataka using (*a*) SAM and (*b*) MLC with AVIRIS-NG and LISS-IV convoluted data.

Crop classification

The SAM algorithm was applied over AVIRIS-NG data to classify crop type in homogeneous (Kota site) and heterogeneous (Jhagdia site) agricultural areas. For this 25 training datasets were prepared using *in situ* spectral observations convoluted according to the spectral bandwidth of AVIRIS-NG data in the form of ROIs (regions of interest) for various crop types. The generated classified images showed classification accuracy of 86.4% and 80.8% with kappa coefficient of 0.84 and 0.77 for Kota and Jhagadia agricultural sites respectively, with 15 independent *in situ* data. The crop-type classification and spectral behaviour of AVIRIS-NG and *in situ* data (Figure 4 *a* and *b*) reveal that curvature (slope) of both the spectral remains same for various crop types but difference in magnitude exists. This may due to (i) exposure of soil within plant canopy, (ii) two or more crops within a pixel and (iii) atmospheric perturbations.

After PCA, five from 420 bands of AVIRIS-NG and two from three bands of LISS IV were selected to classify crop types using SAM and MLC algorithms. Totally 25 ROIs were generated from *in situ* data as training dataset. AVIRIS-NG-based classification showed better accuracy compared to LISS IV equivalent multispectral data (Figure 5). This is due to the presence of specific narrow bands in AVIRIS-NG, which possess information of crop chlorophyll, protein, lignin, cellulose and nitrogen contents as well as biophysical information. Whereas LISS IV only provides information in GREEN, RED and NIR broadbands and thus is unable to address crop-specific biochemical and biophysical properties. The results show that convolution of LISS IV equivalent broadband data results in loss of crucial information essential for accurate crop discrimination, while narrow contiguous bands of AVIRIS-NG data contain this critical information. The confusion matrix generated from SAM-based AVIRIS-NG and LISS IV classified image with 15 independent *in situ* data showed accuracy of 77.7% and 42.8% and kappa coefficient of 0.75 and 0.34 respectively. While MLC-based classification showed classification accuracy 94.3% and 55.6% and kappa coefficient 0.93 and 0.46

Table 2. Cross-classification between mango and sapota orchards

User accuracy (%)	AVIRIS-NG all bands		AVIRIS-NG MNF space		LISS IV	
	Mango	Sapota	Mango	Sapota	Mango	Sapota
SAM	100	90.7	–	–	47.62	32.89
MLC	–	–	100	100	47.37	86.44
SVM	–	–	96.46	100	50.5	73.21
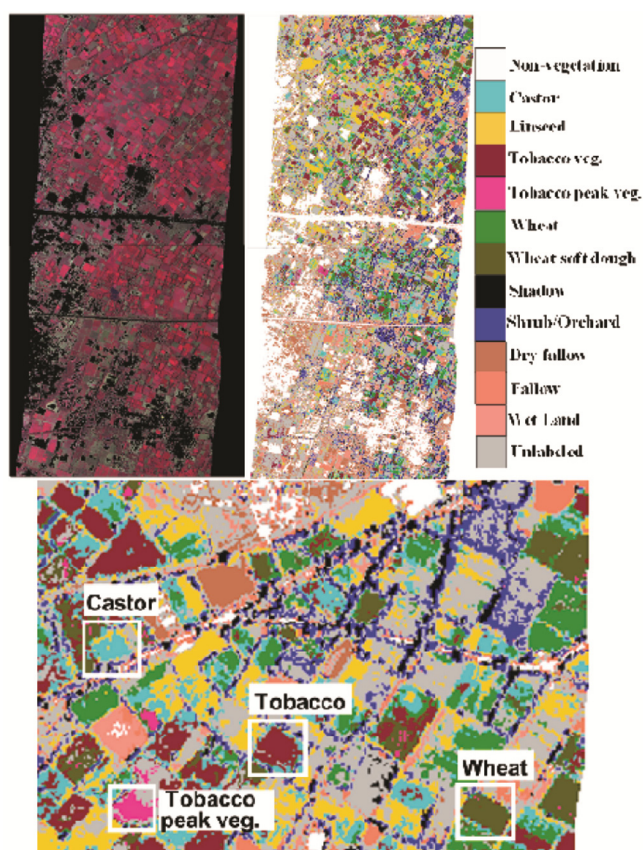
Figure 6. FCC and classified image of parts of Anand site.

respectively, for AVIRIS-NG and LISS IV equivalent multispectral bands. The higher accuracy is observed in MLC classifier suggests that the hyperspectral data derived from the optimal band configuration of the airborne sensor have a sufficiently Gaussian distribution. This gives a full and representative description of the respective classes (spectrally separable crop type), and fulfils the requirements for such a parametric algorithm[31]. The accuracy of classification can be further improved with multi-temporal hyperspectral data[32].

The PCA and DNN-based methods showed overall accuracy for 20, 40 and 80 optimum bands were 96% and 93–98% respectively, but average class accuracy ranged from 92% to 94% at Anand site. Total 3,775 and 164,226 AVIRIS-NG pixels from 12 different classes were used as training and validation datasets respectively. Overall accuracy for 20, 40 and 80 optimum bands selected with

DNN ranged between 93% and 98%, and average class accuracy ranged from 88% to 97%. Maximum average class accuracy as obtained with 40 bands was 92% with PCA and 97% with DNN-based selection method. Average JM distance (1.41) between classes was higher for optimum bands selected with DNN than those selected with PCA. Final classification was carried out with 40 bands selected using DNN method after consideration of accuracy and separability tests. Figure 6 shows the classified image. Through this classification algorithm identification of wheat at vegetative and soft dough stages, tobacco at vegetative and peak vegetative stages, castor, linseed and shrubs, dry and wet fallow lands have been done. The features not correctly identified were marked as unclassified pixels. Confusion matrix generated from training classes gave an overall accuracy of 95% and average class accuracy of 93% with kappa coefficient of 0.94. Validation of classified pixels was done by matching their spectral profiles with the respective references. Figure 7 shows the classified pixels and their reference profiles.

