



# Spectral analysis of winter wheat leaves for detection and differentiation of diseases and insects



Lin Yuan<sup>a,b</sup>, Yanbo Huang<sup>c</sup>, Rebecca W. Loraamm<sup>d</sup>, Chenwei Nie<sup>a,b</sup>,  
Jihua Wang<sup>a,b</sup>, Jingcheng Zhang<sup>a,b,\*</sup>

<sup>a</sup> Beijing Research Center for Information Technology in Agriculture, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China

<sup>b</sup> Institute of Remote Sensing and Information Application, Zhejiang University, Hangzhou 310058, China

<sup>c</sup> USDA-ARS, CPSRU, 141 Experiment Station Road, Stoneville, MS 38776, USA

<sup>d</sup> Department of Geography, Environment and Planning, University of South Florida, 4202 E. Fowler Ave, Tampa, FL 33620, USA

## ARTICLE INFO

### Article history:

Received 25 June 2013

Received in revised form

12 November 2013

Accepted 13 November 2013

### Keywords:

Hyperspectral

Powdery mildew

Yellow rust

Aphid

Fisher linear discrimination analysis (FLDA)

Partial least square regression (PLSR)

## ABSTRACT

Yellow rust (*Puccinia striiformis* f. sp. *Tritici*), powdery mildew (*Blumeria graminis*) and wheat aphid (*Sitobion avenae* F.) infestation are three serious conditions that have a severe impact on yield and grain quality of winter wheat worldwide. Discrimination among these three stressors is of practical importance, given that specific procedures (i.e. adoption of fungicide and insecticide) are needed to treat different diseases and insects. This study examines the potential of hyperspectral sensor systems in discriminating these three stressors at leaf level. Reflectance spectra of leaves infected with yellow rust, powdery mildew and aphids were measured at the early grain filling stage. Normalization was performed prior to spectral analysis on all three groups of samples for removing differences in the spectral baseline among different cultivars. To obtain appropriate bands and spectral features (SFs) for stressor discrimination and damage intensity estimation, a correlation analysis and an independent *t*-test were used jointly. Based on the most efficient bands/SFs, models for discriminating stressors and estimating stressor intensity were established by Fisher's linear discriminant analysis (FLDA) and partial least square regression (PLSR), respectively. The results showed that the performance of the discrimination model was satisfactory in general, with an overall accuracy of 0.75. However, the discrimination model produced varied classification accuracies among different types of diseases and insects. The regression model produced reasonable estimates of stress intensity, with an  $R^2$  of 0.73 and a RMSE of 0.148. This study illustrates the potential use of hyperspectral information in discriminating yellow rust, powdery mildew and wheat aphid infestation in winter wheat. In practice, it is important to extend the discriminative analysis from leaf level to canopy level.

© 2013 Elsevier B.V. All rights reserved.

## 1. Introduction

Diseases and insects cause reduced crop yields every year, posing a significant risk to food production worldwide (Strange and Scott, 2005). In dealing with crop diseases and insects, application of an appropriate amount of fungicide/insecticide is important. Misuse and overuse of fungicide/insecticide could result in failure of disease/insect control, and even soil contamination (Sankaran et al., 2010). With the capability of observing the ground surface in a spatially continuous manner, remote sensing provides an

alternative to conventional means for disease and insect surveys. As plant stress may be characterized using specific responses in the visible (VIS), near infrared (NIR) and shortwave infrared (SWIR) spectral domains, it is possible to detect or map plant response to diseases/insects with remotely sensed data (Sankaran et al., 2010; Escalante-Ramírez, 2012).

To conduct remote sensing of crop diseases/insects, it is important to identify bands or spectral features that are sensitive to specific stressors. Delwiche and Kim (2000) assessed the spectral characteristics of fusarium head blight disease in winter wheat, and found that the reflectance at 550, 568, 605, 623, 660, 697, 715 and 733 nm are the best indicators of disease presence. Moshou et al. (2004) reported that the reflectance at 680, 725 and 750 nm could be used to detect yellow rust in winter wheat. Yang et al. (2005) identified that the band centered at 694 nm can be used to detect greenbug infestation. Graeff et al. (2006) successfully distinguished powdery mildew and take-all disease infected wheat leaves from normal wheat leaves with reflectance at 490, 510, 516, 540, 780, and

**Abbreviations:** VIs, vegetation indexes; SFs, spectral features; FLDA, Fisher's linear discriminant analysis; PLSR, partial least square regression; VIS, visible; NIR, near infrared; SWIR, shortwave infrared; PM, powdery mildew; YR, yellow rust; AH, aphid; DI, damage index.

\* Corresponding author. +Tel.: +86 10 51503215.

E-mail address: [zhangjc@nercita.org.cn](mailto:zhangjc@nercita.org.cn) (J. Zhang).

1300 nm at an early stage. Apart from the original bands, vegetation indices were also shown to be useful in diagnosis of plant diseases and insects, given that most VIs can enhance spectral features at specific positions through certain transformations (e.g. subtraction, dividing, and normalization). It was found that yellow rust could be detected using NDVI (Normalized Difference Vegetation Index) or PRI (Photochemical Reflectance Index) (Bravo et al., 2003; Huang et al., 2007). In detection of aphid infestation, Mirik et al. (2007) proposed an index called AI (Aphid Index) to indicate the abundance of Russian wheat aphids in the field.

As studied above, most studies so far focus on spectral response or detection of one specific disease or insect; they seldom consider spectral mixing between different diseases or insects. However, in practice, different diseases and insects often occur simultaneously in the field, posing a challenge to their detection and discrimination. For example, serious diseases and insects (including powdery mildew yellow rust and aphids) of winter wheat in northern China tend to occur simultaneously during the jointing to grouting stages because their geographical regions overlap. To avoid possible spectral mixing among these diseases and insects, prior to estimation of their severity, it is necessary to compare spectral responses and identify spectral features that are able to differentiate these conditions. However, in reviewing the literature it was found that this comparison of spectral signatures was generally missing. To address this issue, hyperspectral measurements were made for all three types of diseases and insects together.

The objectives of this paper are: (1) to understand the difference of spectral responses among PM (powdery mildew), YR (yellow rust) and AH (aphid) at leaf level on winter wheat; (2) to identify efficient spectral features for differentiating the three stressors; and (3) to develop a discriminant model for differentiating PM, YR and AH, and a regression model for estimating their severities.

