

# Monitoring and prediction of wheat stripe rust based on hyperspectral and meteorological data

Chunyan Ma (✉ [mayan@hpu.edu.cn](mailto:mayan@hpu.edu.cn))

Henan Polytechnic University

Jinhui Dong (✉ [1284234391@qq.com](mailto:1284234391@qq.com))

Henan Polytechnic University

Fan Ding

Henan Polytechnic University

Changchun Li

Henan Polytechnic University

---

## Research Article

**Keywords:** wheat, stripe rust, monitoring, vegetation index, meteorological factors, spectral sensitivity

**Posted Date:** March 2nd, 2022

**DOI:** <https://doi.org/10.21203/rs.3.rs-1410519/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

This study integrated spectral sensitivity values and absorption characteristic parameters based on canopy hyperspectral data and meteorological data, e.g. air volume, average air temperature, relative humidity, sunshine hours, and precipitation. These include the total area of absorption peaks, the area at the left end thereof, absorption depth, and symmetry. This study analyses correlations between these factors and a disease index based on partial least squares method and constructs the small of flag picking period, early and middle grouting period, and the entire growth period. The disease condition inversion model of wheat stripe rust was used to verify the accuracy of the model. The correlation of NDVI, RDVI, TVI, DVI, RVI, and meteorological factors with stripe rust were analysed simultaneously. DVI and average temperature values with a high correlation degree and stripe rust occurrence probability were selected, a prediction model of wheat stripe rust occurrences was established, and the accuracy of the model was verified. Results showed that the average  $R^2$  coefficients of the inversion model for the disease index in three growth periods and the entire growth period were 0.94 and 0.92, respectively; the average RMSE values were 0.068 and 0.063, respectively, with high inversion accuracy, which indicated that the prediction model based on spectral data and meteorological data had high prediction accuracy and good prediction effect. It provides new technology and method for monitoring and forecasting crop diseases and insect pests.

## 1 Introduction

Crop diseases and insect pests can cause huge losses to agricultural production. The United Nations Food and Agriculture Organization estimates that the annual reduction of world food production caused by diseases and insect pests is as high as 24% (Lu et al., 2019). Wheat stripe rust is a type of serious biological disaster that has a long-term impact on wheat production safety in China. This type of rust has a wide incidence area, strong epidemic capabilities, high incidence rates, and can cause considerable agricultural harm. Its spores can be spread widely by wind. Wheat rust can occur many times in the process of wheat growth, resulting in the large-scale reduction of wheat yield and quality. Therefore, effective real-time and accurate monitoring and prediction of wheat stripe rust and timely control measures can effectively reduce the impact of disease. At present, the field sampling survey is the main method for monitoring and forecasting crop diseases and insect pests in China (Gao et al., 2009). These traditional methods cost considerable human and material resources, and the representativeness and timeliness of disease points are not satisfactory during investigation processes. Investigators are inevitably subjective, and meeting the needs of large-scale real-time monitoring and predictions of diseases and pests is difficult (Wang et al., 2007). Remote sensing technology has a non-contact requirement and a high throughput and can provide an effective scientific method for rapid, efficient, economic, non-destructive, and timely monitoring and predictions of wheat stripe rust. Hyperspectral technology has both a hyperspectral and spatial resolutions. The health status of crops can be determined as rapidly as possible through map analysis, and severity of diseases and insect pests in a

large area can be obtained, data support can be provided for timely and accurate prevention and control of diseases and insect pests, and a foundation can be set for crop disease monitoring.

In recent years, many studies have monitored and predicted crop diseases and insect pests. Huang et al. (2019) extracted 26 characteristic factors from Sentinel-2 image data, analysed their correlations with wheat stripe rust and constructed a disease index inversion model based on a BP neural network. Moshou et al. (2004, 2005) used the vehicle imaging hyperspectral remote sensing system to monitor wheat diseases in real time. Jing et al. (2019) used the hyperspectral data of the canopy to perform differential processing, build a reflectivity differential spectral index, and build an inversion model of wheat stripe rust disease based on the random forest and BP neural network algorithms. Chen et al. (2011) used a TM image to identify and monitor the severity of cotton disease by establishing a spectral index. Yuan et al. (2013) discussed the feasibility of using spectral analysis to distinguish, identify and monitor the severity of two diseases based on the spectral data of powdery mildew and stripe rust. Zhan et al. (2010) established a relationship model between a rice ear spectrum and a disease grade by using the ear spectrum of rice glume blight. Luo et al. (2010) and Jiang et al. (2010) measured the canopy spectrum of wheat stripe rust with an ASD ground object spectrometer; next, they constructed a spectrum index suitable for monitoring the disease that more effectively retrieved the incidence degree and range of stripe rust compared to existing methods. Guo et al. (2009) constructed a stripe rust sentiment index regression model at 690 and 850 nm bands with different spectral characteristics such as reflectivity, SDr, NDVI, and RVI. Jiang et al. (2010) used the first-order differential maximum method to obtain the red edge and yellow edge positions of the spectrum, and compared and analysed the red edge reflectance, yellow edge reflectance, and the prediction ability of a combined disease index, and identified wheat stripe rust 12 days in advance. Hu et al. (2019) predicted the possibility of wheat stripe rust in Gansu province based on meteorological data from 1995 to 2012. Zhang et al. (2019) analysed the impact of climate factors such as sea temperature, El Niño events, atmospheric circulation, temperature, rainfall, humidity, wind, and light at the occurrence time, infection speed, transmission, and distribution of diseases and insect pests; they also constructed an inversion model for diseases and insect pests. To improve the practicability of the mathematical statistical model, Wang et al. (2010) used the meteorological data of diseases in Shanxi province from many years and established a meteorological condition prediction model for the occurrence of winter wheat diseases in Shanxi Province. Hu et al. (2000) studied the effect of rainfall in April and the average temperature in April on the occurrence of wheat stripe rust in Hanzhong. Qian et al. (2005) explored the relationship between the epidemic of wheat powdery mildew in China and the key atmospheric circulation characteristic factors and their possible influence mechanisms; they concluded that the key circulation characteristic factors in the early stage were strong indicators of the occurrence and epidemic potential of wheat powdery mildew in China. Xu et al. (2018) analysed the relationship between the average temperature and the occurrence of wheat stripe rust and established a statistical model of overwintering meteorological conditions based on the disease index and disease field rate of wheat stripe rust.

These existing studies only used spectral characteristics or meteorological factors to analyse the relationships between them and the degrees of crop diseases and insect pests, as well as to build a

disease index estimation model and a disease occurrence prediction model. There are few studies on disease degree monitoring and predictions based on comprehensive spectral characteristics and meteorological factors. When wheat is infected by stripe rust fungus, as caused by a lack of water and nutrition, the content of photosynthetic pigments and the internal structure of the leaves change to a certain extent, and the reflection spectral curve of the canopy also changes to a certain extent. The severity of the disease can differ between plants, and the spectral curve may change at different extents. At the same time, the occurrence and epidemic of wheat stripe rust is closely related to environmental meteorological factors, which is the basic tenet for monitoring and predicting the occurrence of the disease. Therefore, this study uses the hyperspectral data of wheat canopies combined with meteorological data to explore the spectral characteristics of said canopies in different growth stages and different susceptibility levels, extract the spectral band characteristics with high sensitivity to wheat disease, and analyse the response mechanism of remote sensing spectrum of wheat stripe rust. Concurrently, this study analyses regional meteorological data such as average wind speed, average temperature, and relative humidity observed during periods of wheat stripe rust, sunshine hours, rainfall, and other factors related to the disease status among regions. The meteorological environment transmission mechanism of the wheat stripe rust disease was analysed. The disease status, degree, and trend of wheat stripe rust events were monitored and predicted.

## 2 Materials And Methods

### 2.1 Overview of the study area

The study area is located in Longnan, Tianshui, Pingliang, Dingxi, and Qingyang in southeast Gansu Province. The specific location is N32.551°–37.072°, E103.363°–108.705°. As shown in Fig. 1, the study area is part of the Gansu Qinghai region, which is one of the 9 endemic regions in the northwest extreme of North China. The area has a complex terrain, cool and rainy summers, large temperature differences between day and night, and large amounts of rain and dew. The wheat fields with different altitude gradually mature from the end of June to the end of September. The self-produced wheat seedlings as well as late mature winter and spring wheat overlap and coexist, which greatly facilitates the transmission of individual pathogens. The most bacteria are present in summer, and the study area has become the main base for the pathogens of winter wheat autumn seedlings in North China and Northwest China.