The different classification algorithms were used to classify homogeneous horticultural crops at Talala region using 20 *in situ* data as training dataset and 15 independent dataset for accuracy assessment. Both MLC and SVM provided classification accuracy of nearly 86% (86.64% for MLC and 85.02% for SVM), and mango versus sapota discrimination (ground truth testing data) of about 100% (except for sapota using SVM). The classification accuracy obtained by applying SAM on AVIRIS-NG all-bands data was also higher than LISS-IV multispectral image, as expected. Applying SAM classifier on AVIRIS-NG data using all bands provided a classification accuracy of ~72%. Table 2 shows inter-species cross-classification accuracy between mango or sapota. It is clearly observed that the cross-classification is much higher for AVIRIS-NG image in comparison to LISS IV data, thus highlighting the efficiency of using hyperspectral data over multispectral data. Figure 8 shows the classified map (vegetation classes only) generated using MLC over high SNR 18 bands of AVIRIS-NG data in MNF transformed space. It shows high classification accuracy and user accuracy with lowest intermixing between mango and sapota classes.

At Jhagdia site, hierarchical decision tree and continuum-removed absorption depths were used to discriminate
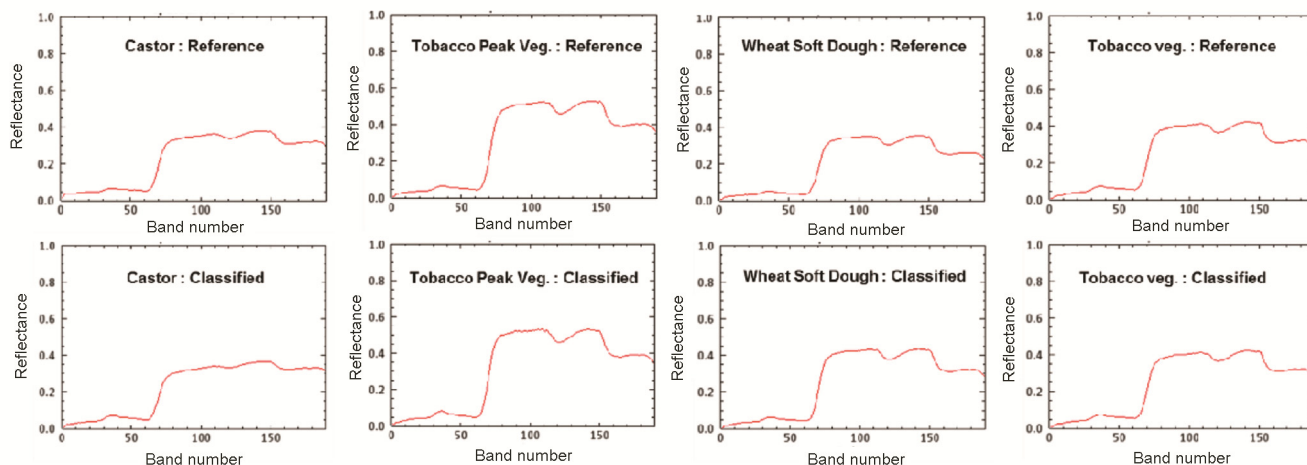
Figure 7. Spectral profiles of classified data and *in situ*-based profiles.
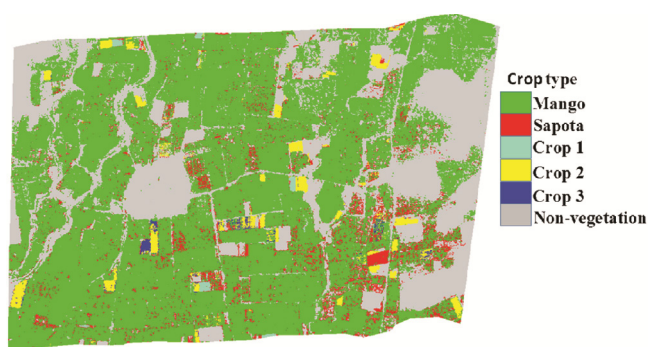

Figure 8. Classified map of Talala site state generated from AVIRIS-NG.

fresh and ratoon sugarcane. A total of 15 and 10 *in situ* datasets were used for training and validation respectively. NDVI, WBI and NDII were used in hierarchical decision tree. The threshold of these indices were generated from *in situ* data. These indices are adequate to describe the biochemical or biophysical interactions between light and matter, and have the potential for crop-type discrimination[33]. The identified spectral bands based on absorption depths computed from normalization of continuum analysis (660–702 nm and 947–998 nm) were used to discriminate fresh and ratoon sugarcane crop. The bands showing maximum difference in absorption depths were selected for discrimination analysis. The spectral fitting score with independent data sets showed maximum value of 0.87 and 0.75 for ratoon and fresh sugarcane respectively, for absorption depths between 947 and 998 nm spectral bands (Figure 9).

Retrieval of crop parameters

Figure 10 *a* and *b* shows one-dimensional sensitivity of canopy radiative transfer model for LAI and C_{ab}

respectively. Here, typical mean values of parameters over observations at Kota site have been considered for sensitivity analysis. LAI and C_{ab} varied from a fixed value of 2.5 and 30 µg cm^{-1} with an increment and decrement of 0.5 and 5 respectively. The visible and near-infrared band reflectances showed variation from –28% to 40% and from –8% to 16% respectively, for different LAI intervals. C_{ab} showed sensitivity only to visible band and the variation from its fixed value yielded –90% to 56% variation in reflectances. The analysis showed that all visible and near-infrared bands were sensitive to LAI and C_{ab}, and their fluctuation could be captured through simulated reflectance values.