## 2. Materials and methods

### 2.1. Data acquisition

#### 2.1.1. Study area and disease inoculation

The experiment was conducted at Beijing Xiaotangshan Precision Agriculture Experimental Base, China (40° 10.6'N, 116° 26.3'E) during the wheat growing seasons in 2011 and 2012. Three different wheat cultivars, Jingshuang16, Jing9843, and Zhongmai16, were used in the experiment, and they were susceptible to PM, YR and AH, respectively.

To introduce YR and PM pathogens to wheat plants, an inoculation technique was used. Given that aphids tend to occur naturally in the experimental field every year, no pesticide treatment was used to allow aphid infestation. For each disease/insect treatment, three replicates were set with a unit plot in size of 5 m × 5 m. In addition, three control plots were also used as a reference containing exactly the same cultivars but with the use of pesticides to prevent infections. To avoid cross infection among different types of diseases/insects, the distance between each treatment plot was no less than 20 m. In both years, the PM and YR showed visible leaf symptoms beginning in the early grain filling stage. The AH also exhibited clear symptoms at this stage. Symptoms of the three stressors are shown in Fig. 1. For this reason, the early grain filling stage was chosen for data collection in this study on May 23 in 2011 and May 16 in 2012.

#### 2.1.2. Leaf sampling and spectral measurement

Leaf spectral measurements were taken using a FieldSpec® UV/VNIR spectroradiometer (ASD Inc., Boulder, Colorado, USA) over the 350–2500 nm wavelengths, coupled with an ASD Leaf Clip, an accessory of the ASD spectroradiometer. The device included a

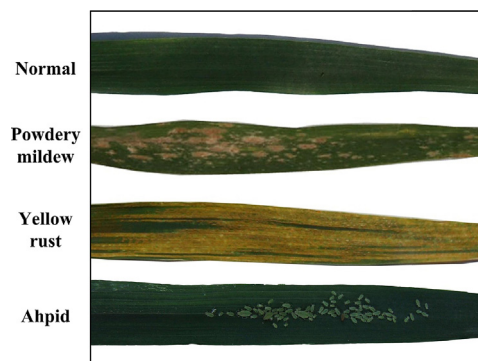


Fig. 1. Normal leaves and leaves infected with powdery mildew, yellow rust and aphid of winter wheat.

built-in light source, which allowed for indoor measurements. The spectrum of a white Spectralon reference panel (99% reflectance) was measured once for every 10 leaf measurements. To convert the radiance signals to spectral reflectance, a calibrated white panel was used. The dark reference panel was placed behind the leaf when taking measurement given it can absorb transmitted lights. This avoids possible disruption from multiple transmitted lights. Leaf reflectance was generated by dividing the sample radiance with the radiance of the white Spectralon with a multiplication of the reflectance of white reference:

$$\text{Ref}_{(\text{target})} = \frac{\text{Rad}_{(\text{target})}}{\text{Rad}_{\text{white reference}}} \times \text{Ref}_{\text{white reference}} \times 100\% \quad (1)$$

where  $\text{Ref}_{(\text{target})}$  is the reflectance of the observing target;  $\text{Rad}_{(\text{target})}$  is the radiance of the observing target;  $\text{Rad}_{(\text{white reference})}$  is the radiance of the white panel; and  $\text{Ref}_{(\text{white reference})}$  is the reflectance of the white panel.

For each leaf, 15 readings were recorded and then averaged to obtain a spectral curve for the leaf. A digital color photo was also taken immediately after each spectral measurement with a white paper background to facilitate later determination of disease severity. For YR, PM and AH treatments, a total of 66, 33 and 50 spectral measurements were made for infected leaf samples, whereas 26, 14 and 16 measurements were made for non-infected leaf samples, respectively. For calibration and validation of the models in each group, samples were randomly split into sub-groups representing 60% and 40% of the samples in each group.

#### 2.1.3. Quantification of leaf damage levels

In this study, damage index (DI) was used to indicate the damage levels caused by different diseases/insects. For YR, PM and AH, the DI was determined by a visual estimation of the percent coverage of pustules/aphids on the leaf (Graeff et al., 2006; Zhang et al., 2012a; Luo et al., 2012). For this estimation, the DI value was recorded in intervals of 5% ranging from 5% to 100%. Leaves with a DI less than 5% were difficult to visually separate from normal ones; these were classified as normal leaves. All estimations were made according to leaf photos visualized by a trained technician to minimize subjective error.

#### 2.1.4. Spectral features for stressor discrimination and detection

In addition to raw reflectance, various forms of spectral features were also included for discriminating and detecting the PM, YR and AH given their physical and biological significance. As a result, a total of 30 spectral features were chosen, including the first derivative transformed spectral features, continuous removal transformed spectral features and vegetation index. The first derivative transformed SFs and continuous removal transformed SFs were considered as they were efficient indicators of

**Table 1**

Summary of first derivative spectral features, continuous removal transformed spectral features and vegetation index that are used for disease detection of powdery mildew, yellow rust and aphid at leaf level.

