Detection of rice leaf folder, *Cnaphalocrocis medinalis* (Guenée) (Lepidoptera: Crambidae) infestation using ground-based hyperspectral radiometry

Bhubanananda Adhikari[1,2], Radhakrushna Senapati[1], Minati Mohapatra[2], Laxminarayan Mohapatra[2], Rahul Nigam[3] and Shyamaranjan Das Mohapatra[1,*]

[1]ICAR-National Rice Research Institute, Cuttack 753 006, India
[2]Odisha University of Agriculture and Technology, Bhubaneswar 751 003, India
[3]Space Application Centre, Indian Space Research Organisation, Ahmedabad 380 015, India

Hyperspectral remote sensing is a useful technique for detecting spatio-temporal changes in crop morphological and physiological health. In order to identify the pest-sensitive bands for rice leaf folder (RLF), the ground-based hyperspectral data were recorded at varying damage levels. The first- and second-order derivatives were correlated with correlation coefficient *r* and per cent leaf damage. The common region identified were recognized as sensitive regions (508–529, 671–680, 721–759, 779–786 and 804–820 nm). The absorption dips were also found using Sensitivity and Continuum Removal Analysis. Combining all, a total of nine spectral bands (518, 549, 661, 674, 678, 731, 789, 816 and 898 nm) were identified as pest-sensitive bands for RLF. The feature-selection method was employed using RELIEFF algorithm to find out the band combinations and bands 518, 661 and 731 nm yielded maximum accuracy of 81.67%.

Keywords: Hyperspectral sensing, rice leaf folder, sensitive spectral bands, spectroradiometer.

RICE plays a vital role in India's food security and economy. Among the various factors reducing the rice yield, biotic stress is the major one. The prevalence of disease and insect pests are two of the biological factors that have the greatest impact on rice yield[1]. Currently, rice farmers in India are exposed to health risks due to the application of large amounts of pesticides to prevent crop damage due to pests. However, crop loss due to pest attacks has increased significantly during the last few decades. To reduce the use of pesticides in crop protection and make agriculture more environment-friendly and sustainable, early detection of pest damage is essential for timely pest management measures. This is also crucial to accurately and confidently monitor the location, range and severity of crop pests and diseases. Identifying visual symptoms has frequently allowed acceptable precision assessments of crop pests and diseases[2]. The visual assessment methods, which rely on the farmers' skill are overly subjective, labour-intensive and time-consuming. The manual detection of crop pests and diseases may be replaced with an efficient remote sensing (RS) technique. RS can be used to track the spatio-temporal changes in crop morphology and physiological status due to pest damage over a large area in a short period. When employing remote sensing data to identify crop pests and diseases, two strategies are frequently used[3,4]. To examine the spectral changes of crop pigment, water content and canopy structure under the pressures of pests or diseases, one of the techniques uses hyperspectral spectroscopy or imaging. The other method uses a variety of remote sensing data to retrieve habitat information, such as land surface temperature, soil water content and crop growth, in order to further predict the spatial distribution of pests or diseases. The occurrence of pests and diseases requires suitable habitat conditions[5–8]. The ability of these technologies to practically direct the precise application and reduction of pesticides may be constrained.

According to estimates, diseases and insect pests cause farmers to incur losses to 37% of their rice production each year[9]. Among the rice leaf-feeding insects, the rice leaf folder (RLF) *Cnaphalocrocis medinalis* (Guenée), once regarded as a minor pest, has now attained a major pest status in several parts of the Indian continent, causing significant loss to the rice crop. It causes visible damage to the rice crop by folding leaves and scraping off the green mesophyll tissue. Numerous rice leaves are defoliated or their chlorophyll removed by RLF larvae, which prevents photosynthesis. Frequent feeding results in the green leaves turning yellow, curled, or stunted[10]. The larvae feed by scrapping off the green mesophyll tissues within the folded leaves after folding the leaf margins lengthwise. This feeding results in a narrow, horizontal white stripe. Membranous patches appear as a result of the injury. Significant loss in

*For correspondence. (e-mail: s.dasmohapatra@icar.gov.in)

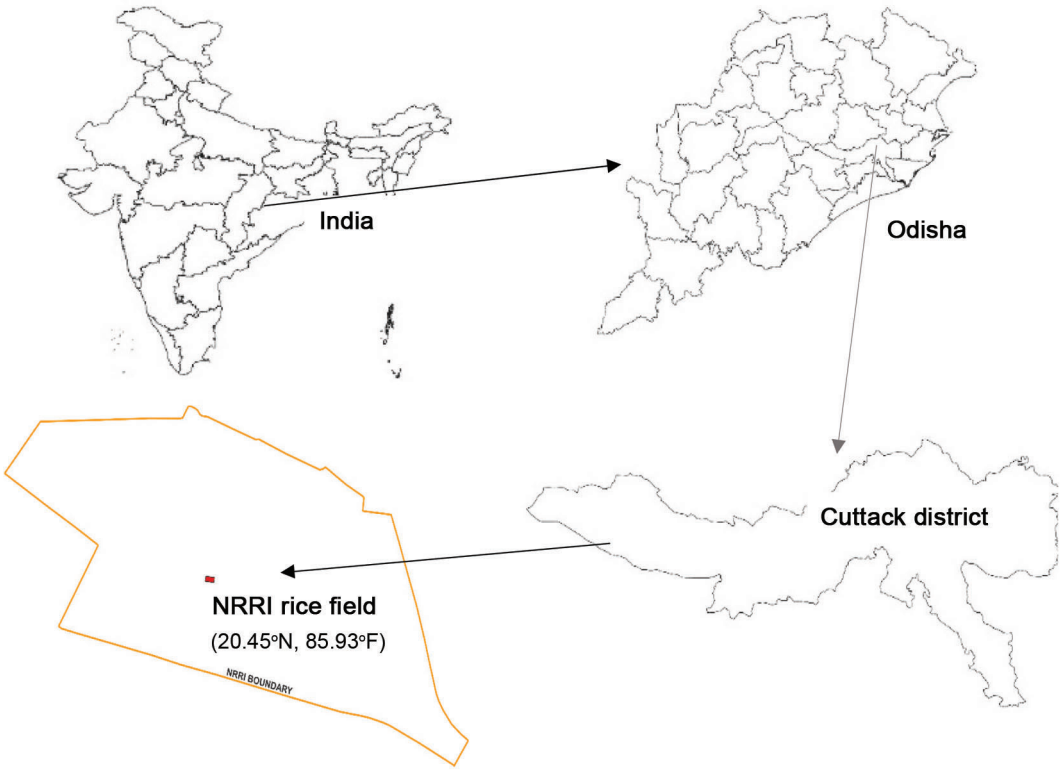

Figure 1. Data recording site.

leaf chlorophyll and carotenoids occurs due to RLF infestation[11]. Majority of early second-instar larvae are gregarious in nature. From the late second instar stage, the larva feeds on the folded leaves and solitude. The loss in yield due to the leaf folder outbreak is between 30% and 80% (ref. 12). Although they can emerge at any point in the life cycle of the rice crops, leaf folder infestations tend to be more prevalent during the reproductive and ripening periods[13].

The information-rich hyperspectral (HS) data comes with challenges such as information redundancy removal and optimal information identification. Spectral band selection depends on the input data properties and criteria defined to discriminate the most optimal feature subset of bands. Identification of the most informative spectral bands in the HS data relies on statistical methods like partial least squares in conjunction with techniques like correlation coefficient analysis, elimination of non-informative variables, stepwise regression variable selection and exhaustive band combination[14]. Most works in band selection associated with HS RS are related to classification problems, while relatively fewer works deal with regression problems considering vegetation characteristic properties. Among six types of HS band (feature) selection methods for classification problems, viz. ranking-based, searching-based, clustering-based, sparsity-based, embedding learning-based and hybrid scheme-based, the sparsity-based methods performed best in terms of accuracy, but ranking-based methods were more suitable for large HS datasets due to low complexity[15]. Feature selection methods are also classified based on search organization, sub-setting and evaluation methods like filter, wrapper and embedded methods[16]. Filter methods are shown to be faster and better suited to high-dimensional datasets. Estimation of the number of bands for band selection is a challenge as well because less number of bands will not allow enough spectral information to be preserved within a selection, whereas more bands cause band redundancy.