The decorrelation technique showed that ten spectral bands had maximum decorrelation in the range 400–1000 nm. Among them, six (451, 551, 677, 797, 857, 882 nm) and four (451, 551, 656, 677 nm) narrow bands of AVIRIS-NG showing maximum sensitivity to LAI and chlorophyll content were selected for retrieval of these parameters. The model was run in forwarded mode by integration of the respective AVIRIS-NG bands to generate LUT for LAI and C_{ab} for the selected bands separately. The retrieval of LAI and C_{ab} was done by inversion of AVIRIS-NG reflectances using the generated LUT. LAI and C_{ab} were retrieved over the heterogeneous agriculture area of Jhagdia covering crops such as wheat, sugarcane, banana, onion and pigeon pea. At Jhagdia site, 50% of agricultural area is dominated by wheat and sugarcane crops. LAI and C_{ab} varied from 1 to 5 and 5 to 40 µg cm^{-2} respectively (Figure 11 *b* and *c*). The validation from *in situ* data showed deviation of 27.5% and 29.54% from the mean for LAI and C_{ab} respectively (Figure 12), for various crop types. The LAI and C_{ab} were also retrieved over the homogeneous agricultural area of Kota, covering crops such as wheat, mustard, beans, garlic, fenugreek, coriander, peas and onion. At Kota, wheat crop area is about 70% while 30% is covered by other
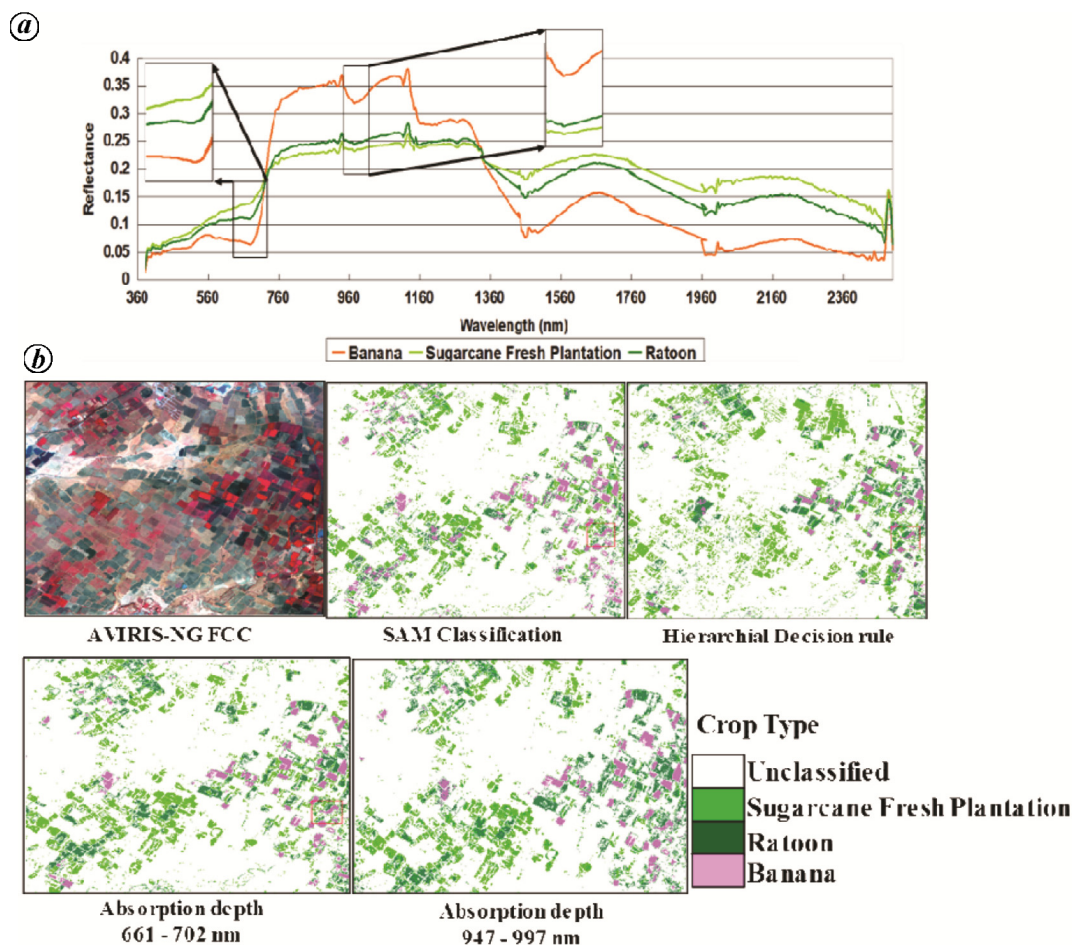
Figure 9. *a*, Spectral profile of sugarcane and ratoon crops. *b*, Spatial distribution of types of sugarcane using different classification techniques.
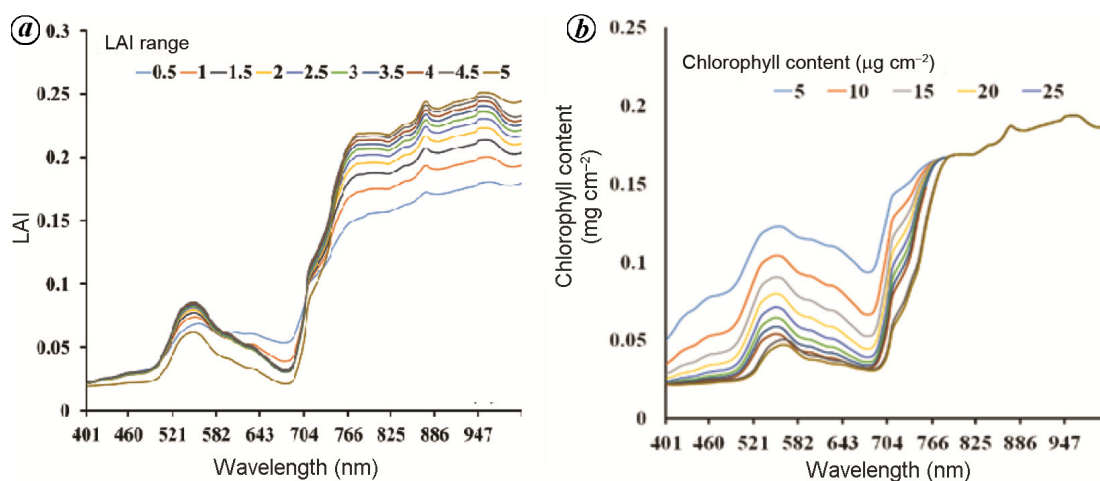

Figure 10. Sensitivity analysis of canopy radiative transfer model for (*a*) Leaf Area Index (LAI) and (*b*) chlorophyll content.

crops. LAI and C_{ab} varied from 1 to 7 and 10 to 50 μg cm^{-2} respectively (Figure 11 *e* and *f*). The validation from *in situ* data showed deviation of 19.75% and