Variable	Definition	Description	Literatures
Derivative transformed spectral variables			
$D_b$	Maximum value of 1st derivative within blue edge	Blue edge covers 490–530 nm. $D_b$ is a maximum value of 1st order derivatives within the blue edge of 35 bands	Gong et al. (2002)
$\lambda_b$	Wavelength at $D_b$	$\lambda_b$ is wavelength position at $D_b$	Gong et al. (2002)
$SD_b$	Sum of 1st derivative values within blue edge	Defined by sum of 1st order derivative values of 35 bands within the blue edge	Gong et al. (2002)
$D_y$	Maximum value of 1st derivative within yellow edge	Yellow edge covers 550–582 nm. $D_y$ is a maximum value of 1st order derivatives within the yellow edge of 28 bands	Gong et al. (2002)
$\lambda_y$	Wavelength at $D_y$	$\lambda_y$ is wavelength position at $D_y$	Gong et al. (2002)
$SD_y$	Sum of 1st derivative values within yellow edge	Defined by sum of 1st order derivative values of 28 bands within the yellow edge	Gong et al. (2002)
$D_r$	Maximum value of 1st derivative within red edge	Red edge covers 670–737 nm. $D_r$ is a maximum value of 1st order derivatives within the red edge of 61 bands	Gong et al. (2002)
$\lambda_r$	Wavelength at $D_r$	$\lambda_r$ is wavelength position at $D_r$	Gong et al. (2002)
$SD_r$	Sum of 1st derivative values within red edge	Defined by sum of 1st order derivative values of 61 bands within the red edge	Gong et al. (2002)
Continuous removal transformed spectral features			
DEP550–750	The depth of the feature minimum relative to the hull	In the range of 550 nm–750 nm	Pu et al. (2003, 2004)
DEP920–1120		In the range of 920 nm–1120 nm	
DEP1070–1320		In the range of 1070 nm–1320 nm	
WID550–750	The full wavelength width at half DEP (nm)	In the range of 550 nm–750 nm	Pu et al. (2003, 2004)
WID920–1120		In the range of 920 nm–1120 nm	
WID1070–1320		In the range of 1070 nm–1320 nm	
AREA550–750	The area of the absorption feature that is the product of DEP and WID	In the range of 550 nm–750 nm	Pu et al. (2003, 2004)
AREA920–1120		In the range of 920 nm–1120 nm	
AREA1070–1320		In the range of 1070 nm–1320 nm	
VI-based variables			
GI	Greenness index	$R_{554}/R_{677}$	Zarco-Tejada et al. (2005)
NDVI	Normalized difference vegetation index	$(R_{NIR} - R_R)/(R_{NIR} + R_R)$	Rouse et al. (1973)
TVI	Triangular vegetation index	$0.5 \times [120 \times (R_{750} - R_{550}) - 200 \times (R_{670} - R_{550})]$	Zhao et al. (2004)
PRI	Photochemical reflectance index	$(R_{570} - R_{531})/(R_{570} + R_{531})$	Huang et al. (2007), Naidu et al. (2009)
CARI	Chlorophyll absorption ratio index	$\frac{[(a670 + R_{670} + b)] / (a^2 + 1)^{1/2} \times (R_{700}/R_{670})}{a = (R_{700} - R_{550})/150, b = R_{550} - (a \times 550)}$	Kim et al. (1994)
RVSI	Red-edge vegetation stress index	$[(R_{712} + R_{752})/2] - R_{732}$	Merton (1998)
SIPI	Structural independent pigment index	$(R_{800} - R_{445})/(R_{800} + R_{680})$	Devadas et al. (2009)
NPCI	Normalized pigment chlorophyll ratio index	$(R_{680} - R_{430})/(R_{680} + R_{430})$	Devadas et al. (2009)
WI	Water index	$R_{900}/R_{970}$	Naidu et al. (2009)
NDWI	Normalized difference water index	$(R_{860} - R_{1240})/(R_{860} + R_{1240})$	Gao (1996)
AI	Aphid index	$(R_{740} - R_{887})/(R_{691} - R_{698})$	Mirik et al. (2006)
DSSI2	Damage sensitive spectral index2	$(R_{747} - R_{901} - R_{537} - R_{572})/(R_{747} - R_{901} + R_{537} - R_{572})$	Mirik et al. (2006)

absorption characteristics and curve shapes (Gong et al., 2002; Pu et al., 2003, 2004; Weng, 2011). Further, some hyperspectral VIs were included for their physical mechanism in reflecting plant physiological status. Photochemical reflectance index (PRI) and Aphid Index (AI) have been used for disease/insect detection (Naidu et al., 2009; Mirik et al., 2006). The definitions, formulas, and corresponding literature of all SFs are summarized in Table 1.

## 2.2. Data processing and analysis

In this study, data processing and analysis was carried out in three parts: spectral normalization, identification of efficient spectral features, and the development of models for disease/insect discrimination and severity estimation (Fig. 2).

### 2.2.1. Spectral normalization

As the leaf spectra of yellow rust, powdery mildew and aphid infestation were collected from different cultivars, a spectral normalization was implemented at first to adjust the spectral data from different groups to an identical baseline, which facilitated subsequent spectral analysis and comparisons. A detailed procedure for this normalization is described in Zhang et al. (2012b). In the present study, the non-infected samples of YR group were

treated as a spectral standard. Spectral ratios were calculated using an averaged spectral curve of non-infected samples of the YR group divided by an averaged non-infected spectral curve of the PM or AH group.

The spectral ratio curves were shown to reflect the difference in spectral baseline between the two groups of measurements. Therefore, a spectral adjustment was implemented by multiplying the spectral curve on each spectra of the PM or AH group:

$$Ref'_{(PM/AH)} = Ref_{(PM/AH)_i} \times Spectral\ Ratio_i \quad (2)$$

where  $Ref'_{(PM/AH)_i}$  is the normalized reflectance at  $i$  band of PM or AH group;  $Ref_{(PM/AH)_i}$  is the original reflectance at  $i$  band of PM or AH group;  $Spectral\ Ratio_i$  is the spectral ratio between non-infected reflectance of YR group and PM or AH group at  $i$  band.

The above normalization procedures do not alter the inherent spectral differences among normal and diseased samples from one group given that exactly the same adjustment is implemented for all spectra from one group. This procedure makes the three groups have the same spectral baseline, and thereby facilitates spectral comparison between different stressors. All subsequent comparisons and analysis were based on the normalized spectra.

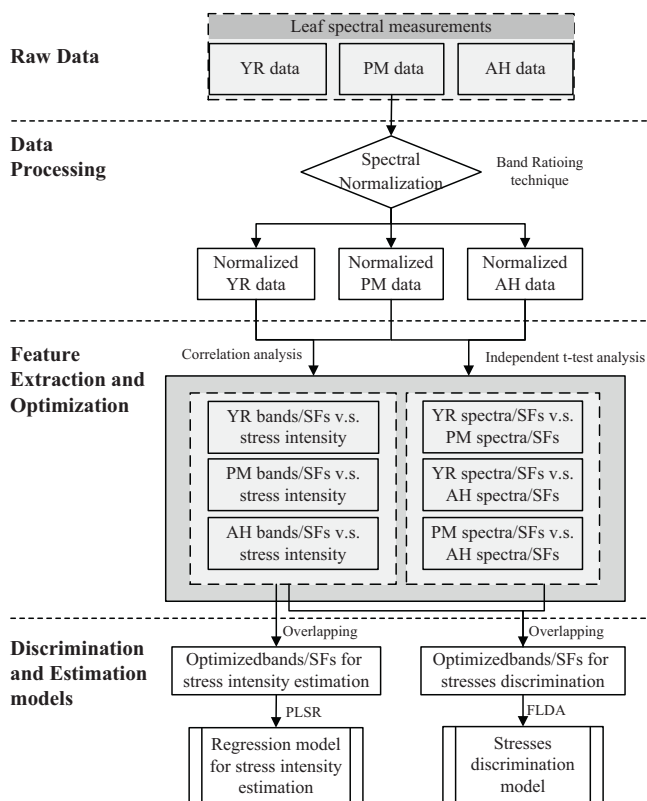


Fig. 2. Flowchart of data analysis and processing.