Although a few researchers primarily collected indoor- or field-HS data using ground-based spectroscopic sensors to study the viability, sensitive bands and usable spectral indices for identifying agricultural pests and damages[7,8], studies on RLF are meagre. Hence, the present study aimed at identifying the RLF-sensitive electromagnetic spectral bands from the optical spectral region to assess pest-infested crop areas in the targeted location.

Materials and methods

Study site

For the present study, an experiment was conducted at the ICAR-National Rice Research Institute (NRRI), Cuttack, Odisha, India (20.45°N, 85.93°E) during the rainy season of 2019 to generate different levels of leaf folder damage in randomized block design with four replications over rice crop (Figure 1). Insect occurrence is subject to suitable

Table 1. Different levels of rice leaf folder infestation at National Rice Research Institute based on Standard Evaluation System (SES)[18]

Crop stage	Scale	Damaged leaves (%)	Larvae released per microplot	Leaf damage (%) in the microplot
Vegetative	0	No damage	0	0
	1	1–10	10	4
	3	11–20	20	17
	5	21–35	40	34
	7	36–50	60	48
	9	51–100	100	86

weather parameters, as they regulate the presence of pests in the field. Weather parameters are known to have a significant impact on the occurrence, growth and development, and population build-up of insect pests in crop ecosystems, and thus on the extent of crop damage and yield loss. Before predicting the presence of pests, it is necessary to know the relationship between the prevalence and accumulation of various insect pests and their natural enemies and weather parameters[17].

Climatic conditions of the study site

The NRRI research farm has a subtropical climate with hot, wet summers and cool, dry winters. The weather conditions did not differ substantially during the study period. The total rainfall was 2292.5 mm during 2019–20 (July to December). The average sunshine hours per day was 4.57 h in 2019–20 (July to December). The average relative humidity was about 81.42%. Maximum, minimum and average temperatures ranged from 21.6°C to 36.2°C, 11.0°C to 27.3°C and 17.5°C to 30.15°C respectively, during the study.

Test insect

Leaf folder colony was maintained on the susceptible rice variety Taichung Native 1 (TN 1) under greenhouse conditions (27° ± 5°C temperature and 70% ± 10% relative humidity). Ten adult pairs were released separately for oviposition on 20- to 25-day-old TN 1 plants covered with mylar cage (75 cm height and 30 cm diameter). Honey solution (25%) was provided as a source of food for the adult moths. They lay eggs singly or in rows and mostly concentrated at the tip of the leaves. Potted TN 1 plants with eggs were separated and kept for further development. Eggs hatched in 4–5 days with 90% viability. Fresh TN 1 plants were provided as and when necessary for the transfer of larvae. Neonate and third-instar larvae used in the present study were taken from this stock culture.

Experimental set-up

Twenty two-day-old rice seedlings of TN 1 rice variety were transplanted in microplots (1 m × 1 m) with six levels of infestations (Table 1). The recommended dose of fertilizers was applied following standard agronomic practices to raise good crops. The experimental microplots were covered with nylon nets to protect the plants from other non-target pests. Variable numbers of third-instar larvae of RLF were released in the microplots for the desired level of leaf folder damage. The International Rice Research Institute's (IRRI) Standard Evaluation System (SES) was followed[18]. The whole flag leaf area of rice plants was considered as 100%; if one-fourth area was damaged, it was considered as 25%. If half the flag leaf area was damaged, it was considered as 50%, and so on.

Based on the visual symptoms on leaves (Figure 2), the sampled plants were graded into five levels of infestation, starting from score 0 (healthy), score 1 (low infestation), to score 9 (severe damage) (Table 1). In addition, using a graphical method, the percent leaf damage for various samples was calculated from RLF scale 1 as 4. It significantly increased as the damage level increased, reaching 86% for the RLF scale 9 (ref. 19).

Data collection and pre-processing

A portable spectroradiometer (FieldSpec3, Analytical Spectral Devices (ASD), Boulder, CO, USA) was used to record the spectral reflectance of rice crops with differential leaf folder damage at 1 nm intervals from 350 to 2500 nm. The instrument was remotely connected with ASD software installed in a laptop, which captured the reflectance data. Before recording the reflectance data from the rice leaves, the instrument was calibrated using a white reference panel named (Spectralon; Labsphere, Inc., North Sutton, NH, USA). The sensor's probe fitted in the pistol grip was kept at the height of about 50 cm above the leaf surface. Observations from all microplots were made in bright sunshine hours (often between 10.00 and 14.00 h) from fixed sites[20]. The RLF often damages the rice crop at the plant booting stage. Hence, in this stage, spectral reflectance from the microplots was recorded for each waveband during the cropping season. Before analysis, data pre-processing was done to remove the atmospheric noise perturbations between 1350 and 1450, 1780 and 2000, 2350 and 2500 nm (ref. 21).

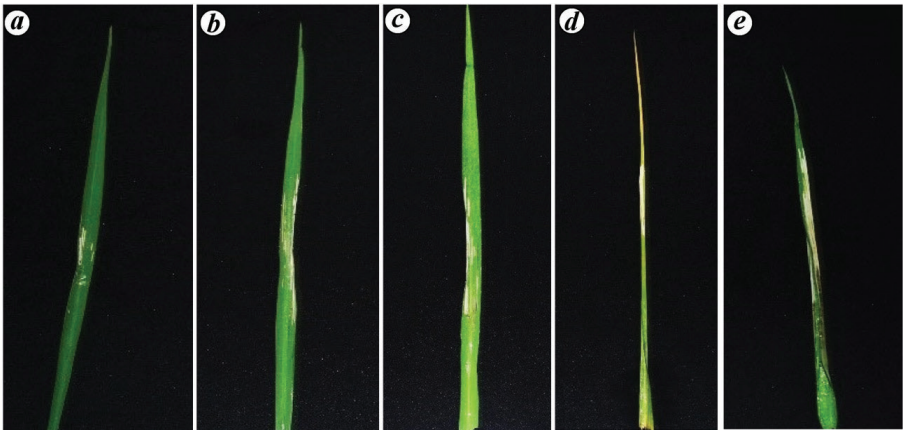

Figure 2. Field photographs of leaf folder-infested sample at different levels of damage. (***a***) LF scale 1, 1–10% damage, (***b***) LF Scale 3, 11%–20% damage, (***c***) LF scale 5, 21%–35% damage, (***d***) LF Scale 7, 36%–50% damage and (***e***) LF scale 9, 51%–100% in accordance with the standard evaluation system (SES) for rice leaf folder[18].

Spectral band analysis using derivative and continuum removal

The first-order derivative spectrum explains how the slope of the original reflectance spectrum varies with wavelength. Similarly, the second-order derivative (eq. (2)) determines the slope of the first derivative (eq. (1)), or the rate of change of slope of the spectrum. Accordingly, spectral derivatives are calculated from spectral reflectance data to reduce the unpredictability caused by variations in illumination or soil/land reflectance[22,23], and to detect the change in curvature of the spectral band.

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{f(x+\mathrm{d}x) - f(x)}{\mathrm{d}x}, \tag{1}$$

$$\frac{\mathrm{d}y^2}{\mathrm{d}x^2} = \frac{\mathrm{d}}{\mathrm{d}x}\frac{(dy)}{(dx)}. \tag{2}$$

When the spectral sample size is small, the first-order derivative can effectively capture the identification of various spectral features because, on increasing the sample size normalization may occur, which in turn can reduce the power of the HS data[24–26]. According to several studies, first-order derivative spectra at the red edge have two or more peaks which can lead to a bimodal distribution of the red edge positioning (REP) values that correlate to low and high chlorophyll (Chl) concentration[27,28]. As a result, REP often jumps at a specific Chl threshold, prohibiting the formation of REP and Chl predictive correlations. The distinctive chloride absorption spectra could be a factor in the existence of several peaks in the first-order derivative spectra[29,30].