23.05% from mean data for LAI and C_{ab} respectively (Figure 12). The root mean square error and percentage of deviation from mean was high in heterogeneous agricultural
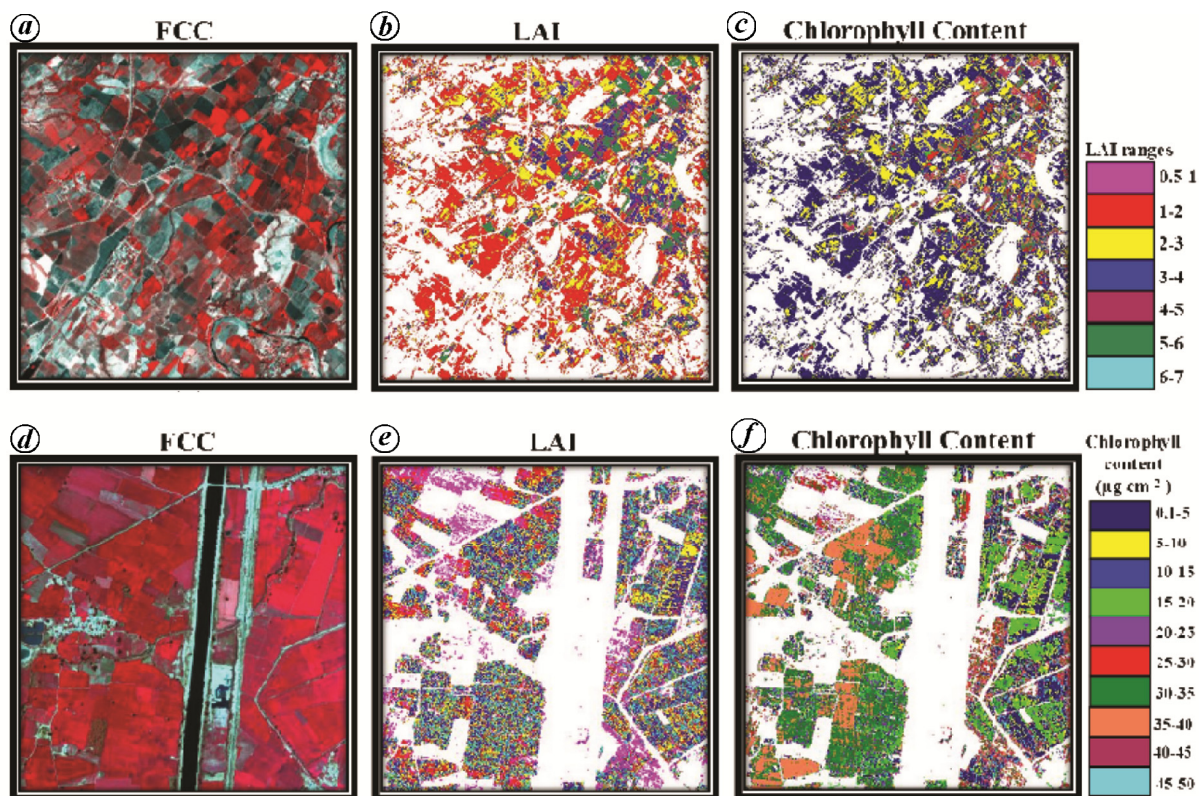
Figure 11. Spatial distribution of LAI and chlorophyll content at (*a*–*c*) heterogeneous and (*d*–*f*) homogeneous agricultural regions.
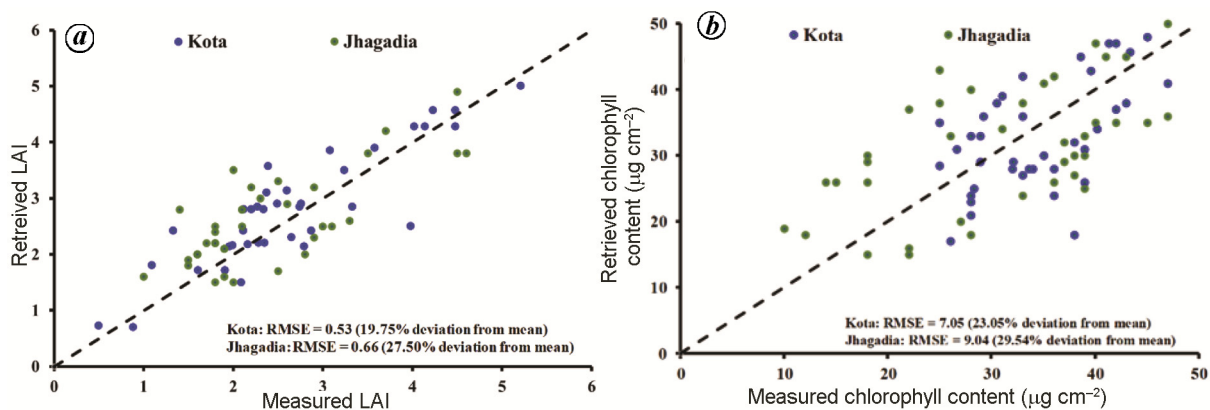

Figure 12. Validation of retrieved LAI with ground-measured LAI (*a*) and chlorophyll content (*b*) over homogeneous and heterogeneous agricultural areas.

area due to low canopy density leading to mixing of soil background reflectance with crop and more than one crop within a pixel.

To retrieve crop nitrogen correlation between measured N at Anand site and narrow band indices from AVIRIS-NG was computed. The correlation with different indices varied from –0.3 to 0.44. Highest correlation of 0.44 was found with photochemical reflectance index (PRI)[34] followed by normalized difference nitrogen index (NDNI; 0.4)[35]. The NDWI and carotenoid reflectance index

(CRIndex) 1 and 2 showed negative correlation[36]. These indices were further used to develop the multivariate model that resulted in R^2 of 0.81. The validation with independent dataset showed R^2 of 0.71. The model structure is given in eq. (2).

$$\text{Plant N content (\%)} = a_1 * X_1 + a_2 * X_2 + \cdots + a_{12} * X_{12}, \quad (2)$$