### 2.2.2. Comparison of spectral changes caused by PM, YR and AH

To demonstrate the spectral changes that are caused by different stressors, a ratio procedure was conducted for all three groups, which divided the averaged spectra of stressed samples by the averaged spectra of normal samples:

$$\text{Ratio}_{(\text{stress})} = \frac{\overline{Ref}_{(\text{stress})}}{\overline{Ref}_{(\text{Healthy})}} \quad (3)$$

where  $\overline{Ref}_{(\text{stress})}$  is the average reflectance of the stressed samples;  $\overline{Ref}_{(\text{Healthy})}$  is the average reflectance of corresponding healthy samples.

To compare the spectral difference between different stressors, an independent *t*-test was also used to assess the significance of spectral differences between each pair of stressors for both the original bands and SFs (only stressed samples were used in this analysis). In addition, to illustrate the spectral response to different stressors, a correlation analysis was used to assess the strength of correlation between original bands/SFs and sample's DI for PM, YR and AH groups, respectively.

### 2.2.3. Discrimination of diseases/insects by spectral features

To differentiate stressors by means of their spectral measurements, the bands and SFs were selected with high sensitivity to DI and significant difference among stressors. The *p*-values of the correlation analysis and independent *t*-test were used to indicate the strength of correlation between stressors and significance of spectral difference between stressors, respectively. Then, an overlapping procedure was conducted to screen for bands and SFs that are sensitive to all three stressors (*p*-value <0.001), while also having significant difference between stressors (*p*-value <0.001). Based on the remaining bands and SFs, a discriminant model for differentiating PM, YR and AH was constructed using Fisher linear discriminate analysis (FLDA) (Mclachlan, 2004). With the 40%

validation samples, five statistics were calculated from the confusion matrix to reflect the accuracy of the discriminant model: overall accuracy (OA), average accuracy (AA), producer's accuracy, user's accuracy, and *kappa* coefficient.

### 2.2.4. Estimation of damage index by spectral features

In addition to the differentiation among two diseases and one insect, the bands/SFs were identified that are most sensitive to all three types of stressors to construct a model for estimating a sample's DI. To achieve this goal, only bands/SFs exhibiting significant correlation (*p*-value <0.001) to all the PM, YR and AH stressors were identified by the overlapping procedure. To further eliminate information redundancy among retained bands/SFs, the correlation between each two variables was checked. For a pair of variables with an  $R^2$  over 0.8, the one with stronger correlation to DI (indicated by  $R^2$  of the correlation analysis) was retained. This procedure was repeated until there were no pairs of variables having an  $R^2$  over 0.8. Based on these remaining bands/SFs, the Partial Least Square Regression (PLSR) analysis was adapted to develop multivariate models for estimating sample DI. An advantage of PLSR over traditional regression methods is its capability to eliminate possible multicollinearity among variables. This feature makes PLSR a favorable regression method in spectral analysis and modeling (Zhang et al., 2012a).

According to the validation sample, the performance of the PLSR model was evaluated by two accuracy indices: the coefficient of determination ( $R^2$ ) and the relative root mean square error (RMSE).

## 3. Results

### 3.1. Comparisons among spectral signatures of PM, YR and AH

After the normalization process, the averaged original spectra of different diseases/insect, as well as normal treatment are demonstrated in Fig. 3a. To demonstrate the spectral changes produced by different stressors, the ratio curves of PM, YR and AH (using Eq. (3)) are illustrated in Fig. 3b. As the ratio curve reflects both change direction (increase or decrease) and change magnitude of reflectance, it can be treated as a spectral signature of a specific stressor. Generally, the two diseases and aphid presence lead to an increase in reflectance over most of the spectrum.

The shapes of spectral curves for PM, YR and AH generally increase in reflectance at 500–690 nm, 1390–1520 nm, and 1860–2080 nm, with a ratio peak over 620–670 nm. However, as a clear difference of change direction, the PM and AH produced a decrease of reflectance at 750–1300 nm, whereas the YR produced an increase of reflectance over this range. In addition, the magnitude of changes varied significantly among different diseases/insect. Generally, the PM resulted in the strongest spectral change over entire spectrum region, followed by YR and AH. The ratio curves of the two diseases and insect exhibited the most significant difference at 350–550 nm, 580–700 nm, 730–1320 nm, 1400–1540 nm, and 1870–2140 nm (Fig. 3b). For example, within 580–700 nm, the PM, YR and AH induced 66%, 42% and 15% spectral increase, respectively.

To quantify the bands' sensitivity to different stressors along the spectrum, correlation curves are displayed in Fig. 3c. It is evident that the spectral response differed significantly among PM, YR and AH. Using a threshold of *p*-value <0.001, sensitive bands and SFs were identified corresponding to the two diseases and insect (Table 2, Fig. 4). Moreover, to assess the significance of difference between stressors, curves of *p*-values were drawn for each pair of diseases/insect (PM vs. YR, PM vs. AH, AH vs. YR) (Fig. 3d). With a threshold of *p*-value <0.001, the bands and SFs were identified for a good potential to differentiate stressors (Table 2, Fig. 4).