Derivative spectra were employed in qualitative and quantitative assessments of the spectral data by calculating the change in slope and curvature at various spectral regions. Qualitative data on pigment composition were generated based on the wavelength position of the absorption characteristics in the derivative spectra. We examined how to quantify the levels of healthy and RLF-damaged samples using second-order derivative peaks based on the change in curvature of the slope. Reflectance and spectral derivative data from healthy and RLF samples were correlated at various damage levels using Pearson's correlation method to identify the most critical bands. In the present study, the sensitive spectral regions were generated using derivative analysis with correlation coefficient determination, and the region that was common to both the first- and second-order derivatives with a correlation coefficient value greater than 85% was used to distinguish between the RLF damage from healthy leaf samples (Figure 3)[4,31,32].

Spectral band analysis using sensitivity analysis and continuum removal

To determine the sensitive spectral regions, the sensitivity analysis and continuum removal analysis were performed using the spectral regions of interest with absorption qualities. The difference in reflectance between healthy and RLF-damaged samples demonstrate how reflectance responds to variations in insect occurrence[33]. As a result, sensitivity is the ratio of reflectance changes between infected and healthy samples to the reflectance of the healthy sample[21].

$$\mathrm{SS} = \frac{R_\mathrm{d} - R_\mathrm{h}}{R_\mathrm{h}}, \tag{3}$$

where SS is the spectral sensitivity, R_d the reflectance of the RLF infested sample and R_h is the reflectance of the healthy sample.

The spectrum libraries were created after resampling, and the reproductive 'ENVI' was used to determine the

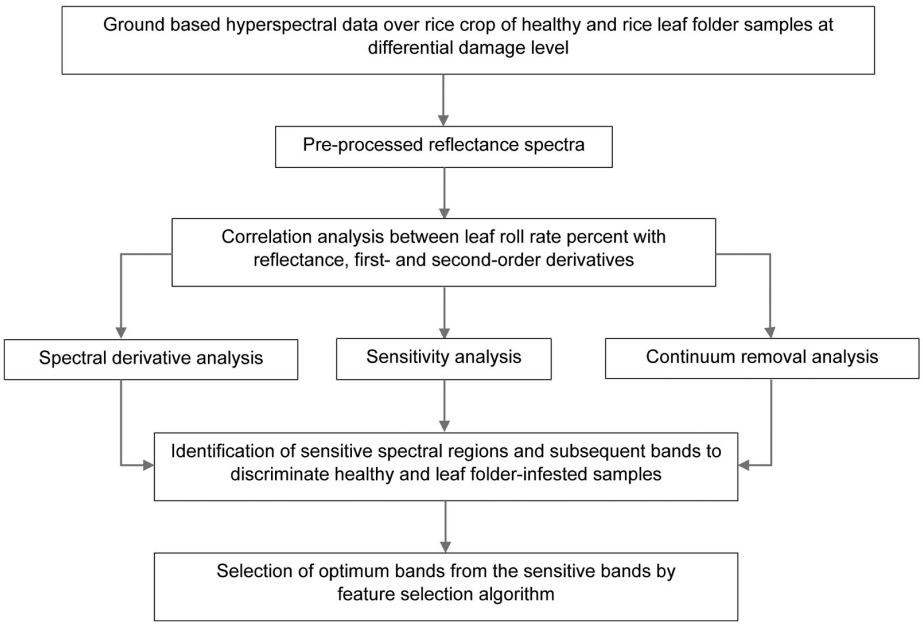

Figure 3. Flowchart of data analysis for leaf folder detection.

spectral signatures of healthy and RLF-infested samples at various damage levels. Before assessing absorption properties, it is best to eliminate a general concave shape of the spectrum. When the least sensitive parts with the least payload information were deleted, the most sensitive spectral region was developed. To quantify absorption features in the spectra, the generally concave contour of a spectrum should be removed. The absorption dips in the sensitive spectral areas were then found using ENVI. This procedure, also known as continuum elimination or the convex-hull transform makes it easier to compare spectra recorded using different equipment or under different lighting conditions. The continuum is a convex hull that fits over the top of a spectrum and connects local spectra maxima with straight-line segments. As the initial and final spectral data values are on the hull, the first and last bands of the output continuum removed data file are marked as 1.0. Equation (4) is a representation of the continuum equation.

$$S_{cr} = S/C, \tag{4}$$

where S_{cr} is the continuum removed spectra, S the original spectrum and C is the continuum curve.

When the continuum and spectra are aligned, the image spectrum equals 1.0, and it is less than 1.0 when absorption features are present. The spectral bandwidth with maximum absorption characteristics or dips was chosen as the best for continuum removal. The continuous removal method is critical for analysing the spectrum features of absorption dips. This normalization technique compares spectra obtained from different samples under different lighting conditions by modifying the convex continuum hull. The continuum has straight sections connecting its convex hull along the top of the spectrum to maximize local spectra. Absorption feature depths were measured by fitting a continuum to vegetation reflectance spectra[34]. This approach of neglecting the bands with lesser payload information to identify the prominent absorption dips can thereby cause normalization of the band depth. Despite the fact that derivative analysis gives us absorption peaks and dips based on the change in slope and curvature, the continuum removal approach scores over the derivative-based approach for assessing tree canopy biochemical characteristics utilizing airborne imaging spectrometry data[35].

In the present study, hyperspectral signatures of healthy and infested RLF samples were evaluated using a non-imaging spectroradiometer. Total 600 samples (300 for training and 300 for testing) with 50 samples each of healthy and RLF at all damage scales were used to find the best possible bands responsible for the detection of RLF using MATLAB. To reduce the data redundancy, the RELIEFF algorithm was used[36]. This technique has the advantage of appropriately estimating the quality of features with strong relationships.

Using RELIEFF, the significant and most relevant wavelengths and two band normalized changes from 400 to 900 nm describing the impact of the pest were retrieved from the dataset[37]. The optimally weighted combination of a single wavelength and a normalized wavelength difference was sought to determine the best bands responsible for the identification of RLF. The RLF samples were identified with higher accuracy and sensitivity. The search algorithm with the best performance depended on the database and the evaluation function, which were used to determine the

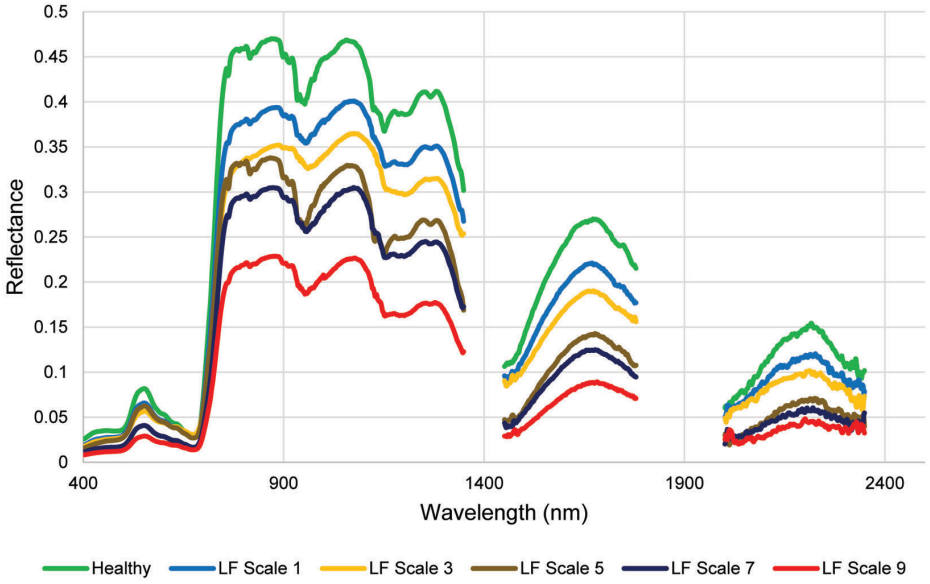

Figure 4. Typical ground-based spectral signature of rice crop with different severity levels of infestation due to rice leaf folder.

optimal feature subset for a particular database. Random backward sequential selection was used with at least two different error rate evaluation functions to determine the optimal set of features. As the forward selection identifies a weaker subset of features, the backward sequential feature selection method was used to find the most sensitive bands with higher accuracy.