where X_1 is NDVI, X_2 is SRI (simple ratio index)[37]; X_3 is EVI (enhanced vegetation index)[38], X_4 is ARVI
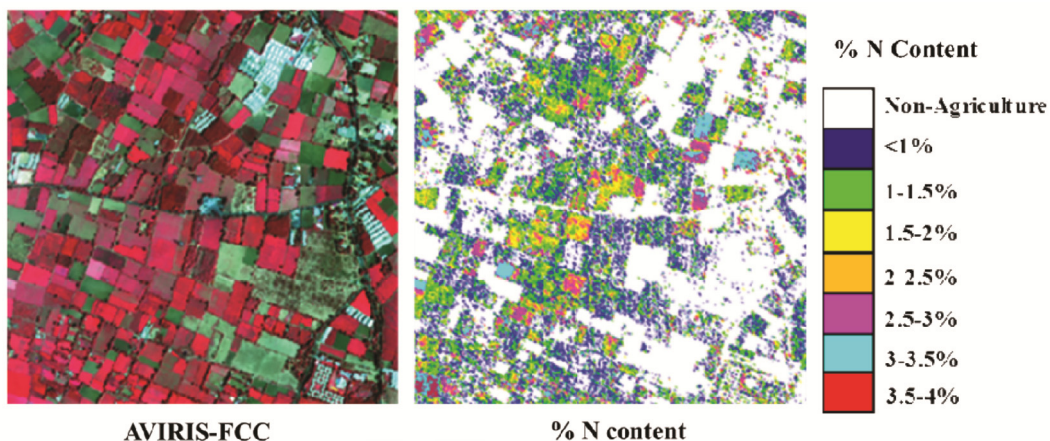
Figure 13. Spatial distribution of % N content over agricultural area of Anand site.
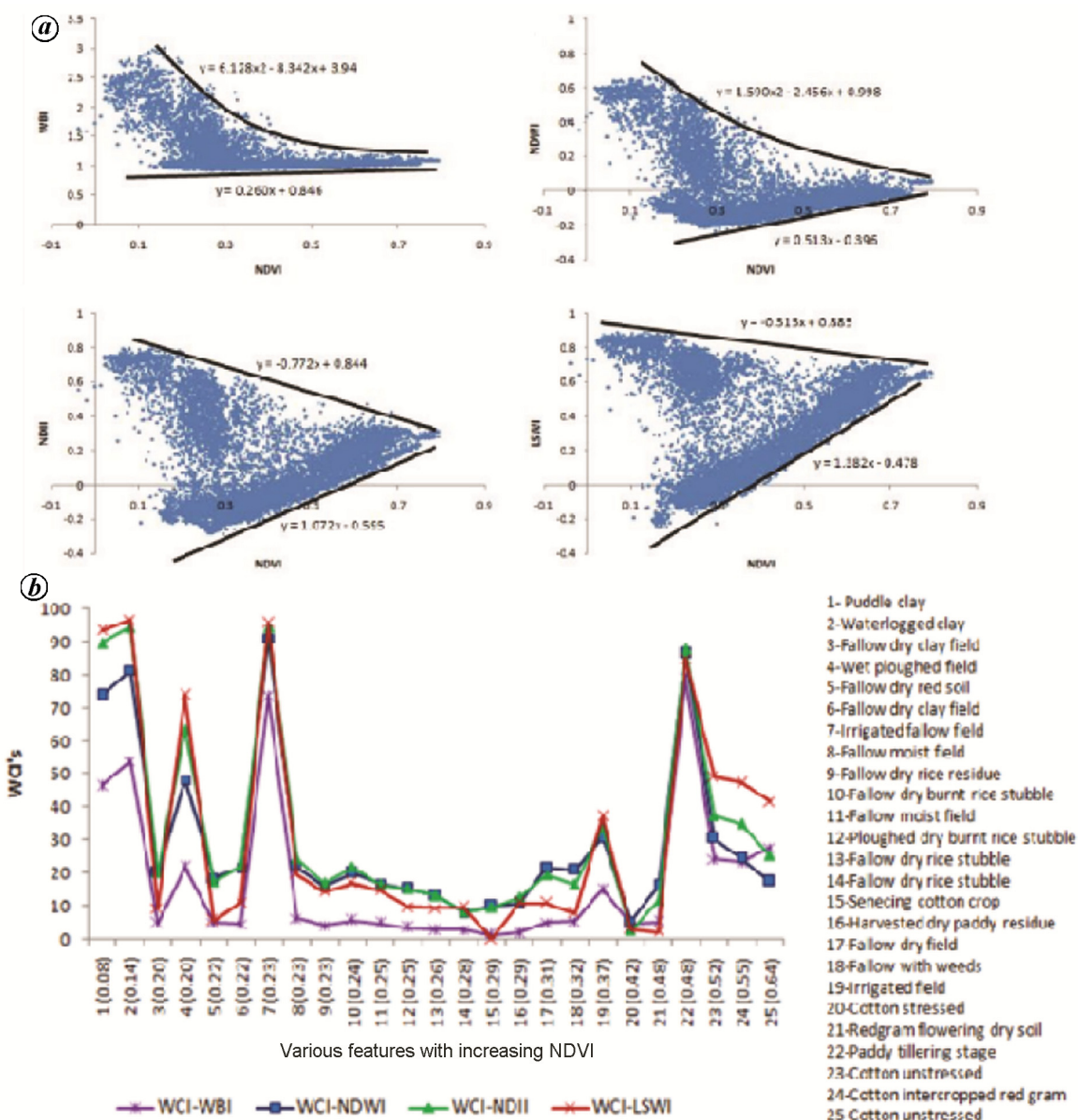

Figure 14. *a*, Scatter plot of NDVI versus WBI, NDWI, NDII and LSWI. *b*, WBI, NDWI, NDII and LSWI plotted for different ground-truth features.
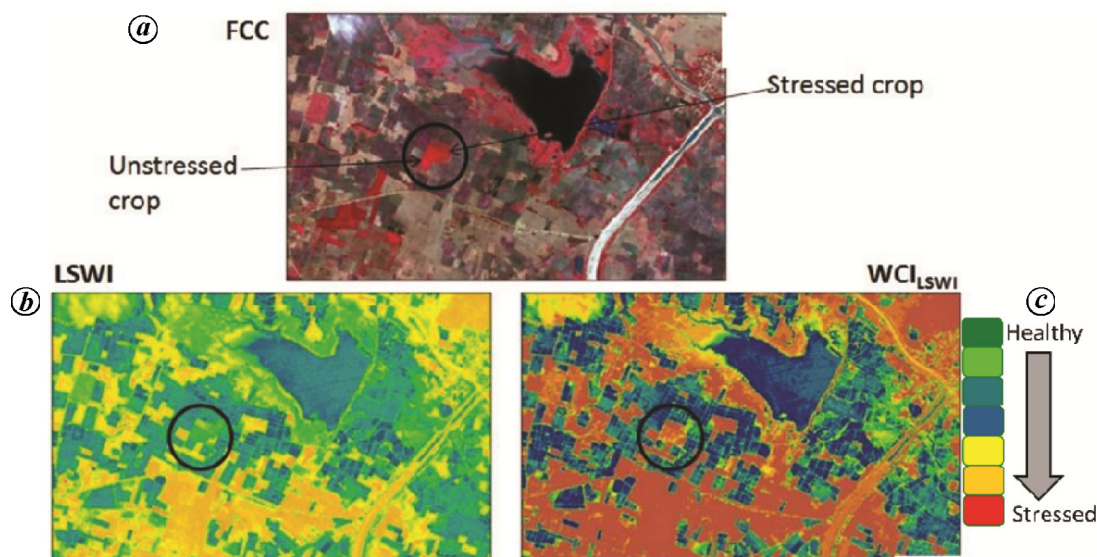
Figure 15. Discrimination of stress and unstressed crop by WCI$_{LSWI}$.