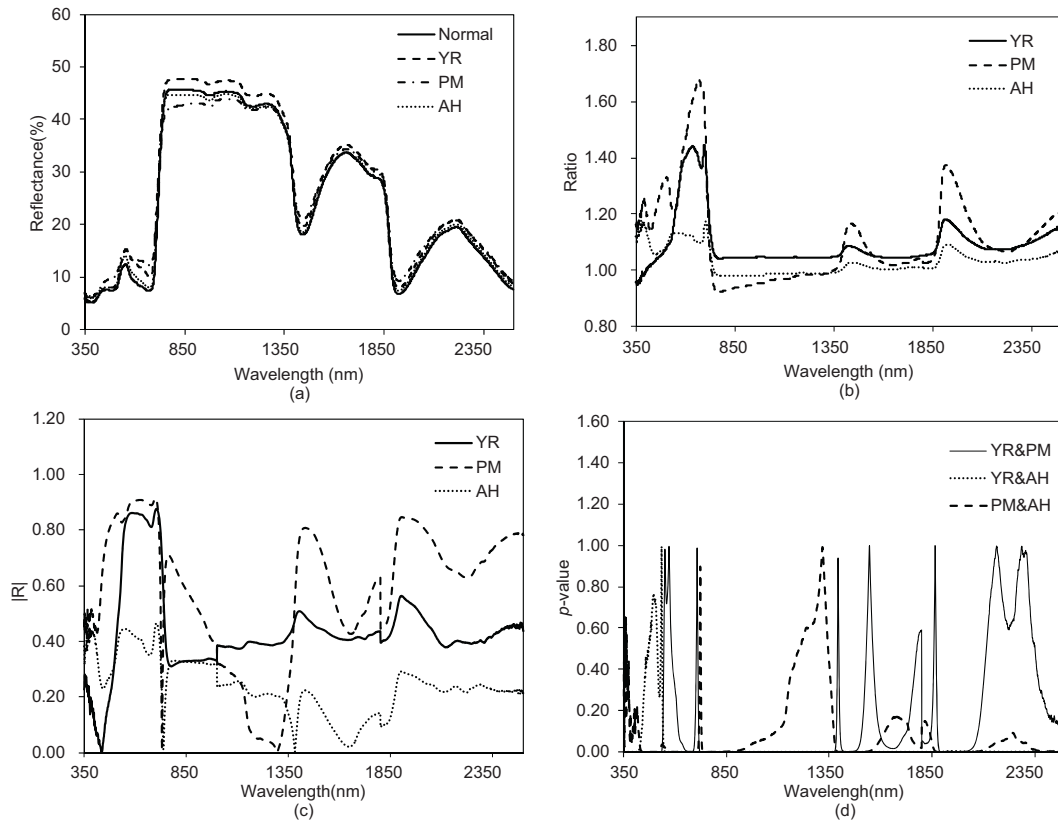


Fig. 3. Comparisons of spectral responses between normal and different diseases/pest.

3.2. Discrimination among PM, YR and AH based on spectral features

As shown in Fig. 4, according to the overlapping process, three original bands at 666–683 nm (Ref<sub>666–683nm</sub>), 752–758 nm (Ref<sub>752–758nm</sub>), and 1893–1905 nm (Ref<sub>1893–1905nm</sub>) were

identified, which are both sensitive to PM, YR and AH, and also exhibited significant difference among the two diseases and the insect. Similarly, for SFs, the overlapping procedure resulted in four SFs, D<sub>y</sub>, GI, NDVI, and PRI. Finally, using an autocorrelation analysis, the four bands and the SFs were retained, which were Ref<sub>752–758nm</sub>, Ref<sub>1893–1905nm</sub>, D<sub>y</sub>, and GI. Based on these features a

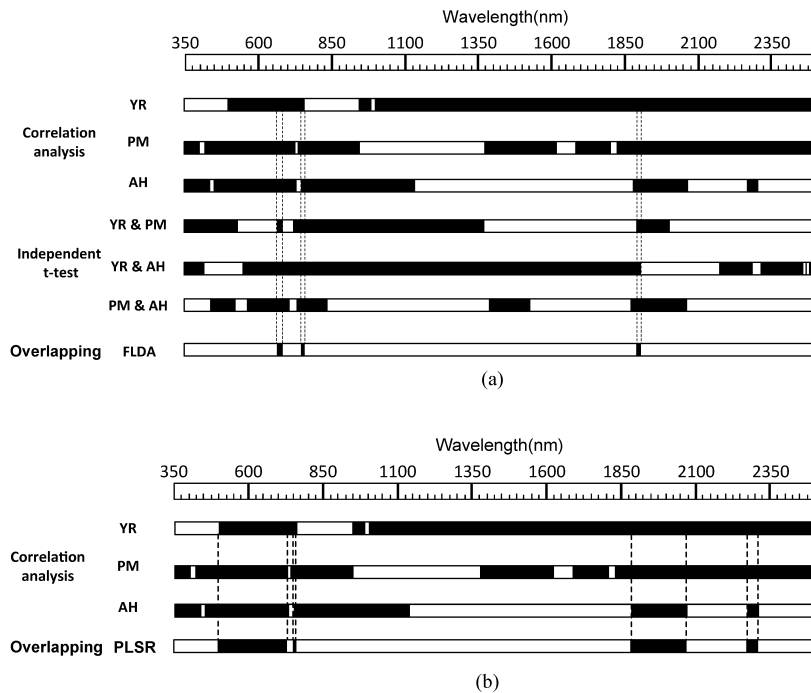


Fig. 4. Overlapping band selection.

**Table 2**  
Summary of disease sensitivity and discriminative capability of diseases and pest for spectral features.

PLSR modle Spectral features	Correlation			Independent <i>t</i> -test		FLDA modle	
	YR	PM	AH	YR&PM	YR&AH	PM&AH	Spectral features
$D_b$	+		+	+		+	$D_b$
$\lambda_b$	+				+	+	$\lambda_b$
$SD_b$	+		+	+	+		$SD_b$
$D_y$	+	+	+	+	+	+	$D_y$
$\lambda_y$							$\lambda_y$
$SD_y$	+	+	+		+	+	$SD_y$
$D_r$	+	+	+	+		+	$D_r$
$\lambda_r$	+	+	+				$\lambda_r$
$SD_r$		+		+		+	$SD_r$
DEP550–770	+	+	+	+		+	DEP550–770
AREA550–770	+	+	+	+		+	AREA550–770
WID550–770	+	+	+		+	+	WID550–770
DEP920–1120			+				DEP920–1120
AREA920–1120		+	+	+			AREA920–1120
WID920–1120			+				WID920–1120
DEP1070–1320		+	+	+	+	+	DEP1070–1320
AREA1070–1320		+		+	+	+	AREA1070–1320
WID1070–1320							WID1070–1320
GI	+	+	+	+	+	+	GI
NDVI	+	+	+	+	+	+	NDVI
TVI				+		+	TVI
PRI	+	+	+	+	+	+	PRI
CARI	+	+	+		+		CARI
RVSI	+	+		+	+	+	RVSI
SIPI	+	+	+	+		+	SIPI
NPCI	+	+	+		+	+	NPCI
WI		+		+		+	WI
NDWI		+	+	+		+	NDWI
AI	+	+	+				AI
DSSI2	+						DSSI2

“+” Indicates the selected SFs (for correlation analysis, the significant  $>0.001$  were chosen; for independent *t*-test, *p*-value  $>0.001$  were chosen); YR, PM and AH represent yellow rust, powdery mildew and aphid, respectively.  $\square$  and  $\square$  represent SFs that were selected for PLSR and FLDA, respectively.

discriminant model was constructed using FLDA (Table 3). Using an independent validation sample, the OA(%) and kappa of the FLDA model reached 0.76 and 0.67, respectively. As for the three stressors, the model produced the highest accuracy in identifying YR, with OA reaching 0.96. For PM and AH, the OA was approaching 0.85 and 0.90, respectively. In contrast, the model failed to produce accurate identification for normal samples, which only yielded an OA of 0.35.