Results

Differential damage of rice leaf folder

The RLF damage on leaves was quantified, and the damage scale was determined based on the per cent leaf damage (Table 1). According to the SES for the leaf folder by IRRI, the damage was assigned as 0 (healthy) followed by 1 (4%), 3 (17%), 5 (34%), 7 (48%) and 9 (86%) in the present study (Table 1). The leaf damage per cent was classified using the graphical mode of analysis, calculating the number of grids covered when the damaged sample was superimposed on graph paper. This result showed that the rate of leaf damage per cent is directly proportional to the level of injury.

Removal of noise

As depicted in Figure 4, water vapour absorption significantly influences spectral curves at 1350–1450, 1780–2000 and 2350–2500 nm. Water vapour absorption generated several anomalous values on the curves; nonetheless, the effective spectral curves were flattened. After removing the perturbated spectral bands from the ground-measured spectral curve, the curve gets flattened, which helps detect only sensitive bands from the near-infrared (NIR) region, i.e. between 700 and 900 nm. Nevertheless, any difference in the visible (VIS) region (400–700 nm) was obtained using spectral derivative analysis (SDA).

Identification of sensitive regions by spectral derivative analysis

The reflectance value and first- and second-order derivatives of reflectance for all samples are associated with the leaf damage per cent, and the correlation coefficient (R) of leaf damage with the above parameters was determined (Figure 5 *a–c*). The correlation coefficient against reflectance reached a maximum value of 60%, whereas when correlated against the first- and second derivatives, the correlation coefficient increased to more than 85%. The identification of sensitive regions is based on the correlation coefficient and if the value exceeds 85% with some observable changes in the spectral curve, then the spectral region will be considered as sensitive subject to the commonality of the first- and second-order derivatives. Based on the aforementioned criteria, 11 sensitive regions were identified after determining the correlation coefficients with the first- and second-order derivatives of reflectance. These sensitive regions are 508–540, 671–680, 691–761, 782–789, 794–820, 886–900 nm (for the first-order derivative) and 500–529, 670–695, 721–759, 779–786 and 804–820 nm (for the second-order derivative), found on the basis of correlation coefficient and leaf damage per cent. The common regions which were identified after finding the correlation coefficient between the first- and second-derivatives were 508–529, 671–680, 721–759, 779–786 and 804–820 nm.

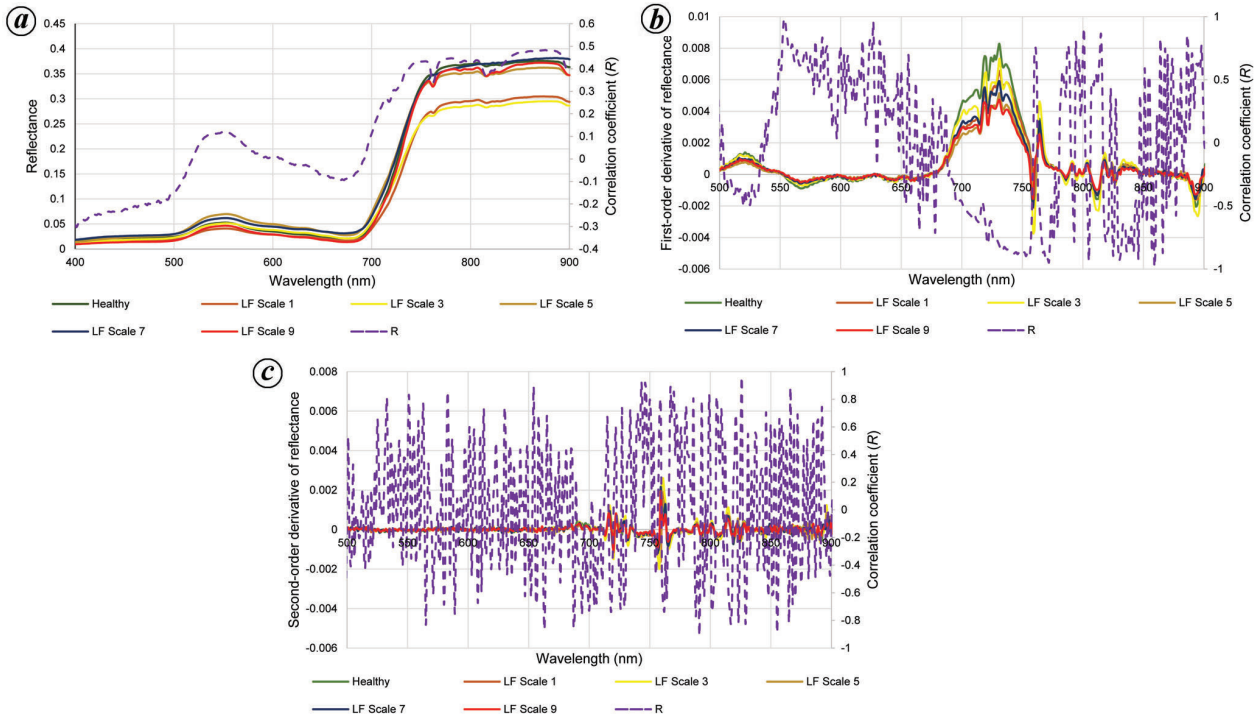

Figure 5 ***a–c.*** Spectral reflectance from healthy and infested leaves, and correlation coefficient (*R*) between (***a***) reflectance and per cent leaf damage, (***b***) between the first-order derivative and per cent leaf damage, and (***c***) between the second-order derivative and per cent leaf damage.

Identification of central wavelength

Spectral derivative analysis method: After identifying the sensitive region, the SDA was performed to determine the central wavelength (Figure 6 *b–f*), which was then incremented up to 100 nm to test for sensitivity using the unpaired *t*-test. The performance of the *t*-test for significance analysis was the basis of obtaining the central wavelength. The healthy rice leaves showed two reflectance peaks in the VIS and NIR regions at 518 and 731 nm respectively in the first-order derivative spectral curves. The red edge appeared between 670 and 750 nm, and the smallest reflectance spectrum was 661 nm. Two dips at 816 and 898 nm were also found in the NIR region, which could further be determined as sensitive bands (Figure 6 *d* and *e*). Based on these findings, the variation in spectral reflectance was significant in different damage levels. which is obvious in different spectral areas (VIS, NIR and short wave infrared (SWIR) regions). The spectral difference in the NIR range was the most apparent[38]. Due to changes in normal soil reflectance and atmospheric variables, a derivative spectral value containing the first- and second-orders can more precisely portray the spectral features of healthy and infested rice canopies. The analytical results revealed that as the damage level increased, the absolute values of the derivative spectra decreased since the reflectance value varies for healthy and RLF-affected samples. The reflectance value for healthy samples increased in both VIS and NIR regions in comparison to RLF samples, indicating that the derivative values should decrease as infestation increases.

Sensitivity analysis to predict sensitive bands: Sensitive bands are essential for hyperspectral remote sensing data in order to avoid data duplication, selection of useful bands and develop a conversion model for the prediction of vegetative indices. Hence precisely identifying the most sensitive bands or spectral regions is crucial. In the present study, sensitivity analysis and continuum removal method were used concurrently to identify sensitive bands and regions. The sensitivity analysis identifies peaks and dips, eventually predicting the sensitive band for RLF (Figure 6 *a*). The extent of shift decreases rapidly as infestation levels increase. The sensitivity values in the VIS and NIR regions of the rice leaf folder (damage score 7 and 9 respectively) were positive, suggesting that the reflectance of the affected rice with higher intensity was greater than that of infested rice leaves with lower intensity. The sensitivity value was more in the blue (centred at 501 nm) and red bands (centred at 674 nm), but less near the green bands (centred at 549 nm). When the wavelength reached 674 nm the sensitivity value was positive, suggesting that healthy rice has a higher reflectance value than RLF-infested rice leaves. It was also observed that the sensitivity curve dropped after the red edge zone, generating less payload information while ignoring more sensitive regions.