(atmospherically resistant vegetation index)[39], X_5 is modified red edge SRI[40], X_6 is Vogelmann red edge index – 1 (ref. 41), X_7 is REPI (red edge position index)[42], X_8 is PRI, X_9 is NDNI, X_{10} is CRIndex1, X_{11} is CRIndex2 and X_{12} is NDWI. a_1, a_2 ,…, a_{12} are the coefficients of X_1, X_2, … X_{12}. The above model was applied to the 12 indices to generate spatial plant N content map from AVIRIS data (Figure 13). The plant N content in different crops varied from 0.5% to 4% of the plant dry weight.

Abiotic stress

In Nagarjuna Sagar command area, transplanted rice field (waterlogged) had the lowest NDVI; however, it had the highest value in all the water indices. In order to capture the continuous change in water indices for a given NDVI, portion of the study area was chosen which exhibited wide range of moisture conditions and crop cover. The corresponding values of NDVI were plotted against WBI, NDWI, NDII and LSWI (Figure 14 a). The shape of the scatter plot between NDVI and water indices was similar to the triangular space of LST and NDVI[43]. In NDII–NDVI triangles, slope of the wet and dry edge was much steeper wet edge than LSWI–NDVI. This may be due to limitations of the LSWI in mixing the response between the wet surface and healthy crop. In this study NDVI–NDII triangle has been used during all stages of the crop to discriminate wet and healthy crops.

The ground observations showed that WBI and NDWI were more sensitive to water content in the soil compared to crop water content at different NDVI values. The normalization of these indices was carried out by deriving the WCI. One of the important assumptions in the estimation of WCI is that the soil moisture and vegetation water

content are the main contributing factors for variation in all WIx. The dry and wet edges were computed using the scatter plots of NDVI and WIx. Using the wet and dry edge with current NDVI values WCI is computed for the study area. The main advantage of deriving the WCIs is to bring the four indices to a common scale (Figure 14 b). Among the WCIs, WCI-LSWI was able to explain the variations from dry soil/crop conditions to the wet saturated soil and healthy crop conditions. The use of WCIs, helped in overcoming the saturation of WBI and NDWI at higher NDVI. This technique also helped in discriminating the healthy crops from the stressed crops at peak vegetative stage. The LSWI showed whole area as healthy crop area as shown in Figure 15 even though it comprises healthy crop and irrigated fallow fields. The low WCI$_{LSWI}$ values (Figure 15 c), indicate that a part of the crop is under stressed condition while a small portion is healthy. The surrounding wet fallow field showed high WCI$_{LSWI}$ value and was able to discern all four cases of moisture condition, like dry fallow, wet fallow, healthy crop and stressed crop.

Biotic stress

The threshold of disease indices LRDSI-1 and LRDSI-2 in the range 7.5–7.9 and 8–8.4 respectively, derived from ground-observed spectra were used to discriminate yellow rust-infested wheat crop at Rupnagar site. Further, after continuum removal absorption band depth from ground-based and AVIRIS-NG spectra revealed that two characteristic spectral regions one each in visible (662–702 nm) and SWIR (2155–2175 nm) were able to discriminate infested crops (Figure 16 a). For visible domain 0.25–0.35; 0.20–0.23 and SWIR domain 0.07–0.14; 0.03–0.06,
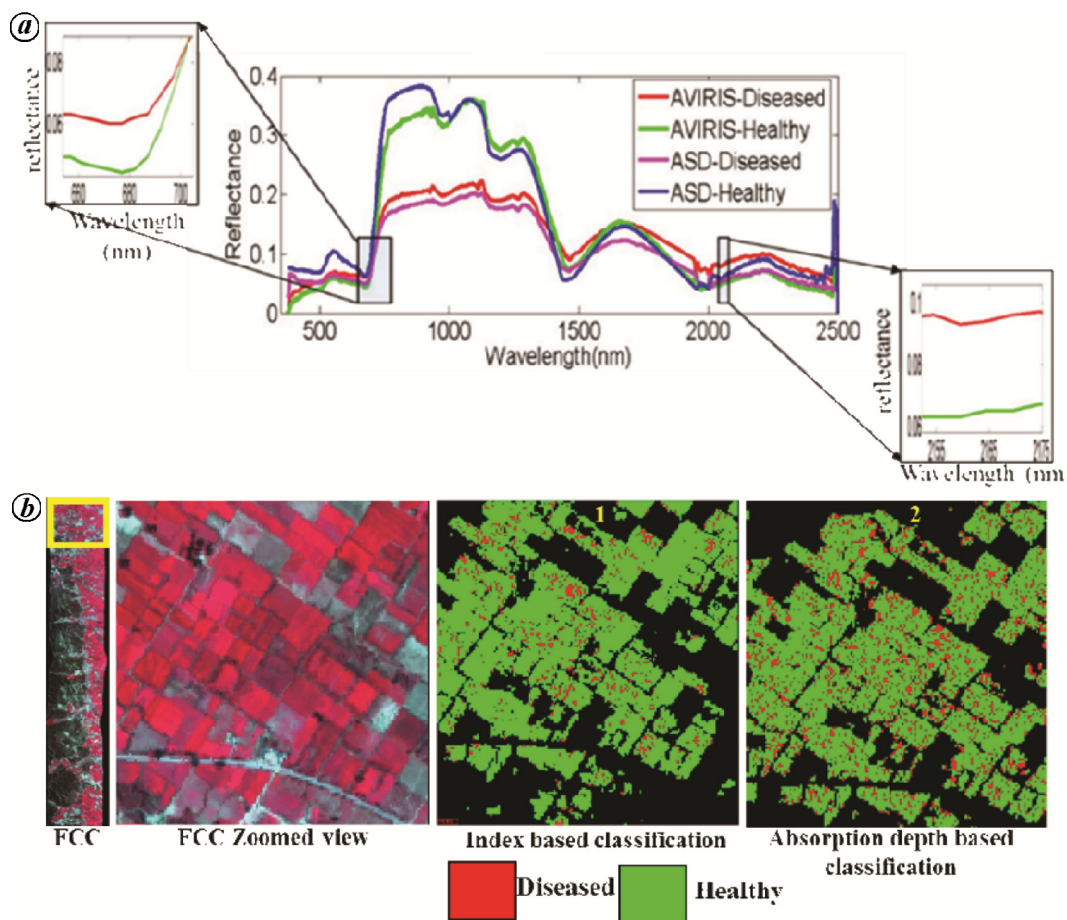
Figure 16. *a*, Selection of spectral band for wheat disease identification. *b*, Classification of healthy and diseased wheat crops using different classification techniques.