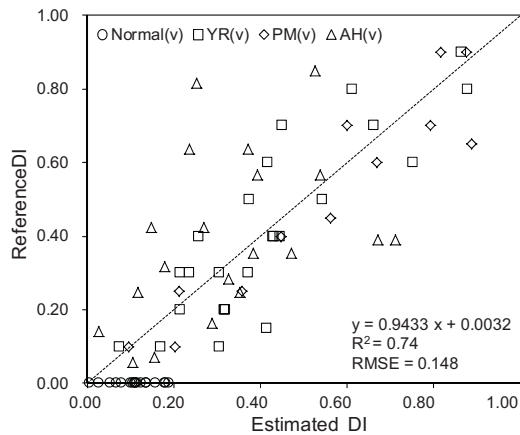
### 3.3. Regression model for DI estimation for YR, PM and AH

As shown in Fig. 4(b), all disease and insect sensitive bands were overlapped, yielding four bands: 500–725 nm, 752–758 nm, 1879–2063 nm and 2268–2303 nm. In addition, 14 SFs were identified that were sensitive to all the PM, YR and AH, which were  $D_y$ ,  $SD_y$ ,  $D_r$ ,  $\lambda_r$ , DEP550–770, AREA550–770, WID550–770, GI, NDVI, PRI, CARI, SIPI, NPCI, and AI (Table 2). Then, an autocorrelation analysis

**Table 3**  
Confusion matrix and classification accuracies for both the calibration and validation of the FLDA model.

	Reference			Sum	U.'s a. (%)	OA (%)	Kappa		
	Normal	YR	PM						AH
Calibration	Normal	15	0	4	0	19	78.95	74.80	0.66
	YR	2	37	0	2	41	90.24		
	PM	4	2	16	4	26	61.54		
	AH	12	1	0	24	37	64.86		
	Sum	33	40	20	30	123			
	P.'s a. (%)	45.45	92.50	80.00	80.00				
Validation	Normal	8	0	1	1	10	80.00	75.61	0.67
	YR	1	25	0	1	27	92.59		
	PM	4	0	11	0	15	73.33		
	AH	10	1	1	18	30	60.00		
	Sum	23	26	13	20	82			
		P.'s a. (%)	34.78	96.15	84.62	90.00			

OA (%) is overall accuracy; P.'s a. represents producer's accuracy; U.'s a. represents user's accuracy.



**Fig. 5.** Scatter plots between the reference DI and the estimated DI based on PLSR model.

was implemented to the above chosen bands and SFs, which lead to eight variables,  $Ref_{752-758nm}$ ,  $Ref_{2268-2303nm}$ ,  $\lambda_r$ ,  $AREA_{550-770}$ , GI, CARI, SIPI, and AI. All bands/SFs retained were used to construct a partial least square regression analysis (PLSR) model for DI estimation. Using an independent validation, the model achieved satisfactory accuracies, with an  $R^2$  of 0.73 and RMSE of 0.148 (Fig. 5).

#### 4. Discussion

As shown in Table 2 and Fig. 3, the leaves that were infected by PM, YR and AH exhibited some consistent spectral responses across the spectrum. The increase of reflectance in visible bands might be associated with a breakdown of chloroplast and some visible foliar symptoms (Devadas et al., 2009; Sankaran et al., 2010), whereas the increase of reflectance in mid-infrared bands and shortwave near-infrared bands might be related to loss of water content (Pu et al., 2004; Weng, 2011).

The biology of the PM and YR pathogens and AH insect, and the characteristics of host-pathogen/insect interaction can present different physiological changes and symptoms among diseases/insects (Cooke et al., 2006). It should be noted that the spectral responses of the two diseases and the insect in the visible spectral region (VIS, 400–700 nm) are closely associated with their visual symptoms (Devadas et al., 2009). For example, leaves that are infected by PM and YR have noticeable pustules, which produce a clear spectral response in the VIS region. However, as shown in Fig. 3b, the shapes of their ratio curves varied significantly in this region, given that PM produces white, fluffy mycelia, whereas YR produces yellow or reddish-purple longitudinal streaks on leaves. Compared to YR and PM, AH has a relatively weaker response in the VIS region (Fig. 3b), given that the piercing and sucking of aphids produces less apparent injury on leaves. In addition to the original bands, four VIs ( $D_y$ , GI, NDVI, and PRI) were identified as the most discriminating features as they exhibited significant difference between each stressor pair. It was also found that bands corresponding to these VIs are all located in the VIS spectral region.

Another major difference among PM, YR and AH is located in the near-infrared region. YR exhibits an increase in reflectance whereas PM and AH show a decrease in reflectance in this spectral region. The leaf tissue affected by the two diseases and insect tended to be destroyed; as a result, multiple scattering became weaker (Riedell and Blackmer, 1999; Graeff and Claupein, 2007). This led to a lower reflectance than non-infested leaves in the NIR region. For PM, cellular destruction tended to lower the multi-scattering of leaves, resulting in a decrease of the reflectance in NIR region.

The responses to foliar water content at 1390–1520 nm (mid-infrared region) and 1860–2080 nm (SWIR region) varied among the three stressors. The PM and YR exhibited varied responses, given the different extent of water loss due to disease severity. AH caused a loss of foliar water content given that aphids are fed by piercing the leaf and sucking out the juice from the infested leaf. This effect produced a much weaker spectral response in the SWIR region than the other two diseases. The different spectral responses among the two diseases and insect as stated above became the basis for distinction between them from interpretation of the spectrum.

In this study, by implementing the overlapping procedure to both the original bands and SFs, the most appropriate and effective features were identified for stressor discrimination. Validation results for the discriminant model showed that a generally satisfactory differentiation among PM, YR and AH can be achieved based on hyperspectral measurements and analysis. However, the model produced unequal success rates among different stressors (Table 3). YR could be accurately identified, with the P.'s a. (%) and U.'s a. (%) over 90% for both the training and validation samples. The classification accuracies for PM and AH were lower than YR. As shown in Fig. 3d and Fig. 4a, more bands exhibited significant difference between YR and the other two stressors than between PM and AH. Particularly, the different change direction in the NIR region between YR and the other two stressors (increase or decrease) also indicated that YR might easily differentiate from other stressors. For PM and AH, the two stressors had generally similar spectral response over the spectrum with varied response magnitude. This characteristic might be associated with a specific confusion between them as reflected by low U.'s a. (%) for both the training and validation samples of PM and AH (Table 3). Moreover, it was found that the samples with lower damage index ( $5\% < DI < 30\%$ ) had a higher chance of misclassification (OA = 54%), compared to severely damaged samples ( $DI > 30\%$ , OA = 91%). According to the validation results, another problem with the discriminant model is that the normal samples were frequently misclassified as AH samples, resulting in low P.'s a. (%) for normal samples and high U.'s a. (%) for AH samples (Table 3). Comparing the two diseases, the physiological changes that are produced by AH are relatively moderate, which lead to weaker spectral response and made AH infested samples more difficult to distinguish (see Table 3, the U.'s a. (%) of AH is relatively low for validation samples).