Please put Fig. 1 about here

### 2.2 Test design

The test field is 100 m long and 50 m wide. It is divided into 18 small areas, with 9 control areas and 9 disease areas, each with an area of 24 m<sup>2</sup>. In the disease area, a spraying method was used to inoculate against the stripe rust pathogen, and the number of bacteria inoculated in each plot was equal. Inoculation began at 17:00 p.m. and the prepared spore suspension was dispersed onto wheat leaves with a small sprayer. After inoculation, the plastic film was covered overnight, and the plastic film was

removed around 9:30 in the morning to complete the inoculation. After 10 days of inoculation, the wheat in the inoculated area showed symptoms gradually. At the flag picking stage, all diseases had occurred in the inoculated area. In the control area, the stripe rust was not inoculated, and wheat was infected under natural conditions.

## 2.3 Data acquisition and processing

The canopy hyperspectral data collection process and disease index investigation were each performed in the flag picking stage, early filling stage, and middle filling stage.

### 2.3.1 Canopy hyperspectral data acquisition and processing

An ASD FieldSpec Pro spectrometer was used to collect the hyperspectral data of the canopy. The measurement time was 10:30 – 14:30 Beijing time. The weather was clear and cloudless. During the observation period, the probe of the spectrometer was placed vertically downward and always maintained a distance of 1.6 m from the ground. The field angle of the probe was 25° and the field of view of the ground was 50 cm in diameter. The crop canopy reflectance was measured and averaged 20 times in each plot. Before and after the measurement of each cell, the reference plate was calibrated.

### 2.3.2 Disease index survey

After symptoms occurred, the disease index was investigated immediately after the canopy hyperspectral data were measured every 10 days. The five points method was used to investigate the disease index. Five symmetrical investigation points were selected for each plot; each point covered an area of approximately 2 m<sup>2</sup>, and 20 wheat plants were selected for each point to investigate the disease degree. The degree of occurrence was expressed by severity, i.e., the relative percentage of the area occupied by *Fusarium* on the diseased leaves and the total area of leaves. The severity was divided into 8 gradients, i.e., 1%, 5%, 10%, 20%, 40%, 60%, 80%, and 100%. The number of wheat leaves for each severity level was recorded, and the disease index (DI) was calculated according to Eq. (1).

$$DI = \frac{\sum (x \times f)}{n \times \sum f} \times 100 \% (1)$$

where  $x$  is the extreme value of each gradient,  $n$  is the highest gradient value, and  $f$  is the number of blades of each gradient.

### 2.3.3 Meteorological data collection and processing

The daily precipitation, relative humidity, temperature, and average wind speed from March 1 to July 15 of each year were analysed as meteorological data provided by the National Meteorological Centre. The daily meteorological data was interpolated into weekly data, the relative humidity and temperature were taken as the average value, and the precipitation data was taken as the weekly total value.

The daily wind speed data was used for the calculation of the wind direction influence value. The influence value of the wind direction is defined as the cosine value of the angle between the local wind

direction and other two directions. The wind direction influence value of one place to other two places is equal to the product of wind speed of one place and the wind direction influence value of other two places., The wind direction influence values of Longnan on Tianshui, Dingxi, Pingliang and Qingyang were calculated according to the time sequence and wind speed of the initial occurrence of the disease in the study area.

## 2.4 Methods

### 2.4.1 Calculation of spectral sensitivity

Kobayashi et al. (2001) proposed the concept of spectral sensitivity to analyse the different spectral responses disease index plants under ear neck blast stress; this was done to more easily compare the spectral curves of different disease indexes. When the spectral sensitivity is positive in a certain band, the spectral reflectance of the stressed plant is higher than that of the normal plant, and the higher the spectral sensitivity, the more significant the difference between the spectral reflectance of the stressed plant and that of the normal plant, and vice versa. The calculation method of spectral sensitivity is shown in Eq. (2):

$$S_S = \frac{S_T - S_N}{S_N}$$

2

where  $S_S$  represents spectral sensitivity;  $S_T$  represents the spectral reflectance of stressed plants, and  $S_N$  represents the spectral reflectance of normal plants.

### 2.4.2 Removal of canopy spectral continuum

Continuum removal, also known as envelope removal, uses the actual spectral band value divided by the corresponding band value on the continuum. After continuous removal of the original spectral curve, the important characteristics of the responses of various groups in the leaf tissue structure, pigment content, water, and protein to the reflection spectrum can be obtained. These absorption characteristic parameters mainly include the total area "a" of the absorption peak, the area "A<sub>1</sub>" of the left end of the absorption peak, the absorption depth "D," and the symmetry degree "S" (Zheng et al.,2019).

### 2.4.3 Leave-one-out cross validation

Leave-one-out cross validation (LOOCV) can avoid the problem of having to choose between modelling samples and validation samples in an experimental design. Every sample participates in the test and training, which can address the drawbacks of using training samples with low richness. LOOCV is an effective method to evaluate the generalization ability and reliability of a regression model. Its basic principle is to assume that there are  $N$  samples, with each sample is regarded as a test set, and the rest of the samples ( $N-1$ ) are regarded as a training set, which are cycled  $N$  times.

## 2.4.4 Adaptive network fuzzy reasoning

An adaptive network fuzzy reasoning system, also known as an adaptive network-based fuzzy inference system (ANFIS), is a type of fuzzy inference system that integrates neural network adaptability (Zhang et al., 2019). This system synthesizes the neural network learning algorithm and the concise form of fuzzy reasoning; it also generates a numerical solution by learning the training data group. The structure of ANFIS includes five layers, as shown in Fig. 2. The inference system has two inputs of  $X$  and  $Y$ , the output is  $f$ , and each node in the same layer has similar functions.  $O_{1,i}$  is used to represent the output of the  $i$  node in the first layer, and so on.

Please put Fig. 2 about here

As shown in Fig. 2, the first layer represents the selection and fuzzification of input parameters, which is the first step in establishing fuzzy rules. Each node  $i$  in this layer is a square node represented by a node function.

$$O_{1,i} = \mu_{A_i}(x), i = 1, 2 \quad O_{1,i} = \mu_{B_{(i-2)}}(y), i = 3, 4$$

$O_{1,i}$  represents the membership function of fuzzy set A. The second layer represents the calculation of the excitation intensity of the fuzzy rule, multiplying the membership degree of the input signal; the output is:

$$O_{2,i} = \omega_i = \mu_{A_i}(x) \mu_{B_i}(y), i = 1, 2$$

The second layer is the calculation of the fuzzy rule excitation intensity, which multiplies the membership degree of the input signal, and the output result is as follows,

$$O_{2,i} = \omega_i = \mu_{A_i}(x) \mu_{B_i}(y) \quad i = 1, 2$$

The third layer is the normalization calculation of the applicability of each rule of the node in this layer, i.e., the ratio of the sum of rule  $\omega_i$  of rule I and all rule  $\omega$  of node  $i$  is calculated as follows:

$$O_{3,i} = \frac{\omega_i}{\omega_1 + \omega_2} \quad i = 1, 2$$

The fourth layer indicates that each node  $i$  of this layer is an adaptive node, and its output is as follows:

$$O_{4,i} = \wp_i(p_i x + q_i y + r_i) \quad i = 1, 2$$

The fifth layer indicates that the single node of this layer is a fixed node, and calculates the total output of all input signals; specifically:

$$O_{5,i} = \frac{\sum \omega_i f_i}{\sum \omega_i} \quad i = 1, 2$$

## 2.4.5 Grey relational analysis

Grey relational analysis (GRA) is a type of grey system analysis method. According to the similarity or dissimilarity degree of development trend among factors, i.e. 'grey relational degree', used as a method to measure the degree of correlation among factors, GRA quantitatively describes the change trend between data (Yu et al., 2019). If the change trend in the two datasets is similar, the correlation between the data is considered high; conversely, non-similar trends indicate that the correlation between the data is low.