Continuum removal method: Hyperspectral remote sensing has the ability to access vast amounts of data or thousands of bands. Choosing the most sensitive band, on the other hand, is a difficult task. As a result, data redundancy is the best strategy for detecting sensitive bands. Hence it is

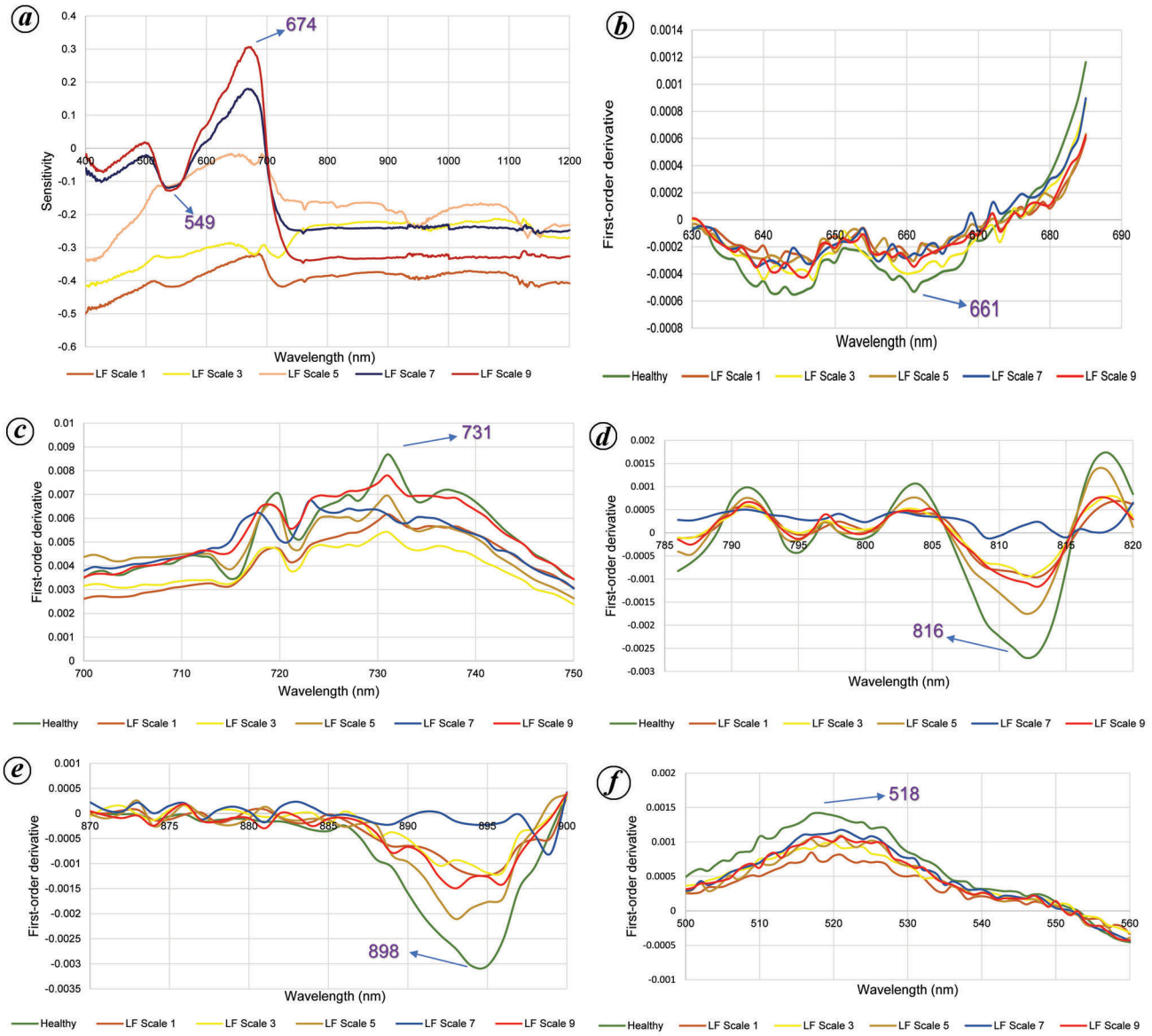

Figure 6. ***a***, Sensitivity analysis of rice leaf folder (RLF) samples for different scales of damage. ***b–f***, Change in derivative from (***b***) 625 to 685 nm, (***c***) 700 to 750 nm, (***d***) 785 to 820 nm, (***e***) 870 to 900 nm and (***f***) 500 to 550 nm.

essential to determine the most feasible bandwidth for discriminating between healthy and infested RLF samples. The most sensitive bands that differentiate between healthy and RLF-infested samples were identified using SDA and continuum reduction. In the VIS and SWIR regions, the amplitude of the shift reduces rapidly with an increase in infestation levels. However, in the NIR region, the reflectance of healthy rice samples is greater than that of RLF-infested samples. ENVI was used to examine the spectral fingerprints of healthy and damaged RLF samples at varying levels of damage. Absorption features of all the damaged levels were examined against a common baseline using the continuum removal method, which normalizes reflectance spectra from 0 to 1. The VIS and NIR regions were analysed to determine the absorption dips. As illustrated in Figure 7, there are three sensitive zones, viz. 558–743, 780–800 and 801–830 nm, and absorption dips occur at 678, 789 and 816 nm (Figure 7 *a–c*).

Identification of sensitive bands

Five sensitive bands were identified based on change in curvature using SDA; three bands were identified on the basis of absorption peaks using the continuum removal method, and two sensitive bands were identified on the basis of sensitivity analysis. The bands at 500–550 and 625–685 nm were sensitive to changes in pigment, while the