In non-contact differentiation among stresses, Kuckenberg et al. (2009) also attempted to detect and differentiate nitrogen-deficiency, powdery mildew and leaf rust in wheat at leaf level by laser-induced chlorophyll fluorescence at 690 nm (F690) and 730 nm (F730). The results showed that the OAA varied from 59% to 85%. While in this study, the OAA in differentiating the three diseases/pests achieved 75%, which is reasonable comparing with the above stated results using a more complex detection system. Also in practice, the fluorescence based instruments are more expensive and less portable than spectral sensors. Therefore, the method developed in this study relying on spectral analysis provides a good alternative in application.

In addition to discriminating diseases and insect, a regression model was constructed to give an estimation of stress intensity, which facilitates recommendations for insecticides/fungicides application rates. A total of eight bands/SFs were identified as the most efficient variables through a combination of correlation analysis and autocorrelation analysis (see Section 3.3). The validation result of the regression model showed that the model could produce promising accuracy in estimating leaves' DI (Fig. 5). The advantageous characteristics of PLSR might account for the reasonable performance of the model. The PLSR is a modification of principal components regression. In PLSR, the eigenvectors are aligned to balance the two objectives of explaining both response

variation and predictor variation as much as possible (Pu, 2012). This method is particularly suitable for processing independent variables that present a strong correlation and redundant information among them (Jones et al., 2010; Song et al., 2011)

In this study, the potential of hyperspectral information in differentiating two diseases and one insect of winter wheat were illustrated, determining the damage severity for leaves. The next step in bringing this method into a real world application is to extend the model from leaf level to canopy level. Several influences at canopy level would complicate spectral characteristics. For example, soil background, the influences of plant organs aside from leaves (e.g. stalk, spike), and leaf morphological factors (e.g., leaf angle) would have certain impacts to canopy spectral characteristics of plants. With the considerations, the discriminant model and the damage severity estimation model can be developed for use at canopy scale, which would extend the application of remote sensing to new precision agricultural implementations.

## 5. Summary and conclusions

In this paper, based on foliar hyperspectral measurements that were made for PM, YR and AH leaf samples, the feasibility has been demonstrated in using spectral means in differentiating two diseases and one insect, and estimating stress intensity at leaf level. Major conclusions can be summarized as follows:

1. Using the spectrum of normal leaves as a reference, each disease/insect possesses a characteristic and specific reflectance curve. Their spectral responses are closely associated with changes in foliar physiological status that are produced by infection of diseases/insect. Based on a correlation analysis and an independent *t*-test, four variables ( $Ref_{752-758nm}$ ,  $Ref_{1893-1905nm}$ ,  $D_y$ , and  $GI$ ) were identified as suitable for stressor discrimination, whereas eight variables ( $Ref_{752-758nm}$ ,  $Ref_{2268-2303nm}$ ,  $\lambda_r$ ,  $AREA_{550-770}$ ,  $GI$ ,  $CARI$ ,  $SIPI$ , and  $AI$ ) were selected for disease severity estimation.
2. Based on the selected bands/SFs above, the discrimination model was established using FLDA. The accuracies for both training and validation samples were satisfactory with  $OA(\%)$  reaching 0.75 and  $kappa$  of 0.67. The discriminant model produced a higher accuracy for YR than PM and AH, with classification accuracy over 90%.
3. In addition to the discrimination of disease/insect, damage intensity can be accurately estimated through a PLSR model. This estimation could be a post-procedure following discrimination among stressors. Validation results of the PLSR model showed that the model is able to produce satisfactory DI estimates, with  $R^2$  of 0.73 and RMSE of 0.148.
4. To facilitate practical implementation, it is necessary to extend the features and models for stressor discrimination and intensity estimation from leaf level to canopy level. An experiment will be conducted for further analysis. The influences from circumstantial effects and plant morphology should be considered and described to address these concerns in the future.

## Acknowledgments

This work was financially supported by the National Natural Science Foundation of China (Project no. 41301476 and 41271412), Beijing Natural Science Foundation (Project no. 4132029). The authors are grateful to Mr. Weiguo Li and Mrs. Hong Chang for data collection.