## 2.4.6 Partial least squares regression

Partial least squares regression (PLSR) is based on the idea of principal component regression. PLSR requires the linear function of the original independent variables  $x_1, x_2, \dots, x_n$ ; concurrently, considering its correlation with the corresponding variables, the regression selects the linear function of  $x_1, x_2, \dots, x_n$  as a new independent variable and the corresponding variables for regression (Jiang et al., 2003). This linear function is strongly correlated with the corresponding variables and is easy to calculate. The algorithm is based on the least squares method. In the algorithm, only the variables related to the strain are selected; not all linear functions of  $x_1, x_2, \dots, x_n$  are considered, and only the parts related to the strain are considered.

## 2.4.7 Akaike Information Criterion

Akaike information criterion (AIC) is an index used to measure the fitting performance of a statistical model (Xu et al., 2018), and its calculation method is shown in Eq. (3):

$$AIC = 2k - 2\ln L$$

3

where  $k$  is the number of model parameters and  $L$  is the likelihood function. The smaller AIC, the stronger the fitting effect.

## 2.4.8 Vegetation index construction

In this study, the ratio vegetation index (RVI), normalized vegetation index (NDVI), difference vegetation index (DVI), triangle vegetation index (TVI), and renormalization vegetation index (RDVI) are used to build the disease monitoring and prediction model. The calculation equation of each vegetation index is shown in Table 1 (Mistele et al., 2010; Kelly et al., 2006; Klein et al., 1972; Xue et al., 2004;)

Table 1  
Calculation method of vegetation index

Name	Calculation formula
RVI	$RVI=(NIR)/R$
NDVI	$NDVI=(NIR-R)/(NIR + R)$
DVI	$DVI = NIR-R$
TVI	$TVI=(0.5)((120)(NIR-G)-200 (R-G))$

Please put Table 1 about here

## 2.4.9 Correlation analysis

Correlation analysis is used to analyse the change trend of two or more groups of data for consistency to obtain whether the relationship between them is close as well as its degree; this approach often uses correlation coefficients for discriminant analysis. The correlation coefficient is obtained by dividing the covariance of two random variables by the standard deviation. If the correlation coefficient is between - 1 and 1, the greater the absolute value and the greater the correlation degree; however, when the coefficient is close to 0, this indicates that there is no correlation. The calculation method is shown in formula (4):

$$\rho_{X, Y} = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - E^2(X)} \sqrt{E(Y^2) - E^2(Y)}}$$

4

where  $\rho_{X, Y}$  represent the correlation coefficient, while  $\text{cov}(X, Y)$  and  $\sigma$  represent the covariance and standard deviation, respectively.

## 2.4.10 Model accuracy evaluation

The coefficient of determination ( $R^2$ ) and root mean square error (RMSE) were selected as the model accuracy evaluation index. The calculation method is shown in equations (5) and (6).

$$R^2 = \frac{(\sum_{i=1}^n Y_i - \bar{y})^2}{(\sum_{i=1}^n X_i - \bar{x})^2}$$

5

$$RMSE = \sqrt{\frac{\sum_{i=1, j=1}^n (x_i - y_j)^2}{n}}$$

6

where  $x_j$  is the measured value of the sample,  $y_j$  is the estimated value of the sample,  $\bar{y}$  is the mean value, and  $n$  is the number of samples. In general, the larger  $R^2$  is, the smaller RMSE is, indicating the superior fitting effect of the model.

## 3 Results And Discussion

### 3.1 Disease index inversion

#### 3.1.1 Analysis of spectral characteristics

The main symptoms of wheat infected with stripe rust fungus are that the leaves turn yellow and lose their green pigments, the chlorophyll is destroyed, water transpiration is greatly increased, the cell size, shape, and structure of the leaves greatly change, and the optical characteristics of the leaves and the canopy also change. The spectral curves of the control area and the disease area in the flag leaf stage, the early filling stage, and the middle filling stage are shown in Fig. 3.

Please put Fig. 3 about here

Figure 3 shows that: In the flag leaf stage, slight disease symptoms occurred. Under the influence of the pathogen, the chlorophyll content of the leaves decreased, the absorption bands of the blue and red areas of chlorophyll decreased, and the red band reflectance of wheat in the disease area increased at a slightly higher rate than that in the control area. In the 750–1300 nm band, because the stripe rust fungus destroyed the internal structure of wheat leaves, the spectral reflectance of wheat canopy in the disease area was lower than that in the control area, and the spectral curve showed the spectral characteristics of healthy and green plants as a whole.

In the early filling stage, the infection degree of stripe rust increased, the leaves turned green and yellow, the structure and tissue of leaves were damaged, the reflectance of near red band of wheat canopy spectrum was significantly lower than that of flag picking stage, the near red band waveform gradually showed a round shoulder shape, and the spectral curve shape of wheat canopy in the control area gradually revealed disease characteristics. In the 550–700 nm yellow light region, the spectral reflectance was slightly higher than that in the flag leaf stage, and the wave form in the near red band was similar to that in the disease region, thereby indicating that the stripe rust pathogen had spread in the flag leaf stage.

In the middle filling stage, the disease index of wheat was increased by 120% compared with that in the early filling stage. The reflectance of the yellow light area increased gradually, and the near red band reflectance was decreased significantly, and was even lower than that of the control area. During the growth period, the wheat in the control area showed slight disease symptoms, the reflectance in the yellow light area continued to rise, the in near red band reflectance continued to decline, the absorption

valley became shallow, the red edge transitioned to blue, and the red edge amplitude decreased significantly, showing obvious morbidity.

### 3.1.2 Spectral sensitivity analysis

By calculating the average value of spectral reflectance, the spectral sensitivity values of wheat in different growth stages can be obtained. The spectral sensitivity values of wheat in different growth stages were calculated and a spectral sensitivity curve was drawn according to Eq. (2). The histogram of spectral sensitivity at different growth stages was drawn. The results are shown in Fig. 4.

Please put Fig. 4 about here

The figure shows that, with the development of the growth period, the difference of spectral reflectance of the canopy hanged gradually as the disease degree worsened. At the flag leaf stage, symptoms of the disease appeared in wheat in the diseased area, but no spread occurred. The spectrum curve of wheat was normal,  $S_S$  was positive, and the reflectivity of wheat in the near red band range decreased,  $S_S$  was negative; At the early filling stage, the symptoms of wheat increased, and the reflectivity of wheat was overall higher,  $S_S$  was positive, it reached its maximum the 660 nm near red light. In the middle filling stage, reflectance gradually decreased in the range of visible light, and the  $S_S$  value increased from a negative to a positive value. The red edge amplitude of wheat in the 700–780 nm red edge band diseased, the  $S_S$  value changed to a negative value.

### 3.1.3 Analysis of absorption characteristics

After the spectral curve was normalized by the continuous removal, the absorption characteristic in each growth stage of the control area and the disease area were obtained. The results are shown in Table 2.

Table 2  
Calculation result of absorption characteristic of infected wheat

	Flag leaf stage		Early filling stage		Middle filling stage	
	CK	YR	CK	YR	CK	YR
	Control check	Yellow rust	Control check	Yellow rust	Control check	Yellow rust
Total area of absorption $A$	136.41	129.20	125.60	116.34	85.54	72.43
Absorption depth $D$	0.93	0.91	0.88	0.81	0.70	0.62
Symmetry degree $S$	0.68	0.69	0.69	0.69	0.68	0.68
The left area of the absorption $A_1$	93.34	88.41	86.81	81.26	58.32	49.33

Please put Table 2 about here

The table shows that when wheat was damaged by disease, it lacked nutrition and water, its spongy tissue was damaged, its leaf pigment proportion changed, the two absorption valleys in the visible light area were not obvious, and the reflection peak value in the near-infrared light area was reduced. With the growth of wheat, the degree of disease increased gradually, and the area and depth of absorption peak decreased gradually; the area and depth of the absorption peak in the control area were slightly lower than that in the disease area, and the degree of symmetry was slightly higher than that in the disease area.

### 3.1.4 Disease index inversion

The correlation coefficients of spectral sensitivity and the disease index of wheat in different wavebands were calculated in flag leaf stage, early filling stage, middle filling stage. The correlation curves of three growth stages were drawn respectively and the results are shown in Fig. 5.