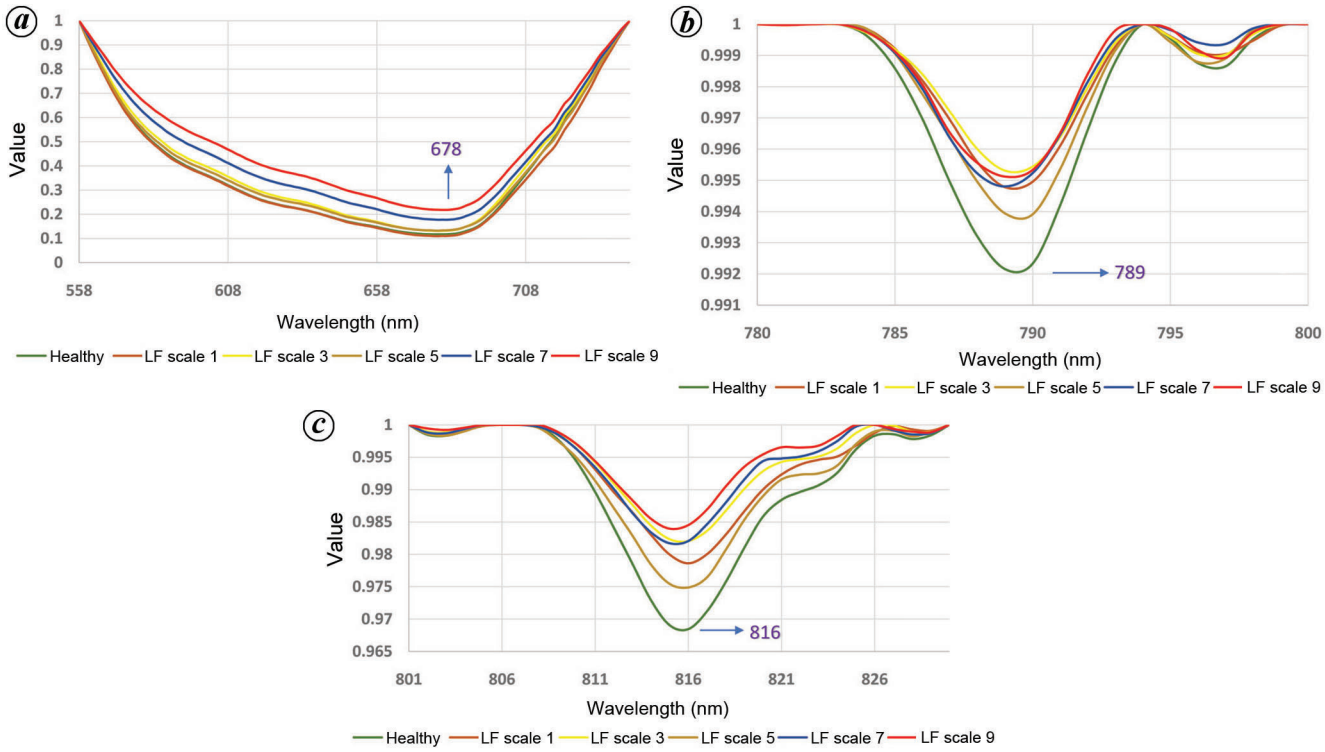

Figure 7. ***a–c***, Continuum removal from (***a***) 558 to 746 nm, (***b***) 780 to 800 nm and (***c***) 801 to 830 nm.

Table 2. Identification of significant bands for leaf folder detection

Method adopted	Sensitive region	Sensitive bands
Spectral derivative analysis (SDA)	625–685	661
	700–750	731
	785–820	816
	870–900	898
	500–550	518
Continuum removal (CR)	558–743	678
	780–800	789
	801–830	816
Sensitivity analysis	400–740	549, 674

Table 3. Combination of bands identified from SDA

Bands	Accuracy (%)	Combination of sensitive bands			
4	78.628	518	661	731	898
3	80.241	518	661	731	
2	73.564	518	661		
1	67.242	518			

longer red portion bands (650–750 nm) were responsible for changes in the leaf area index (LAI) of the healthy and RLF-infested samples[39]. Thus we can infer that the bands at 518 and 549 nm are responsible for carotenoid content, those at 661, 674 and 678 nm for anthocyanins content and the band at 731 nm is responsible for LAI content (Table 2)[39,40]. The regions 785–820 nm (SDA), 780–800 nm and 801–830 nm (continuum removal) showed sensitive bands (789, 816 and 898 nm), but they were neglected due to the presence of atmospheric water perturbations. Also, in satellite images, the NIR region starts from higher wavelengths.

Use of feature selection algorithm to predict the best possible bands

Feature selection procedures are influenced by both the classifier being used and the characteristics of the input data. These approaches call for the definition of a criterion that may be used to assess each feature's quality in terms of its ability to discriminate among the sensitive bands. Then, based on some pre-established criteria, a computational technique is needed to search through the variety of possible subsets of bands and choose the 'best' subset of bands according to their accuracy. This study identifies sensitive bands from the sensitive regions, and then the best bands are selected using the backward sequential feature selection algorithm. Among all the machine learning algorithms, RELIEFF gives the best accuracy results when a combination of sensitive bands is used. It was found that the combination of bands using SDA at 518, 661 and 731 nm gave maximum accuracy of 80.24% (Table 3). The bands 789 and 816 nm were neglected due to water bands which yielded with a single absorption dip (678 nm) identified using CR and its accuracy was 70.36% (Table 4). The combination was extended to the bands identified using SDA

and CR, which yielded that the combination of bands (518, 661 and 678 nm) gave us maximum accuracy of about 79.52% (Table 5). The RELIEFF algorithm was further extended to the combination of sensitive bands using SDA, CR and sensitivity analysis. It was found that the combination of bands at 518, 661 and 731 nm gave the maximum accuracy of 81.67% (Table 6). The RELIEFF algorithm was further extended to a combination of sensitive bands using SDA, continuum removal and sensitivity analysis. It was found that the combination of bands at 518, 661 and 731 nm gave the maximum accuracy of 81.667%.

From Tables 3–6, it is clear that the combination of bands at 518, 661 and 731 nm yields maximum accuracy. Hence, these three bands could be used as the best predictors of RLF.

Discussion

Hyperspectral remote sensing provides crucial data for identifying rice crops affected by RLF. The higher reflection value of light energy for the healthy sample compared to the leaf folder samples in the blue and red regions of the visible spectrum suggests that the plant sample undergoes structural and photosynthetic pigment changes as a result of leaf folder infestation (Figure 4)[41]. The trend in spectral variations in leaf reflectance identified in this study due to leaf folder in rice was similar to those observed due to striped stem borer, leaf folder and a brown spot at different damage levels[4,21,42–44]. To ascertain pest-sensitive bands of RLF from the sensitive spectral region, the spectral reflectance and its derivatives were estimated using Pearson's correlation coefficient method (Figure 5 *a–c*)[4,32].

Interestingly, the reflectance value could not yield any sensitive region as the correlation was around 50–60%. However, when the correlation coefficient was calculated with first- and second-order derivatives and leaf damage per cent, the value increased to about 90%, eventually enabling us to predict the sensitive regions. The spectra generated from the derivatives were less impacted and hence provide higher correlation coefficient compared to values generated from reflectance. This enabled us to identify sensitive regions only on the basis of derivatives and neglect the correlation value obtained from reflectance[45]. The sensitive regions were found from the common regions obtained from both first- and second-order derivatives with distinctive changes in the spectral curve. Finally, five sensitive regions for RLF, viz. 508–529, 671–680, 721–759, 779–786 and 804–820 nm were obtained using SDA. The sensitive regions identified were predominantly confined to the NIR region as the cell structure of the plant was mostly affected due to RLF damage. The red edge region (680–750 nm) showed significant change both in the first- and second-order derivatives, which proves the earlier findings that RLF generally reduces the chlorophyll content of the leaf since chlorophyll detection mostly occurs within this region[46]. This pattern of analysis was almost identical to RLF at the booting stage, where eight sensitive regions were identified, seven at the rice leaf level and one at the rice canopy level[4]. A similar study was carried out for RLF based on correlation coefficients to identify distinct bands sensitive to RLF infestation. Three bands, centred at 424 nm ($r = -0.802$), 758 nm ($r = -0.916$) and 1141 nm ($r = -0.895$) were specially chosen as the most sensitive bands[32]. Pearson's correlation intensity curve for cotton identified four bands at 691, 508, 551 and 710 nm, which were found sensitive for thrips[47]. A similar finding was reported for brown planthopper (BPH), and the study confirmed four sensitive bands (764, 961, 1201 and 1664 nm)[48,49]. The sensitivity analysis also gave significant peaks and dips at 674 and 549 nm in the VIS and NIR regions (Figure 6 *a*), which is inconsistent with the results found for rice brown spot[21,50].

Despite the fact that derivative analysis gives the absorption peaks and dips based on the change in slope and curvature, the continuum removal approach is considered superior for assessing sensitive bands based on band depth[35]. In the present study, the absorption dips (band depth) are found at (678, 789 and 816) nm while the bands 789 and 816 nm were neglected due to water bands and finally 678 nm was considered as the sensitive band[51]. Earlier researchers identified 498 and 678 nm as the sensitive bands for brown spot using continuum removal method[21]. Similarly, adopting continuum removal method for BPH 485, 675 and 750 nm were identified as the sensitive bands[48]. The sensitive bands obtained study were at 518, 549, 661, 674, 678 and 731 nm.