absorption depths were selected for healthy and disease-infested crops respectively. Based on the absorption depth, spatial distribution of healthy and disease-infested wheat crops was generated (Figure 16 *b*). The ground and classified image spectra were used to generate matching score by applying spectral feature fitting. This showed 0.91 and 0.93 scores for healthy and yellow rust-infested wheat crop respectively, for band depth-based classification. Some studies have also reported that wheat aphids-infested crops have low reflectance in near-infrared and high in visible compared to healthy crops and specific band centres at 694 nm and 800 nm respectively[44].

Conclusion

This study examined the performance of AVIRIS-NG hyperspectral narrow band data in many agricultural applications like crop-type discrimination, retrieval of crop biophysical and biochemical parameters, and crop stress assessment. The performance of hyperspectral data varied across homogeneous and heterogeneous agricultural systems. The low accuracy in heterogeneous agricultural area in discrimination and retrieval of crop parameters was due to low crop fraction or overlapping of two crops within a pixel leading to mixing of spectral signature of soil background and other crop with the dominated targeted crop at finer spatial scale. This further alters the unique spectral signal of a particular crop and decreases classification accuracy. In future, modelling of sub-pixel heterogeneity using linear and non-linear approach will improve the accuracy of classification[45]. The study clearly showed that hyperspectral data provide better classification accuracy compared to multispectral LISS IV data in different agricultural settings. It is observed that high-dimensional nature of hyperspectral data introduces many limitations in supervised classifiers, such as the limited availability of training samples, since in order to obtain statistically reliable results, the amount of training data needed to support the results often grows exponentially with dimensionality. Thus, data reduction techniques such as PCA, MNF and DNN provide better accuracy for crop classification. The forward and inversion of canopy radiative transfer model using AVIRIS-NG increase the retrieval accuracy of LAI and C_{ab} over different crop types. The plant N content also showed good retrieved accuracy with multivariate linear modelling of

specific narrow band hyperspectral indices. The unique characteristics of narrow band indices and absorption features in specific bands have been used to discriminate biotic and abiotic stresses. The measurements from AVIRIS-NG at finer scale provided unique absorption features of crop biochemistry. Airborne measurements are largely preserved the variation in spectral shape due to biochemical constituents of each crop type and therefore, could be used to discriminate crop types and retrieval of biophysical and biochemical contents. Despite the promising results obtained in this study, substantial challenges still remain for extensive applications of imaging spectroscopy to quantify all important crop pigments responsible for crop growth and development. The quantification of these pigments or biochemical constituents will provide a pathway for discrimination of crop type and different stresses.

1. Thenkabail, P. S., Enclona, E. A., Ashton, M. S. and Van Der Meer, V., Accuracy assessments of hyperspectral waveband performance for vegetation analysis applications. *Remote Sensing Environ.*, 2004, **91**, 354–376.

2. Meerdink, S. K. *et al.*, Linking seasonal foliar traits to VSWIR-TIR spectroscopy across California ecosystems. *Remote Sensing Environ.*, 2016, **186**, 322–338.

3. Gitelson, A., In *Hyperspectral Remote Sensing of Vegetation* (eds Thenkabail, P. S., Lyon, G. J. and Huete, A.), CRC Press-Taylor and Francis Group, Boca Raton, FL, USA, 2011, pp. 141–166.

4. Ozdogan, M. and Woodcock, C. E., Resolution dependent error in remote sensing of cultivated areas. *Remote Sensing Environ.*, 2006, **103**, 203–217.

5. Asner, G. P. and Martin, R. E., Spectral and chemical analysis of tropical forests: scaling from leaf to canopy levels. *Remote Sensing Environ.*, 2008, **112**, 3958–3970.

6. Jacquemoud, S., Baret, F., Andrieu, B., Danson, F. M. and Jaggard, K., Extraction of vegetation biophysical parameters by Inversion of the PROSPECT + SAIL models on sugar beet canopy reflectance data. Application to TM and AVIRIS sensors. *Remote Sensing Environ.*, 1995, **52**, 63–172.

7. Zarco-Tajada, P. J., Guillin-Climent, M. L., Hernandez-Clemente, R. and Catalina, A., Estimating leaf carotenoid content in vineyards using high resolution hyperspectral imaging acquired from an unmanned aerial vehicle (UAV). *Agric. For. Met.*, 2013, **17**, 281–294.

8. Green, R. O. *et al.*, Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS). *Remote Sensing Environ.*, 1998, **65**, 227–248.

9. Turner, D. P., Ollinger, S., Smith, M.-L., Krankina, O. and Gregory, M., Scaling net primary production to a MODIS footprint in support of Earth observing system product validation. *Int. J. Remote Sensing*, 2004, **25**, 1961–1979.

10. Bremner, J. M., Determination of nitrogen in soil using Kjeldahl method. *J. Agric. Sci.*, 1960, **55**, 11–38.

11. Aguilar, M. A., Saldaña, M. M. and Aguilar, F. J., GeoEye-1 and WorldView-2 pan-sharpened imagery for object-based classification in urban environments. *Int. J. Remote Sensing*, 2012, **34**, 2583–2606.

12. Yonezawa, C., Maximum likelihood classification combined with spectral angle mapper algorithm for high resolution satellite imagery. *Int. J. Remote Sensing*, 2007, **28**, 3729–3737.

13. Paola, J. D. and Schowengerdt, R. A., A review and analysis of backpropagationneural networks for classification of remotely

14. Vapnik, V. N., *The Nature of Statistical Learning Theory*, Springer Verlag, New York, USA, 1995.

15. Melgani, F. and Bruzzone, L., Classification of hyperspectral remote sensing images with support vector machine. *IEEE Trans. Geosci. Remote Sensing*, 2004, **8**, 1778–1790.

16. Krebel, U., Pairwise classification and support vector machines. In *Advances in Kernel Methods-Support Vector Learning* (eds Schölkopf, B., Burges, C. J. C. and Smola, A. J.), The MIT Press, Cambridge, MA, USA, 1999, pp. 255–268.