Please put Fig. 5 about here

The correlation analysis results showed that the correlation between the spectral sensitivity and disease index was the highest at 987 nm, 1362 nm and 1386 nm, and the correlation coefficients were 0.80, 0.87, and - 0.84, respectively; all of these reached a significance level of 0.05. Therefore, based on PLSR method the spectral sensitivity values of these three bands were selected as independent variables and the disease index for linear regression analysis, and the disease index inversion model of wheat stripe rust at different growth stages was obtained. At the same time, the absorption characteristics were selected as independent variables for linear regression with the disease index to obtain the inversion model for the whole growth period. The results are shown in Table 3.

**Table 3** Inversion model of stripe disease index

Growth stage	Model	$R^2$	RMSE
Flag leaf stage	$Y_1=0.0602x+0.1133$	0.642	0.051
Early filling	$Y_2=-0.8313x+0.1944$	0.773	0.042
Middle filling	$Y_3=-0.9010x+0.8815$	0.714	0.030
Full growth	$Y=-1.14319-1.40329D-0.0627A_1+0.039823A+4.39297S$	0.954	0.08

Please put Table 3 about here

After testing, the  $R^2$  values of the disease inversion model in the flag leaf stage, early filling, middle filling, and full growth stages were 0.642,0.773,0.714and 0.954,respectively, with RMSE values of 0.051,0.042, 0.030 and 0.080 respectively. The inversion model was verified using the left one cross validation method, and 27 tests and trainings were conducted in total to obtain the average coefficient of determination ( $R^2$ ) and root mean square error (RMSE) of the inversion model. The  $R^2$  values of the exponential inversion models were 0.94 and 0.92, respectively, and the RMSE values were 0.068 and

0.063, respectively, indicating that the model showed a strong fit and displayed a certain reliability and generalization ability.

## **3.2 Prediction of Stripe Rust**

### **3.2.1 Analysis of spatial characteristics of stripe rust**

The survey data of the stripe rust occurrence area and the disease degree were interpolated using the global polynomial interpolation method, and the results are shown in Fig. 6. The area of disease occurrence was 666.6 ha, and the degree of disease was divided into three levels: sporadic occurrence, spot occurrence, and spread. The survey interval was one week.

Please put Fig. 6 about here

Figure 6 shows that in 2010, the total area of disease occurrence in the Tianshui area was the largest and the occurrence was serious, followed by the Longnan area; next, the Pingliang area had less area, but its disease status was the most serious, and the Dingxi and Qingyang areas had the smallest areas and the lowest severity. In 2011, the number of districts and counties in the Longnan area was the highest, and the incidence was mostly epidemic. The number of districts and counties in Tianshui area was less than that of Longnan, but the latter's area and disease status were higher and more serious, respectively, than in the latter. In 2012, the overall disease incidence in the study area was more severe than in the previous two years. In the study area, the Longnan area had the most diseased areas and counties, followed by the Tianshui area. Diseases were found in Pingliang and Qingyang in the same period. The disease area in Pingliang was relatively large and the disease severity was serious. Disease symptoms appeared last in the Dingxi area. However, its disease status was also serious, and the epidemic spread rapidly in the area.

### **3.2.2 Analysis of meteorological factors for occurrence of stripe rust**

For the meteorological data obtained at stations in the study area, after removing the abnormal values, according to the disease survey cycle, the interpolation showed weekly average data, which corresponded to the disease survey point. Radial basis functions were used to interpolate the meteorological data, and the kernel function selected a regular spline function. The results are shown in Fig. 7.

Please put Fig. 7 about here

Figure 7 shows that in 2010, Longnan had a greater impact on wind volume in Dingxi and Tianshui, followed by Qingyang, and Pingliang was the weakest. Longnan, Tianshui, Pingliang, and Qingyang had high average temperatures, and Dingxi had slightly lower average temperatures than in other regions. In most parts of the study area, the average relative humidity was high, and it was highest in Tianshui. The areas with high sunshine hours were mainly concentrated in most areas of Tianshui, Pingliang, Dingxi and Qingyang. The sunshine hours in the Longnan area were slightly lower. The area was mainly

concentrated in the eastern part of Dingxi, the western part of Pingliang, and the northern part of Tianshui. The precipitation in other areas was not considerably different.

In 2011, the wind velocities of Dingxi and Tianshui were affected the most by wind speed in Longnan, followed by the Qingyang area; Pingliang had the weakest effect. The area with the highest average temperature was mainly concentrated in the central part of the research area, and the average temperatures in other areas were between 13 °C to 15 °C; the Dingxi and Pingliang western regions showed a relatively high humidity, with an average humidity of more than 68%, and other regions had an average humidity of about 58%; areas with high sunshine hours were mainly located in western Tianshui, and the sunshine hours in Dingxi and northern Longnan were slightly below those of the Tianshui area. Pingliang and most parts of Longnan had the fewest sunshine hours. The area with the largest average precipitation was located in north-western Dingxi, followed by Longnan and western Tianshui. The other areas had weaker precipitation, and the least precipitation occurred in Qingyang.

In 2012, the wind speed in Longnan had the largest impact on the wind volume in the central area of Tianshui, and the impact values of wind volume in other areas were slightly lower than that in Tianshui. The area with the highest average temperature was located in north-eastern Tianshui; the average temperatures in other areas were between 10 °C and 13 °C. Most areas in Tianshui, Dingxi, and Pingliang had relatively high humidity, and the average humidity in other areas was approximately 45%, which was slightly lower than this area. The sunshine hours in Tianshui, Pingliang, and Qingyang were relatively high. The annual hours were flat and the southern Longnan region had the lowest sunshine hours. The average precipitation was basically the same as in 2011. The areas with the highest precipitation were concentrated in Dingxi and Tianshui; Longnan, Pingliang, and Qingyang were slightly lower in precipitation.

Wheat stripe rust spores drifted with the wind and impregnated wheat at a temperature of 14–15 °C via water droplets and water films. The average temperature of the study area from 2010 to 2012 did not exceed 20 °C, with high humidity, moderate precipitation, and lower sunshine hours than other regions in Gansu Province. The spatial distribution trend in meteorological data was highly consistent with the trend in the disease survey data. A correlation analysis was performed on the meteorological and disease data at the survey sites in the study area, and the disease occurrence status showed a significant correlation with the average temperature (0.831); the occurrence area was significantly correlated with the influence of wind volume, with a correlation coefficient of 0.528. The above analysis results show that the meteorological environment in the study area provides favourable conditions for the spread of wheat stripe rust disease, and the meteorological factors are closely related to stripe rust disease epidemics.

### **3.2.3 Selection of vegetation index and meteorological environment factor**

First, we used the GRA method to analyse the correlation degree between the vegetation indexes NDVI, RDVI, TVI, DVI, RVI, the air volume value, average temperature, relative humidity, sunshine hours,

precipitation, other meteorological factors, and the disease index, and ranked the correlation degree. The results are shown in Table 4.

Table 4  
Ranking of grey correlation degree of vegetation index

Vegetation index	Gray correlation value	PLS - AIC value	Sort
NDVI	0.807	-77.57	4
RVI	0.731	-244.14	5
TVI	0.818	-245.33	3
DVI	0.963	-248.65	1
RDVI	0.883	-235.42	2

Please put Table 4 about here

Second, the combined PLS and AIC method was used to rank the data according to the correlation degree, calculate the AIC value of the PLS model in turn, and evaluate the modelling ability of vegetation index and meteorological factors. The results are shown in Table 5.

Table 5  
Ranking of grey correlation degree of meteorological factor

Meteorological factor	Gray correlation value	PLS - AIC value	Sort
Precipitation	0.874	19.22	5
Mean wind speed	0.964	14.32	2
Sunshine hour	0.958	17.2	3
Mean temperature	0.974	12.91	1
Relative humidity	0.956	18.97	4

Please put Table 5 about here

Tables 4 and 5 show that the vegetation index with the largest grey correlation value was DVI, the grey correlation value was 0.963, and the AIC value was the smallest at -248.65. The meteorological factor with the largest grey correlation value was the average temperature, the grey correlation value was 0.974, and its AIC value was the smallest at 12.91.