RLF infestation can be recognized by how it influences the physiological characteristics of rice plants using their spectral reflectance[4]. It causes foliar damage to badly affect the growth and nutrient apportion, causing internal damage in chlorophyll pigments and tissue structure[32]. Hence, RLF damage can be clearly identified within the NIR region of the spectrum[52]. In the present study, the pest-sensitive bands for RLF were identified and then they were subjected to an accuracy-based model using RELIEFF algorithm. This can avoid data redundancy if the bands identified are in close proximity to each other, just like 674 and 678 nm in

Table 4. Combination of bands identified from CR

Bands	Accuracy (%)	Combination of sensitive bands
1	70.358	678

Table 5. Combination of bands identified from SDA and CR

Bands	Accuracy (%)	Combination of sensitive bands				
5	70.136	518	661	731	898	678
4	76.28	518	661	678	731	
3	79.517	518	661	678		
2	75.346	518	661			
1	67.18	518				

Table 6. Combination of bands identified from SDA, CR and sensitivity analysis

Bands	Accuracy (%)	Combination of sensitive bands								
9	61.328	661	674	678	731	816	549	518	789	898
8	68.419	661	674	678	731	816	549	518	789	
7	74.374	661	674	678	731	816	549	518		
6	78.527	661	674	678	731	816	549			
5	79.428	661	674	678	731	549				
4	79.271	661	674	678	731					
3	81.667	661	731	518						
2	76.419	661	674							
1	55.38	661								

the present study. The bands 518 (green region), 661 (red region) and 731 nm (NIR region) identified from different regions of the spectra showed with maximum accuracy that can predict the presence of RLF which was inconsistent with other accuracy-based hyperspectral models[53].

Summary

The technique of determining spectral sensitive regions using leaf damage per cent and Pearson's correlation coefficient yields a strong base similar to the outcomes mentioned in this study. The results suggest that hyperspectral data acquired remotely may be used to evaluate leaf folder damage in rice grown in field conditions at the booting stage. The VIS and NIR regions constitute majority of the sensitive bands at 661, 549, 674, 731, 898, 518, 678, 789 and 816 nm. Continuum removal method helped find the sensitive bands as it gave the absorption dips in a specified bandwidth. Feature selection approach gave the most accurate combination of bands (518, 661 and 731 nm) which could easily predict the presence of RLF. Future research using satellite data obtained from aircraft or satellite platforms is necessary to extend this study to the full field or landscape level. Pointing out hotspots on the satellite image with specific coordinates could help find samples affected by leaf folder infestation. Hence, identifying hotspots in satellite imagery can result in enhanced production and the use of the right amount of pesticides. This will be beneficial to scientists, researchers and the farming community.

1. Mohapatra, S. D. *et al.*, Current status and future prospects in biotic stress management in rice. *Oryza*, 2021, **58**, 168–193.
2. Martinelli, F. *et al.*, Advanced methods of plant disease detection. A review. *Agron. Sustain. Dev.*, 2015, **35**, 1–25.
3. Mahlein, A. K., Rumpf, T., Welke, P., Dehne, H. W., Plumer, L., Steiner, U. and Oerke, E. C., Development of spectral indices for detecting and identifying plant diseases. *Remote Sensing Environ.*, 2013, **128**, 21–30.
4. Huang, J. R., Liao, H. J., Zhu, Y. B., Sun, J. Y., Sun, Q. H. and Liu, X. D., Hyperspectral detection of rice damaged by rice leaf folder (*Cnaphalocrocis medinalis*). *Comput. Electron. Agric.*, 2012, **82**, 100–107.
5. Calderon, R., Navas-Cortes, J. A., Lucena, C. and Zarco-Tejada, P. J., High-resolution airborne hyperspectral and thermal imagery for early detection of *Verticillium* wilt of olive using fluorescence, temperature and narrow-band spectral indices. *Remote Sensing Environ.*, 2013, **139**, 231–245.
6. Ghyar, B. S. and Birajdar, G. K., Computer vision-based approach to detect rice leaf diseases using texture and color descriptors. In Proceedings of the International Conference on Inventive Computing and Informatics, Coimbatore, 23–24 November 2017.
7. Nigam, R. *et al.*, Ground-based hyperspectral remote sensing to discriminate biotic stress in cotton crop. In *Multispectral, Hyperspectral and Ultraspectral Remote Sensing Technology, Techniques and Applications VI*, *SPIE*, 2016, vol. 9880, pp. 89–98.
8. Huang, W. J. *et al.*, New optimized spectral indices for identifying and monitoring winter wheat diseases. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sensing*, 2014, **7**, 2516–2524.
9. Mohapatra, S. D. *et al.*, Eco-smart pest management in rice farming: prospects and challenges. *Oryza*, 2019, **56**, 143–155.
10. Dash, L., Ramalakshmi, V., Padhy, D. and Tripathy, B., Breeding for resistance against leaf folder in rice. *Indian J. Pure Appl. Biosci.*, 2020, **8**(6), 248–253.
11. Adhikari, B., Mohapatra, L. N., Senapati, R., Mohapatra, M., Muduli, L. and Mohapatra, S. D., Biochemical changes in rice leaves due to rice leaf folder *Cnaphalocrocis medinalis* (Guenee) infestation. *Pharma Innov.*, 2022, **SP-11**(8), 1463–1468.
12. Tanwar, R. K., Singh, S., Singh, S. P., Kanwar, V. K., Kumar, R., Khokar, M. K. and Mohapatra, S. D., Implementing the systems approach in rice pest management: India context. *Oryza*, 2019, **56**, 136–142.
13. Litsinger, J. A., Bandong, J. P., Canapi, B. L., dela Cruz, C. G., Pantua, P. C., Alviola., A. L. and Batay-An, E. H., Evaluation of action thresholds for chronic rice insect pests in the Philippines. *Int. J. Pest Manage.*, 2006, **52**, 181–194.
14. Jin, J. and Wang, Q., Evaluation of informative bands used in different PLS regressions for estimating leaf biochemical contents from hyperspectral reflectance. *Remote Sensing*, 2019, **11**(2), 197.
15. Xu, B., Li, X., Hou, W., Wang, Y. and Wei, Y., A similarity-based ranking method for hyperspectral band selection. *IEEE Trans. Geosci. Remote Sensing*, 2021, **59**(11), 9585–9599.
16. Saranya, G. and Pravin, A., Feature selection techniques for disease diagnosis system: a survey. In *Artificial Intelligence Techniques for Advanced Computing Applications*, Springer, Singapore, 2021, pp. 249–258.
17. Ashrith, K. N., Sreenivas, A. G., Guruprasad, G. S., Patil, N. B., Hanchinal, S. G. and Chavan, I., Influence of weather parameters on the occurrence of major insect pests in conventional rice ecosystem. *Oryza*, 2017, **54**(3), 324–329.
18. IRRI, *Standard Evaluation System (SES) for Rice, 5th edn*, International Rice Research Institute, Manila, 2013.
19. Chintalapati, P., Javvaji, S. and Gururaj, K., Measurement of damaged leaf area caused by leaffolder in rice. *J. Entomol. Zool Stud.*, 2017, **5**, 415–417.
20. Salisbury, J. W., Spectral measurements field guide. Defense Intelligence Agency Central Measurement and Signature Intelligence