17. Green, A. A., Berman, M., Switzer, P. and Craig, M. D., A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Trans. Geosci. Remote Sensing*, 1988, **26**, 65–74.

18. Singh, A. and Harrison, A., Standardized principal components. *Int. J. Remote Sensing*, 1985, **6**, 883–896.

19. Friedl, M. A. and Brodley, C. E., Decision tree classification of land cover from remotely sensed data. *Remote Sensing Environ.*, 1997, **61**, 399–409.

20. Clark, R. N. *et al.*, Imaging spectroscopy: earth and planetary remote sensing with the USGS tetracorder and expert systems. *J. Geophys. Res.*, 2003, **108**, 5131; doi:10.1029/2002JE001847.

21. Chen, Y., Lin, Z., Zhao, X., Wang, G. and Gu, Y., Deep learning-based classification of hyperspectral data. *IEEE J. Appl. Earth Obs. Remote Sensing*, 2014, **6**, 2094–2107.

22. Kussul, N., Lavreniuk, M., Skakun, S. and Shelestov, A., Deep learning classification of land cover and crop types using remote sensing data. *IEEE Geosci. Remote Sensing Lett.*, 2017, doi:10.1109/LGRS.

23. Jacquemoud, S. *et al.*, PROSPECT + SAIL models: a review of use for vegetation characterization. *Remote Sensing Environ.*, 2009, **113**, S56–S66.

24. Barker, D. M., Huang, W., Guo, Y. R., Bourgeois, A. J. and Xiao, Q. N., A three-dimensional variational data assimilation system for MM5: implementation and initial results. *Mon. Weather Rev.*, 2004, **132**, 897–914.

25. Rouse, J. W., Has, R. H., Schell, J. A., Deering, D. W. and Harlan, J. C., Monitoring the vernal advancement of retrodegradation of natural vegetation, NASA/GSFC, Type III, Final report, Greenbelt, MD, 1974, p. 371; Rouse, J. W., Haas, R. S., Schell, J. A. and Deering, D. W., Monitoring vegetation systems in the Great Plains with ERTS. In Proceedings, 3rd Earth Resources Technology Satellite Symposium, 1973, vol. 1, pp. 48–623.

26. Peñuelas, J., Gamon, J. A., Fredeen, A. L., Merino, J. and Field, C. B., Reflectance indices associated with physiological changes in nitrogen and water limited sunflower leaves. *Remote Sensing Environ.*, 1994, **48**, 135–146.

27. Hardisky, M. A., Klemas, V. and Smart, R. M., The influences of soil salinity, growth form, and leaf moisture on the spectral reflectance of *Spartina alterniflora* canopies. *Photogramm. Eng. Remote Sensing*, 1983, **49**, 77–83.

28. Gao, B.-C., NDWI – a normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing Environ.*, 1996, **58**, 257–266.

29. Chen, D., Huang, J. and Jackson, T. T., Vegetation water content estimation for corn and soybeans using spectral indices derived from MODIS near- and short-wave infrared bands. *Remote Sensing Environ.*, 2005, **98**, 225–236.

30. Ashourloo, D., Mobasheri, M. R. and Huete, A., Developing two spectral disease indices for detection of wheat leaf rust (*Puccinia triticina*). *Remote Sensing*, 2014, **6**, 4723–4740.

31. Bolstad, P. V. and Lillesand, T. M., Semi-automated training approaches for spectral class definition. *Int. J. Remote Sensing*, 1991, **13**, 3157–3166.

32. Clark, M. L., Comparison of simulated hyperspectral HyspIRI and multispectral Landsat 8 and Sentinel-2 imagery for multi-seasonal,

sensed multi-spectral imagery. *Int. J. Remote Sensing*, 1995, **16**, 3033–3058.

regional land-cover mapping. *Remote Sensing Environ.*, 2017, **200**, 311–325.

33. Gitelson, A. A., Kaufman, Y. J., Stark, R. and Rundquist, D., Novel algorithm for remote estimation of vegetation fraction. *Remote Sensing Environ.*, 2002, **80**, 76–87.

34. Gamon, J. A., Penuelas, J. and Field, C. B., A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sensing Environ.*, 1992, **41**, 35–44.

35. Serrano, L., Penuelas, J. and Ustin, S. L., Remote sensing of nitrogen and lignin in mediterranean vegetation from AVIRIS data: decomposing biochemical from structural signals. *Remote Sensing Environ.*, 2002, **81**, 355–364.

36. Gitelson, A. A., Zur, Y., Chivkunova, O. B. and Merzlyak, M. N., Assessing carotenoid content in plant leaves with reflectance spectroscopy. *Photochem. Photobiol.*, 2002, **75**, 272–281.

37. Birth, G. S. and McVey, G. R., Measuring colour of growing turf with a reflectance spectrometer. *Agron. J.*, 1968, **60**, 640–649.

38. Huete, A. R., Liu, H., Batchily, K. and van Leeuwen, W., A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote Sensing Environ.*, 1997, **59**, 440–451.

39. Kaufman, Y. J. and Tanre, D., Strategy for direct and indirect methods for correcting the aerosol effect on remote sensing: from AVHRR to EOS-MODIS. *Remote Sensing Environ.*, 1996, **55**, 65–79.

40. Sims, D. A. and Gamon, J. A., Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sensing Environ.*, 2002, **81**, 337–354.

41. Vogelmann, J. E., Rock, B. N. and Moss, D. M., Red edge spectral measurements from sugar maple leaves. *Int. J. Remote Sensing*, 1993, **14**, 1563–1575.

42. Curran, P. J., Windham, W. R. and Gholz, H. L., Exploring the relationship between reflectance red edge and chlorophyll concentration in slash pine leaves. *Tree Physiol.*, 1993, **15**, 203–206.

43. Sandholt, I., Rasmussen, K. and Anderson, J., A simple interpretation of the surface temperature/vegetation index space for assessment of the surface moisture status. *Remote Sensing Environ.*, 2002, **79**, 213–224.

44. Mirik, M., Michels, G. J., Kassymzhanova-Mirik, S. and Elliott, N. C., Reflectance characteristics of Russian wheat aphid (Hemiptera: Aphididae) stress and abundance in winter wheat. *Comp. Electron. Agric.*, 2007, **57**, 123–134.

45. Xu X., Li, J., Wu, C. and Plaza, A., Regional clustering-based spatial preprocessing for hyperspectral unmixing. *Remote Sensing Environ.*, 2018, **204**, 333–346.