## References

- Bravo, C., Moshou, D., West, J., McCartney, A., Ramon, H., 2003. Early disease detection in wheat fields using spectral reflectance. *Biosystems Engineering* 84, 137–145.
- Cooke, B.M., Jones, D.G., Kaye, B., 2006. *The Epidemiology of Plant Diseases*. Springer, Netherland.
- Delwiche, S.R., Kim, M.S., 2000. Hyperspectral imaging for detection of scab in wheat. *Biological Quality and Precision Agriculture II* 4203, 13–20.
- Devadas, R., Lamb, D.W., Simpfordorfer, S., Backhouse, D., 2009. Evaluating ten spectral vegetation indices for identifying rust infection in individual wheat leaves. *Precision Agriculture* 10, 459–470.
- Escalante-Ramírez, B., 2012. *Remote Sensing Applications*. In: Huang, W.J. (Ed.), *Crop Disease and Pest Monitoring by Remote Sensing*. Intech, Croatia, Chapter 2.
- Gao, B.C., 1996. A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment* 58, 257–266.
- Graeff, S., Claupein, W., 2007. Identification and discrimination of water stress in wheat leaves (*Triticum aestivum* L.) by means of reflectance measurements. *Irrigation Science* 26, 61–70.
- Graeff, S., Link, J., Claupein, W., 2006. Identification of powdery mildew (*Erysiphe graminis* sp. *tritici*) and take-all disease (*Gaeumannomyces graminis* sp. *tritici*) in wheat (*Triticum aestivum* L.) by means of leaf reflectance measurements. *Central European Journal of Biology* 1, 275–288.
- Gong, P., Pu, R., Heald, R.C., 2002. Analysis of in situ hyperspectral data for nutrient estimation of giant sequoia. *International Journal of Remote Sensing* 23, 1827–1850.
- Huang, W.J., David, W.L., Niu, Z., Zhang, Y.J., Liu, L.Y., Wang, J.H., 2007. Identification of yellow rust in wheat using in-situ spectral reflectance measurements and airborne hyperspectral imaging. *Precision Agriculture* 8, 187–197.
- Jones, C.D., Jones, J.B., Lee, W.S., 2010. Diagnosis of bacterial spot of tomato using spectral signatures. *Computers and Electronics in Agriculture* 74, 329–335.
- Kim, M.S., Daughtry, C.S.T., Chappelle, E.W., McMurtrey, J.E., 1994. The use of high spectral resolution bands for estimating absorbed photosynthetically active radiation (APAR). In: *In Proceedings of the 6th International Symposium on Physical Measurements and Signatures in Remote Sensing*, France: Val d'Isere, pp. 299–306.
- Kuckenberger, J., Tartachnyk, I., Noga, G., 2009. Detection and differentiation of nitrogen-deficiency, powdery mildew and leaf rust at wheat leaf and canopy level by laser-induced chlorophyll fluorescence. *Biosystems Engineering* 103, 121–128.
- Luo, J.H., Huang, W.J., Yuan, L., Zhao, C.J., Du, S.Z., Zhang, J.C., Zhao, J.L., 2012. Evaluation of spectral indices and continuous wavelet analysis to quantify aphid infestation in wheat. *Precision Agriculture*, 1–11.
- Mclachlan, G.J., 2004. *Discriminant analysis and statistical pattern recognition* Wiley-Interscience. Hoboken, New Jersey, USA.
- Merton, R.N., 1998. Monitoring community hysteresis using spectral shift analysis and the red-edge vegetation stress index. In: *In Proceedings of the Seventh Annual JPL Airborne Earth Science Workshop*, pp. 12–16.
- Mirik, M., Michels Jr., G.J., Kassymzhanova-Mirik, S., Elliott, N.C., Catana, V., Jones, D.B., Bowling, R., 2006. Using digital image analysis and spectral reflectance data to quantify damage by greenbug (Hemiptera: Aphididae) in winter wheat. *Computers and Electronics in Agriculture* 51, 86–98.
- Mirik, M., Michels Jr., G.J., Kassymzhanova-Mirik, S., Elliott, N.C., 2007. Reflectance characteristics of Russian wheat aphid (Hemiptera: Aphididae) stress and abundance in winter wheat. *Computers and Electronics in Agriculture* 57, 123–134.
- Moshou, D., Bravo, C., West, J., Wahlen, S., McCartney, A., Ramon, H., 2004. Automatic detection of 'yellow rust' in wheat using reflectance measurements and neural networks. *Computers and Electronics in Agriculture* 44, 173–188.
- Naidu, R.A., Perry, E.M., Pierce, F.J., Mekuria, T., 2009. The potential of spectral reflectance technique for the detection of Grapevine leafroll-associated virus-3 in two red-berried wine grape cultivars. *Computers and Electronics in Agriculture* 66, 38–45.
- Pu, R., Ge, S., Kelly, N.M., Gong, P., 2003. Spectral absorption features as indicators of water status in coast live oak (*Quercus agrifolia*) leaves. *International Journal of Remote Sensing* 24, 1799–1810.
- Pu, R., Foschi, L., Gong, P., 2004. Spectral feature analysis for assessment of water status and health level in coast live oak (*Quercus agrifolia*) leaves. *International Journal of Remote Sensing* 25, 4267–4286.
- Pu, R., 2012. Comparing canonical correlation analysis with partial least square regression in estimating forest leaf area index with multitemporal landsat TM imagery. *GIScience & Remote Sensing* 49, 92–116.
- Riedell, W.E., Blackmer, T.M., 1999. Leaf reflectance spectra of cereal aphid-damaged wheat. *Crop Science* 39, 1835–1840.
- Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., 1973. Monitoring vegetation systems in the great plains with ERTS. In: *Third ERTS symposium NASA SP-351*, NASA, Washington DC, pp. 309–317.
- Sankaran, S., Mishra, A., Ehsani, R., Davis, C., 2010. A review of advanced techniques for detecting plant diseases. *Computers and Electronics in Agriculture* 72, 1–13.
- Song, S.L., Gong, W., Zhu, B., Huang, X., 2011. Wavelength selection and spectral discrimination for paddy rice, with laboratory measurements of hyperspectral leaf reflectance. *ISPRS Journal of Photogrammetry and Remote Sensing* 66, 672–682.
- Strange, R.N., Scott, P.R., 2005. Plant disease: a threat to global food security. *Annual Reviews Phytopathology* 43, 83–116.



- Weng, Q.H., 2011. *Advances in environmental remote sensing*. In: Pu, R.L., Gong, P. (Eds.), *Hyperspectral Remote Sensing of Vegetation Bioparameters*. CRC Press, Boca Raton, FL, Chapter 5.
- Yang, Z., Rao, M.N., Elliott, N.C., Kindler, S.D., Popham, T.W., 2005. Using ground-based multispectral radiometry to detect stress in wheat caused by greenbug (Homoptera: Aphididae) infestation. *Computers and Electronics in Agriculture* 47, 121–135.
- Zarco-Tejada, P.J., Berjón, A., López-Lozano, R., Miller, J.R., Martín, P., Cachorro, V., et al., 2005. Assessing vineyard condition with hyperspectral indices: leaf and canopy reflectance simulation in a row-structured discontinuous canopy. *Remote Sensing Environment* 99, 271–287.
- Zhao, C.J., Huang, M.Y., Huang, W.J., Liu, L.Y., Wang, J.H., 2004. In: King, R., Datcu, M. (Eds.), *Analysis of winter wheat stripe rust characteristic spectrum and establishing of inversion models*. *Proceedings of Geoscience and Remote Sensing Symposium 6, Alaska, USA*, pp. 4318–4320.
- Zhang, J.C., Pu, R.L., Wang, J.H., Huang, W.J., Yuan, L., Luo, J.H., 2012a. Detecting powdery mildew of winter wheat using leaf level hyperspectral measurements. *Computers and Electronics in Agriculture* 85, 13–23.
- Zhang, J.C., Pu, R.L., Huang, W.J., Luo, J.H., Wang, J.H., 2012b. Using in-situ hyperspectral data for detecting and discriminating yellow rust disease from nutrient stresses. *Field Crops Research* 134, 165–174.