(MASINT) Office, Report No. ADA362372, Fort Belvoir, 1998, pp. 1–82.
21. Zhao, J., Dongyan, Z., Juhua, L., Yingying, D., Hao, Y. and Wenjiang, H., Characterization of the rice canopy infested with brown spot disease using field hyperspectral data. *Wuhan Univ. J. Nat. Sci.*, 2012, **17**, 086–092.
22. Curran, P. J., Dungan, J. L., Macler, B. A., Plummer, S. E. and Peterson, D. L., Reflectance spectroscopy of fresh whole leaves for the estimation of chemical concentration. *Remote Sensing Environ*., 1992, **39**, 153–166.
23. Elvidge, C. D. and Chen, Z., Comparison of broad-band and narrowband red and near-infrared vegetation indices. *Remote Sensing Environ*., 1995, **54**, 38–48.
24. Bao, J., Chi, M. and Benediktsson, J. A., Spectral derivative features for classification of hyperspectral remote sensing images: experimental evaluation. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sensing*, 2013, **6**(2), 594–601.
25. Thorp, K. R., Wang, G., Bronson, K. F., Badaruddin, M. and Mon J., Hyperspectral data mining to identify relevant canopy spectral features for estimating durum wheat growth, nitrogen status, and grain yield. *Comput. Electron. Agric*., 2017, **136**, 1–12.
26. Tsai, F. and Philport, W., Derivative analysis of hyperspectral data. *Remote Sensing Environ*., 1998, **66**, 41–51.
27. Clevers, J. G. P. W., Kooistra, L. and Salas, E. A. L., Study of heavy metal contamination in river floodplains using the red-edge position in spectroscopic data. *Int. J. Remote Sensing*, 2004, **25**, 3883–3895.
28. Smith, K. L., Steven, M. D. and Colls, J. J., Use of hyperspectral derivative ratios in the red edge region to identify plant stress responses to gas leak. *Remote Sensing Environ*., 2004, **92**, 207–217.
29. le Maire, G., Francois, C. and Dufrene, E., Towards universal broad leaf chlorophyll indices using PROSPECT simulated database and hyperspectral reflectance measurements. *Remote Sensing Environ.*, 2004, **89**(1), 1–28.
30. Mageshwaran, M., Srinivasan, M. R. and Sivasamy, R., Detection and estimation of damage caused by rice yellow stem borer *Scirpophaga incertulas* (Walker) by hyperspectral radiometry. In *Proceedings of the National Symposium on Emerging Trends in Eco-friendly Insect Pest Management* (eds Srinivasan, M. R. *et al.*), 2014, pp. 438–440.
31. Tan, Y., Sun, J., Zhang, B., Chen, M., Liu, Y. and Liu, X., Sensitivity of a ratio vegetation index derived from hyperspectral remote sensing to the brown planthopper stress on rice plants. *Sensors*, 2019, **19**, 375.
32. Zhao, J. L., Zhao, C. J., Yang, H., Zhang, D. Y., Dong, Y. Y. and Yuan, L., Identification and characterization of spectral response properties of rice canopy infested by leaf folder. *Int. J. Agric. Biol.*, 2013, **15**, 694–700.
33. Carter, G. A., Seal, M. R. and Haley, T., Airborne detection of southern pine beetle damage using key spectral bands. *Can. J. For. Res.*, 1998, **28**, 1040–1045.
34. Kokaly, R. F. and Clark, R. N., Spectroscopic determination of leaf biochemistry using band-depth analysis of absorption features and stepwise linear regression. *Remote Sensing Environ.*, 1998, **67**, 267–287.
35. Huang, J. R., Sun, J. Y., Liao, H. J. and Liu, X. D., Detection of brown planthopper infestation based on SPAD and spectral data from rice under different rates of nitrogen fertilizer. *Precis. Agric.*, 2015, **16**, 148–163.
36. Ren, J., Wang, R., Liu, G., Feng, R., Wang, Y. and Wu, W., Partitioned relief-F method for dimensionality reduction of hyperspectral images. *Remote Sensing*, 2020, **12**(7), 1104.
37. Pal, M. and Foody, G. M., Feature selection for classification of hyperspectral data by SVM. *IEEE Trans. Geosci. Remote Sensing*, 2010, **48**(5), 2297–2307.
38. Rezaei, Y., Mobasheri, M. R. and Zoej, M. V., Unsupervised information extraction using absorption line in hyperion images. *Int. Arch. Photogramm, Remote Sensing Spat. Inf. Sci.*, 2008, **37**, 383–388.
39. Din, M., Zheng, W., Rashid, M., Wang, S. and Shi, Z., Evaluating hyperspectral vegetation indices for leaf area index estimation of *Oryza sativa* L. at diverse phenological stages. *Front. Plant Sci.*, 2017, **8**, 820.
40. Gu, X., Cai, W., Fan, Y., Ma, Y., Zhao, X. and Zhang, C., Estimating foliar anthocyanin content of purple corn via hyperspectral model. *Food Sci. Nutr.*, 2018, **6**(3), 572–578.
41. Liang, G. C., Ouyang, Y. C. and Dai, S. M., Detection and classification of rice infestation with rice leaf folder (*Cnaphalocrocis medinalis*) using hyperspectral imaging techniques. *Remote Sensing*, 2021, **13**(22), 4587.
42. Fan, Y., Wang, T., Qiu, Z., Peng, J., Zhang, C. and He, Y., Fast detection of striped stem borer (*Chilo suppressalis* Walker) infested rice seedling based on visible/near-infrared hyperspectral imaging system. *Sensors*, 2017, **17**, 2470.
43. Liu, T., Shi, T., Zhang, H. and Wu, C., Detection of rise damage by leaf folder (*Cnaphalocrocis medinalis*) using unmanned aerial vehicle based hyperspectral data. *Sustainability*, *MDPI, Open Access Journal*, 2020, **12**(22), 1–14.
44. Adhikari, B. *et al.*, Discrimination of healthy and damaged rice by leaf folder using hyperspectral Sensing. In *First Indian Rice Congress*, ICAR-National Rice Research Institute, Cuttack, 2020, pp. 681–684.
45. Han, L., Estimating chlorophyll-*a* concentration using first-derivative spectra in coastal water. *Int. J. Remote Sensing*, 2005, **26**(23), 5235–5244.
46. Thenkabail, P. S., Enclona, E. A., Asthon, M. S. and Van der, M. B., Accuracy assessment of hyperspectral wave band performance for vegetation analysis applications. *Remote Sensing Environ.*, 2004, **91**, 354–376.
47. Ranjitha, G., Srinivasan, M. R. and Rajesh, A., Detection and estimation of damage caused by thrips *Thrips tabaci* (Lind) of cotton using hyperspectral radiometer. *Agrotechnology*, 2014, **3**(1), 123–128.
48. Prasannakumar, N. R., Chander, S. and Sahoo, R. N., Characterization of brown planthopper damage on rice crops through hyperspectral remote sensing under field conditions. *Phytoparasitica*, 2014, **42**, 387–395.
49. Xue, L. Z., Xu, L., Tan, Y. and Liu, X. D., Spectral characteristics of different rice cultivars damaged by the brown planthopper *Nilaparvata lugens*. *J. Nanjing Agric. Univ.*, 2015, **38**, 796–803.
50. Liu, Z. Y., Huang, J. F. and Tao, R. X., Characterizing and estimating fungal disease severity of rice brown spot with hyperspectral reflectance data. *Rice Sci.*, 2008, **15**(3), 232–242.
51. Naik, B. B., Naveen, H. R., Sreenivas, G., Choudary, K. K., Devkumar, D. and Adinarayana, J., Identification of water and nitrogen stress indicative spectral bands using hyperspectral remote sensing in maize during post-monsoon season. *J. Indian Soc. Remote Sensing*, 2020, **48**(12), 787–795.
52. Singh, B., Singh, M., Suri, K., Sharma, A. and Mishra, P. K., Use of hyper-spectral data for detection of rice leaf folder infestation. *J. Res. Punjab Agric. Univ.*, 2013, **50**(3 and 4), 147–150.
53. Wu, B., Chen, C., Kechadi, T. M. and Sun, L., A comparative evaluation of filter-based feature selection methods for hyper-spectral band selection. *Int. J. Remote Sensing*, 2013, **34**(22), 7974–7990.

ACKNOWLEDGEMENTS. We thank the Director, ICAR-National Rice Research Institute, Cuttack and the Head, AED (Agriculture and Land Ecosystem Division)/BPSG (Biological and Planetary Sciences Group)/EPSA (Earth, Ocean, Atmosphere Planetary Sciences and Applications Area), Space Application Centre (SAC), Ahmedabad for support. This study was funded by SAC-ISRO, Ahmedabad under the Space Technology Utilization for Food Security, Agricultural Assessment and Monitoring (SUFALAM) project.

Received 3 September 2022; revised accepted 4 January 2023

doi: 10.18520/cs/v124/i8/964-975