### 3.2.4 Predictions of Wheat Stripe Rust

According to the above analysis results, the DVI and the average temperature of meteorological factors were selected as the input variable  $x$ , and the disease index of wheat was selected as the output variable  $y$ . The ANFIS model was established to predict wheat stripe rust. The bell-shaped function “gbellmf” was

selected as the input variable of the model. The number of functions was 5, the type of the output membership function was linear, the error tolerance of the optimization method was 0.01, and the number of training epochs was 20.

The 30 ground points in the study area in the flag leaf stage and the early filling stage were selected as training data, and 15 ground points in the middle filling stage were selected as verification data. See Figs. 8 and 9 for the prediction results and validation results of the model, respectively.

Please put Fig. 8, Figure.9 about here

A fitting curve was drawn between the predicted value of the disease index of the model and the measured value of the ground disease index, and the results are shown in Fig. 10.

Please put Fig. 10 about here

As shown in Figs. 8, 9 and 10, the prediction results of the model were mostly consistent with the verification results; the  $R^2$  value of the model prediction was 0.998 and the RMSE was 0.0012, thereby indicating that the prediction model of wheat stripe rust constructed by the vegetation index and combined with meteorological elements had a high prediction accuracy.

## 4 Discussion

Based on the hyperspectral data and meteorological data, this study analyses the spectral characteristics of different disease indexes of wheat stripe rust and the meteorological factors that affect its occurrence. It also constructs an inversion model of disease indexes in different growth stages and the whole growth stage, and uses the adaptive network fuzzy reasoning method to predict the possibility of occurrence of wheat stripe rust.

The spectral characteristics and sensitivity of infected wheat showed that the differences in the spectral reflectance of the wheat canopy changed gradually with the development of the disease. In the flag leaf stage, the reflectance of wheat in the disease area in the visible band was higher than that in the control area, and the spectral sensitivity was positive. In the near red band, the reflectance of wheat in the disease area was lower than that in the control area, and the spectral sensitivity was negative. In the early filling stage, the reflectance of wheat in the disease area was higher than that in the control area, the spectral sensitivity was positive, and the difference between them was the largest near the red light. In the middle filling stage, the reflectance of wheat in the visible light in the band range and the reflectance difference between the disease area and the control area decreased gradually; the spectral sensitivity increased from a negative value to positive value. In the red band control area, the red edge amplitude of wheat was larger than the disease area, the spectral sensitivity changed to a negative value, and the difference between the two increased gradually.

An analysis of the absorption characteristics of infected wheat shows that in the early stage of the disease, the two absorption valleys in the visible light area were not obvious, and in the near-infrared area, the reflection peak value was reduced; with the degree of the disease increasing, the absorption peak area and absorption depth gradually reduced, the absorption peak area and absorption depth of wheat in the control area were slightly lower than in the disease area, and the degree of symmetry was slightly higher than in the disease area.

The correlation between the spectral sensitivity of different growth stages and the disease index was analysed. The results showed that the correlation coefficient was the largest at 987, 1362, and 1386 nm. The inversion model of the disease index of stripe rust in different growth stages was constructed by using the spectral sensitivity of this band. The  $R^2$  value of the model was 0.94. At the same time, a second inversion model of disease index of the whole growth stage was constructed based on the absorption characteristics. The  $R^2$  value of the model was 0.92. The results showed that the models had a high precision and strong condition inversion.

Using the grey correlation degree analysis method, the correlation degrees of the vegetation index and meteorological factors with stripe rust disease were analysed. The results showed that the vegetation index with the highest correlation degree was DVI, and the meteorological factor with the highest correlation degree was average temperature.

Based on the DVI and average temperature, a prediction model of wheat stripe rust was established, and the accuracy of the model was verified. The results showed that  $R^2$  value of the prediction model reached 0.998, and the prediction accuracy was high.

The meteorological data were gathered from the China meteorological data website; these data are widespread and cannot fully satisfy smaller spatial scales. Thus, describing the transmission process of stripe rust pathogens at the field scale is difficult. Determining how to accurately describe the impact of farmland micro meteorological environments on the occurrence of stripe rust pathogens requires further study.

## 5 Conclusion

Based on wheat canopy hyperspectral data, spectral sensitivity values were calculated and absorption characteristic parameters obtained through the continuous removal of the normalized spectral curve. By analysing the correlation between these parameters and the disease index, the disease inversion models of wheat stripe rust were constructed in the flag leaf, early and middle filling, and whole growth stages and the accuracy of the models was verified. The results showed that the accuracy of the disease inversion model was high and the inversion effect was strong. The daily air volume, average temperature, relative humidity, sunshine hours, precipitation, and other meteorological and disease investigation data were used for interpolation processing. By analysing the spectral characteristics and meteorological factors of wheat stripe rust using the grey correlation degree method, the correlation degrees of NDVI,

RDVI, TVI, DVI, RVI, and meteorological factors with the disease were analysed, and the difference vegetation index and average temperature with high correlation degrees were screened. The prediction model of wheat stripe rust was established; the precision of the model was very high, indicating the coupling of spectral data studied herein. The prediction method of wheat stripe rust with meteorological data is therefore feasible.

## Declarations

## Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Funding

This study was supported by the Natural Science Foundation of China (41871333), the Important Project of Science and Technology of the Henan Province (212102110238) and Scientific and Technological Innovation Team of Universities in Henan Province(22IRTSTHN008).

## Acknowledgments

Thanks to Haikuan Feng for the image data and field sampling collection. Thanks to Beijing Agricultural Information Technology Research Center for providing data. We are grateful to the anonymous reviewers for their valuable comments and recommendations. *We would like to thank Editage ([www.editage.cn](http://www.editage.cn)) for English language editing.*

## References

1. Chen B., Li, S. K., Wang, K. R., et al. (2011). Estimating Severity Level of Cotton Disease Based on Spectral Indices of TM Image. *J. Infrared Millimeter Waves*. 5, 451–457.
2. Elsayed, S., Rischbeck, P., and Schmidhalter, U. (2011). Comparing the performance of active and passive reflectance sensors to assess the normalized relative canopy temperature and grain yield of drought-stressed barley cultivars. *Field Crops Research*. 177, 148–160. doi:10.1016/j.fcr.2015.03.010
3. Gao, L. W., Shen, Z. R., Xia, B., and Jiang, Y. Y. (2009). Research vision on agricultural disease and insect pests monitoring and forecast information techniques chain. *China Plant Protection*. 11, 32–35.
4. Guo, J. B., Huang, C., Wang, H. G., et al. (2009). Disease Index Inversion of Wheat Stripe Rust on Different Wheat Varieties with Hyperspectral Remote Sensing. *Spectroscopy and Spectral Analysis*. 12, 3353–3357.

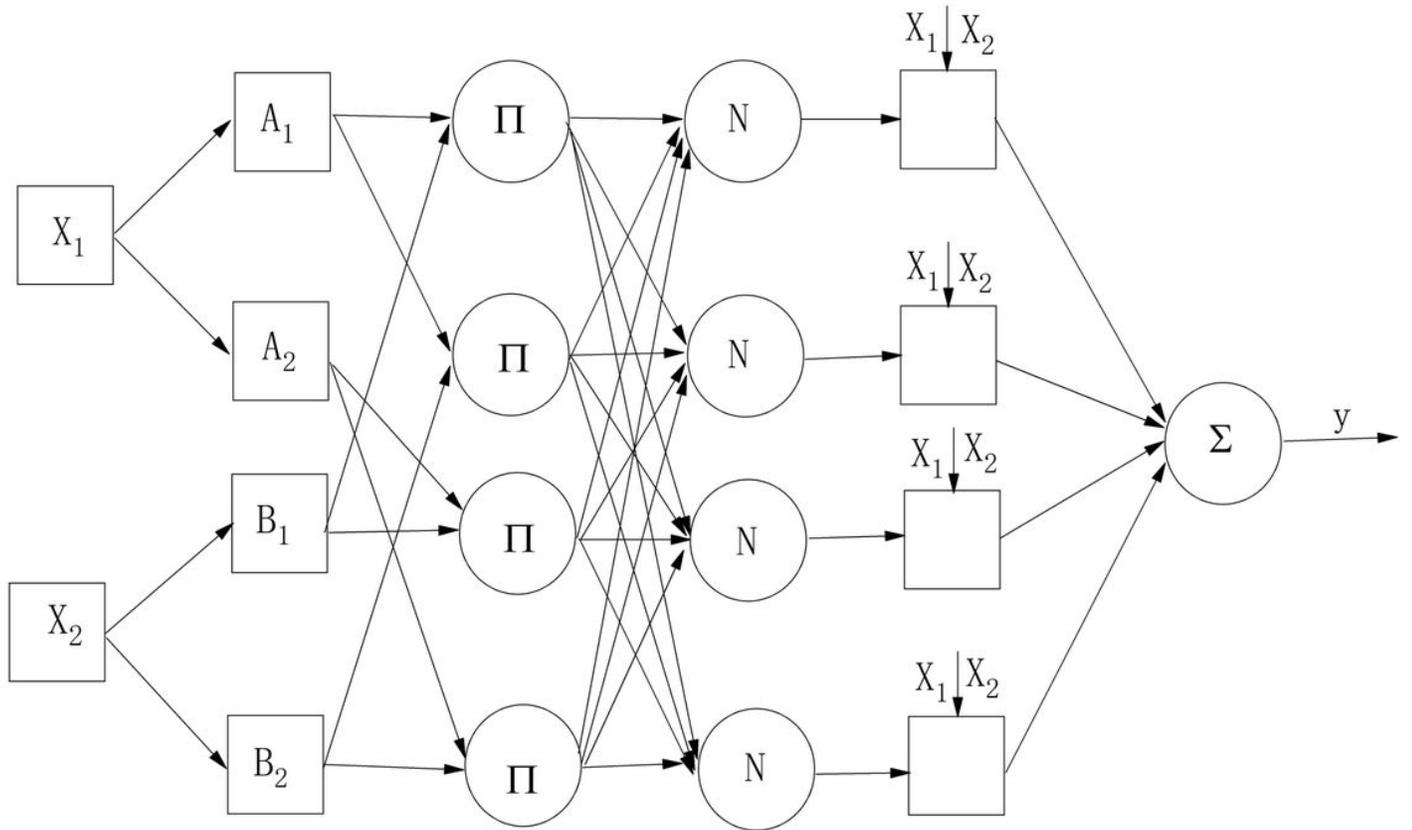
5. Hu, X. P., Cao, S. Q., Cornelius, A., and Xu, X. M. (2019). Predicting Overwintering of Wheat Stripe Rust in Central and Northwestern China. *Plant Disease*. 15, 44–51. doi: 10.1094 / PDIS – 06-19-1148-RE
6. Hu, X. P., Yang, Z. W., Li, Z. Q., et al. (2000). Prediction of Wheat Stripe Rust in Hanzhong Area by BP Neural Network. *Acta Agriculturae Boreali-occidentalis Sinica* .3, 28–31.
7. Huang, L. S., Jiang, J., Huang, W. J., et al. (2019). Wheat Yellow Rust Monitoring Method Based on Sentinel-2 Image and BPNN Model. *Transactions of the Chinese Society of Agricultural Engineering*. 17, 178–185.
8. Jing, X., Bai, Z. F., Gao, Y., et al. (2019). Wheat Stripe Rust Monitoring by Random Forest Algorithm Combined with SIF and Reflectance Spectrum. *Trans. Chin. Soc. Agric. Eng.* 13, 154–161.
9. Jiang, H. W. and Xia, J. L. (2003). Partial Least Square and Its Application. *Journal of the Fourth Military Medical University* .3, 280–283.
10. Jiang, J. B., Chen, Y. H., and Huang, W. J. (2010). Using the Distance between Hyperspectral Red Edge Position and Yellow Edge Position to Identify Wheat Yellow Rust Disease. *Spectroscopy and Spectral Analysis*. 6, 1614–1618.
11. Jiang, J. B., Chen, Y. H., and Huang, W. J. (2010). Using Hyperspectral Remote Sensing to Estimate Canopy Chlorophyll Density of Wheat under Yellow Rust Stress. *Spectroscopy and Spectral Analysis*. 8, 2243–2247.
12. Kelly, R. W., Phinn, S. R., Strong, W. M., et al. (2006). Predicting Grain Protein Content in Wheat Using Hyperspectral Sensing of In-Season Crop Canopies and Partial Least Square Regression. *Int. J. Geoinf.* 1, 93–108.
13. Klein, H., Person, T., and Itil, T. (1972). Family and environmental variables as predictors of social outcome in chronic schizophrenia. *Comprehensive Psychiatry* .4, 317–334. doi:10.1016/0010-440X(72)90002-8
14. Kobayashi, T., Kanda E., Kitada, K., et al. (2001). Detection of Rice Panicle Blast with Multispectral Radiometer and the Potential of Using Airborne Multispectral Scanners. *Phytopathology* .3, 316–323. doi: 10.1094/PHYTO.2001.91.3.316
15. Lu, J. J., Sun, L. G., and Huang, W. J. (2019). Research Progress in Monitoring and Forecasting of Crop Diseases and Pests by Remote Sensing. *Remote Sensing Technology and Application*. 1, 21–32.
16. Luo, J. H., Huang, W. J., Gu, X. H., et al. (2010). Monitoring Stripe Rust of Winter Wheat Using PHI Based on Sensitive Bands. *Spectroscopy and Spectral Analysis*. 1, 184–187.
17. Moshou, D., Bravo, C., West, J., et al. (2004). Automatic Detection of 'Yellow Rust' in Wheat Using Reflectance Measurements and Neural Networks. *Comput. Electron. Agric.* 44, 173–188.
18. Moshou, D., Bravo, C., R. Oberti, et al. (2005). Plant Disease Detection Based on Data Fusion of Hyperspectral and Multi-Spectral Fluorescence Imaging Using Kohonen Maps. *Real-Time Imaging*. 11, 75–83.
19. Qian, S., Huo, Z. G., and Ye, C. L. (2005). Long-Term Meteorological Prediction Research on Epidemic of Wheat Powdery Mildew in China. *Journal of Natural Disasters*. 4, 56–63.

20. Mistele, B. and Schmidhalter, U. (2010). Tractor-based quadrilateral spectral reflectance measurements to detect biomass and total aerial nitrogen in winter wheat. *Agron. J.* 2, 499–506. doi:10.2134/agronj2009.0282
21. Wang, H. G., Ma, Z. H., Wang, T., et al. (2007). Application of Hyperspectral Data to the Classification and Identification of Severity of Wheat Stripe Rust. *Spectroscopy and Spectral Analysis.* 9:1811–1814.
22. Wang, Z. W., Zhang, D. X., Ma, Y. L., and Ban, S. L. (2010). Study on Meteorological Degree Forecast Model of the Main Pests and Diseases of Winter Wheat in Shanxi. *Chin. Agric. Sci. Bull.* 11, 267–271.
23. Xu, L., Chang, G. B., Wang, Y. J., et al. (2018). Application of Akaike Information Criterion in Selecting Polynomial Models for Height Transformation. *Journal of Hefei University of Technology (Natural Science)*. 8, 1074–1078 doi:10.3969/j.issn.1003-5060.2018.08.013
24. Xu, Y. P., Wan, X., Yao, X. H., et al. (2018). Overwintering Climate Suitability of Wheat Stripe Rust in Hedong of Gansu. *Chin. Agric. Sci. Bull.* 12, 142–145.
25. Xue, L., Cao, W., Luo, W., et al. (2004). Monitoring Leaf Nitrogen Status in Rice with Canopy Spectral Reflectance. *Agron. J.* 1, 135–142.
26. Yu, X. B., Chen, H., Ji, Z. H., et al. (2019). Gray Correlation Analysis of Agricultural Meteorological Disasters and Main Crop Yields in Shandong Province. *J. Catastrophol.* 3, 103–108. doi:10.3969/j.issn.1000-811X.2019.03.020
27. Yuan, L., Zhang, J. C., Zhao, J. L., et al. (2013). Differentiation of Yellow Rust and Powdery Mildew in Winter Wheat and Retrieving of Disease Severity Based on Leaf Level Spectral Analysis. *Spectroscopy and Spectral Analysis.* 6, 1608–1614.
28. Zhan, Y., Jing, J., Li, W., et al. (2010). Discrimination of Rice Panicles by Hyperspectral Reflectance Data Based on Principal Component Analysis and Support Vector Classification. *Journal of Zhejiang University-Science B.* 1, 71.
29. Zhang, X. X., Wang, B., Tian, Y. Y., et al. (2019). Research Progress on Forecasting Mechanism and Methodology for Crop Disease and Insect Pest. *J. Agric. Sci. Technol.* 5, 110–120. doi:10.13304/j.nykjdb.2018.0407
30. Zhang, Y. X., Guo, J. Q., Qian, X. Y., et al. (2019). An Adaptive Network-based Fuzzy Inference System with Mixed Data Inputs. *Acta Automatica Sinica.* 9, 1743–1755. doi:10.16383/j.aas.2018.c170698
31. Zheng, Y., Chang, Q. R., Wang, T. T., and Yang, J. (2019). Hyperspectral Estimation of SPAD Value in Oilseed Rape Based on Continuum Removal and Partial Least Squares Regression. *Journal of Northwest A&F University (Natural Science Edition)*. 8, 37–45. doi:10.13207/j.cnki.jnwafu.2019.08.006

## Figures

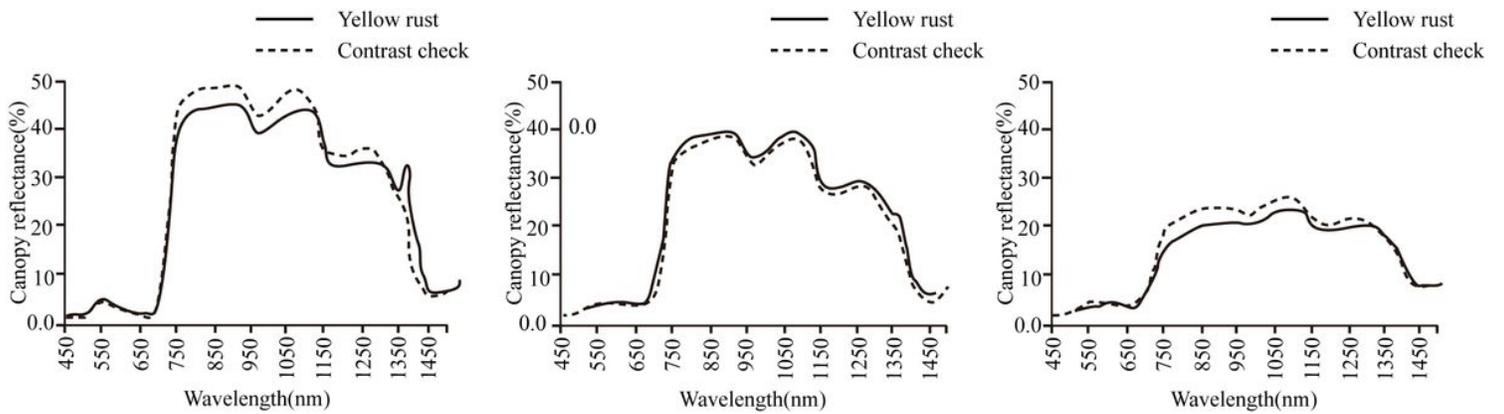
**Figure 1**

Overview of the study area



**Figure 2**

Structure of ANFIS



**Figure 3**

Spectral curve of different growth stage in control area and disease area. (a) flag leaf stage, (b) early filling stage and (c) middle filling stage.

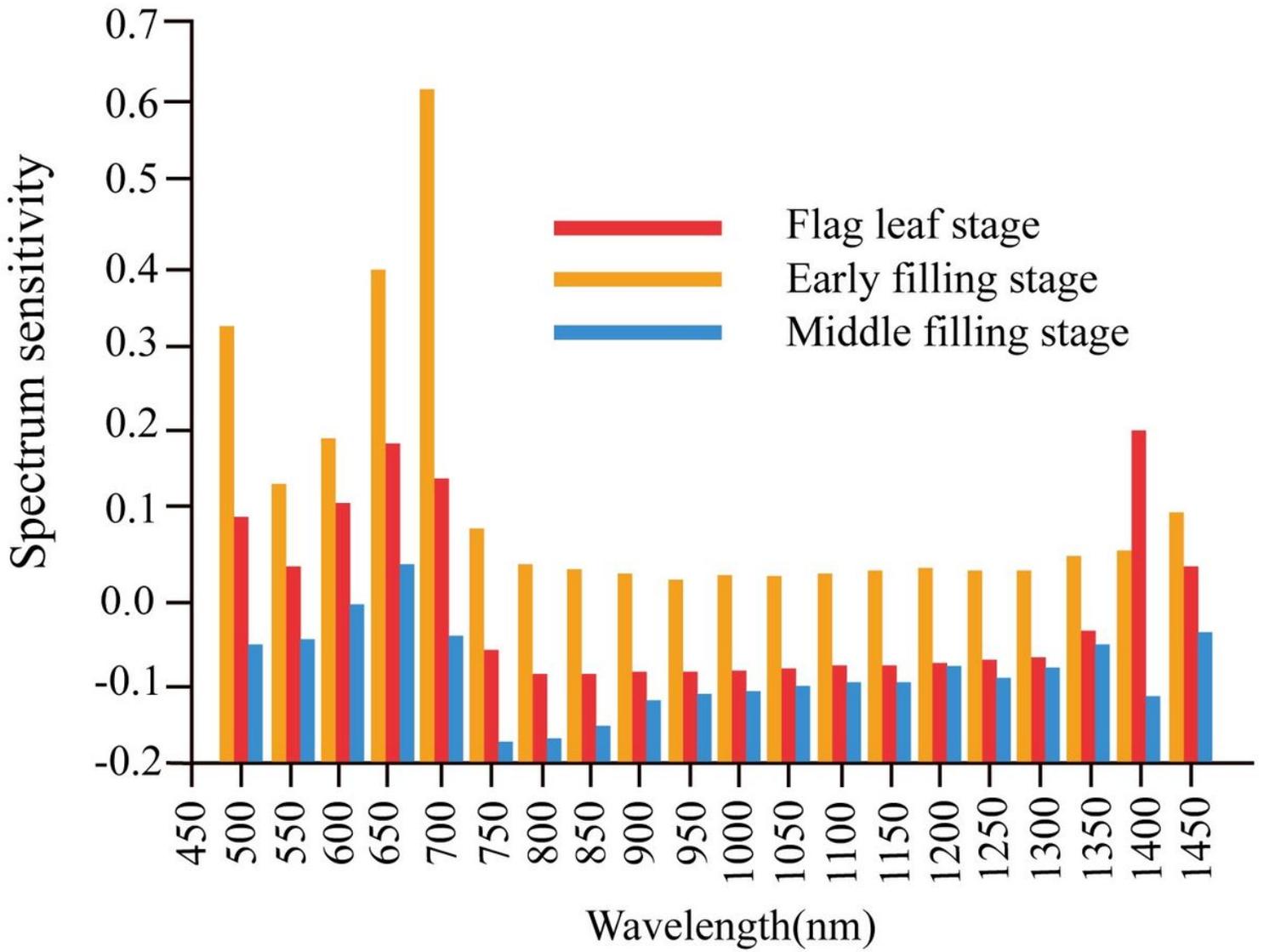
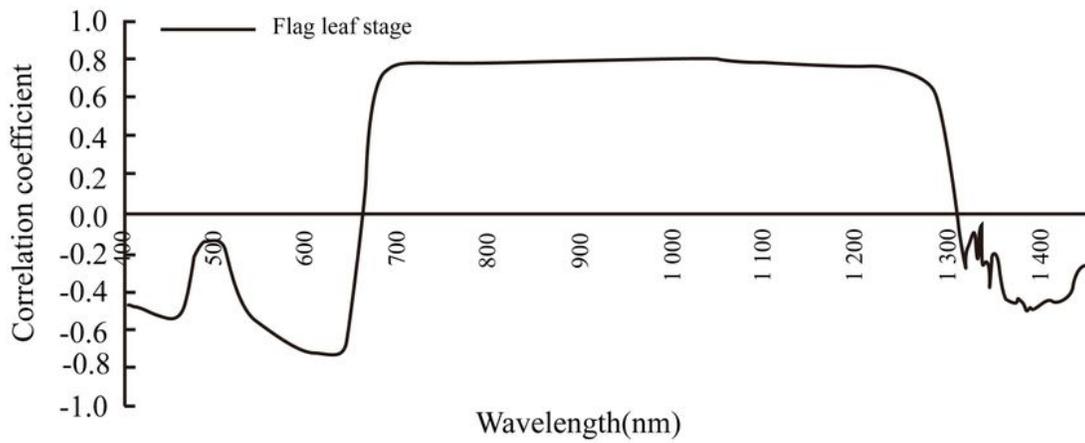
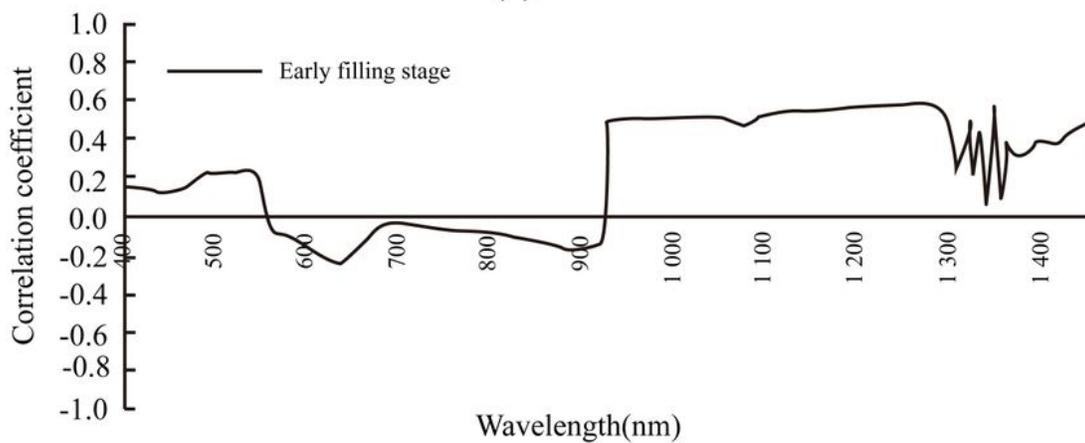


Figure 4

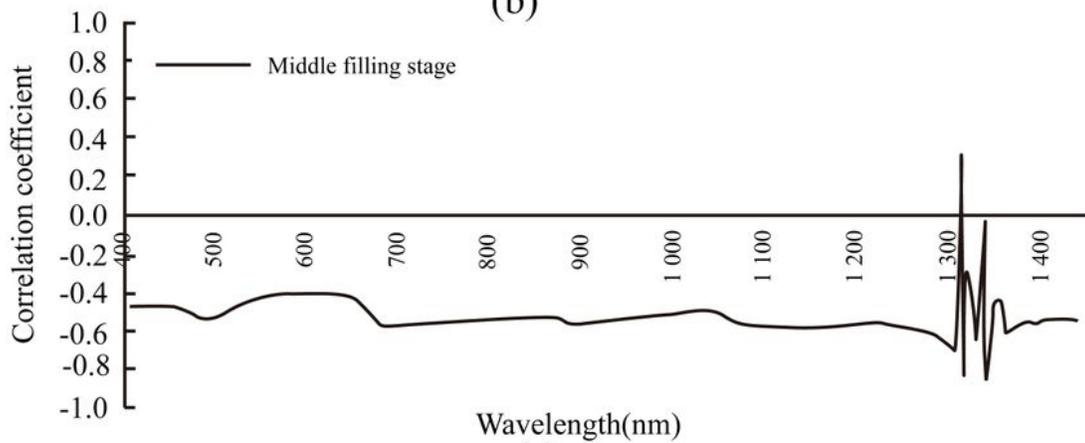
Spectral sensitivity histogram of flag leaf stage, early filling stage and middle filling stage



(a)



(b)



(c)

**Figure 5**

Relationship between spectral sensitivity and disease index in different growth stage.(a) flag leaf stage, (b)early filling stage and (c) middle filling stage.

**Figure 6**

(a)-(c)Area interpolation result and (d)-(f)interpolation result of disease level in 2010-2012.

### **Figure 7**

Interpolation result of meteorological element in 2010-2012. (a)-(c) interpolation result of wind volume value, (d)-(f)interpolation result of average temperature, (g)-(i) interpolation result of relative humidity, (j)-(l) interpolation result of sunshine hour,(m)-(o)interpolation result of precipitation

### **Figure 8**

Model prediction result

### **Figure 9**

Model validation result

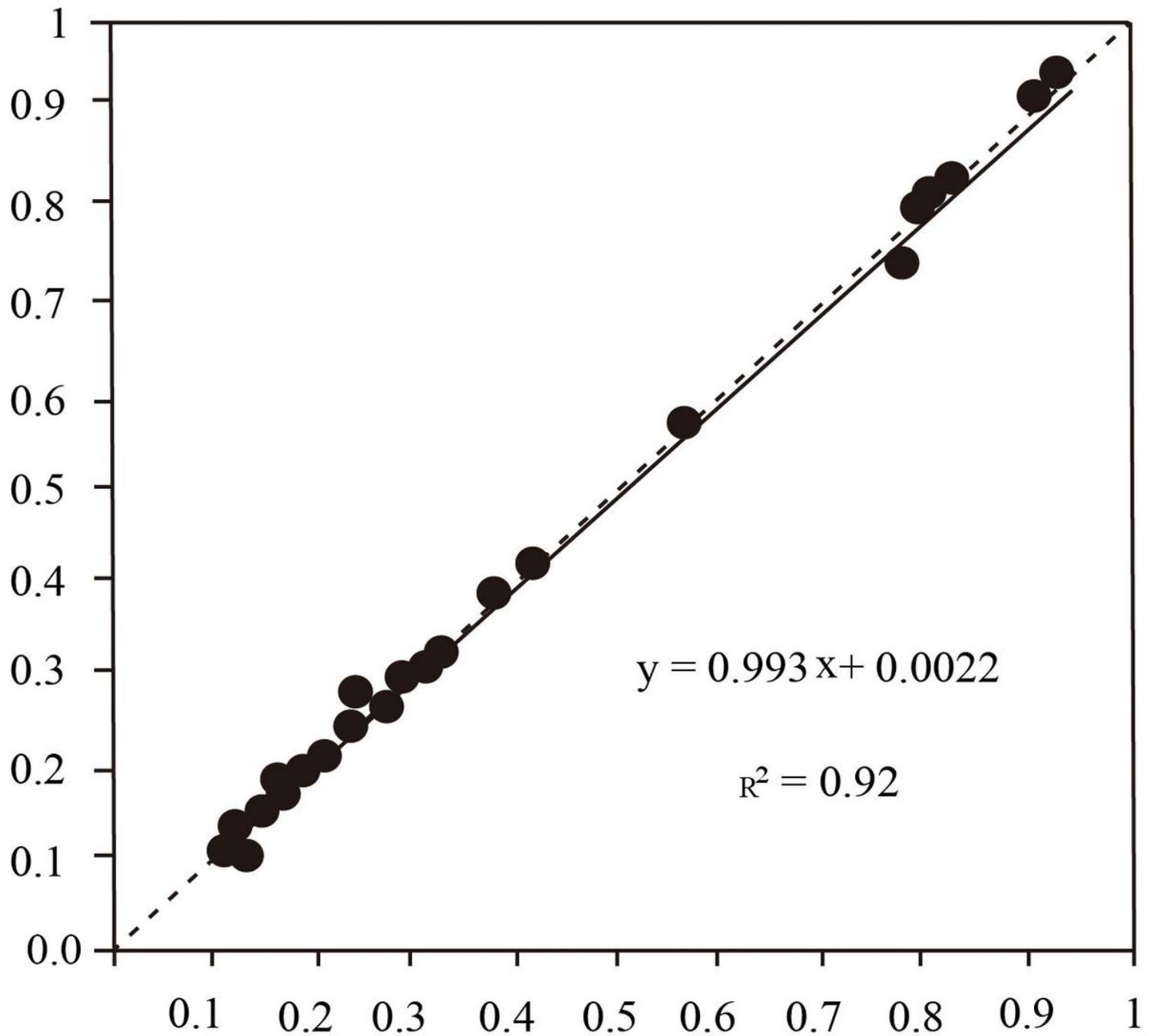


Figure 10

Fitting curve between predicted value and measured value of disease index