Analysis of Rice False Smut Severity Levels in Donggang Using Vegetation Index Normalization

Hengqian Zhao(1)(2)*, Zelong Liu(1), Yifeng Yang(1)

[1] College of Geoscience and Surveying Engineering, China University of Mining and Technology, Beijing 100083, China; zhaohq@cumtb.edu.cn

[2] State Key Laboratory of Coal Resources and Safe Mining, China University of Mining and Technology, Beijing 100083, China

Email: zhaohq@cumtb.edu.cn

Abstract: Timely monitoring of crop disease development is very important for precision agriculture application. Remote sensing-based vegetation indices(VIs) can be good indicators of crop disease severity, but the result of inversion could be affected by the growing environment of crops and has strong uncertainly. Based on VI normalization, crop disease could be identified and has an inherent correlation, and eliminate the influence of variable field condition and the VI range variations. This technique was applied to rice fields infested with false smut in Donggang, Liaoning. Five typical VIs and a custom VI were calculated from the hyperspectral data collected by ground spectral survey using McPin. Six Vegetation indexes were normalized and the regression analysis of normalized result and ground survey result was carried out. The statistical analysis of the results showed that the mean value of normalized VIs had a good rule with the disease grade, and the variance of the disease level with a large number of samples was also very small, most of them were less than 0.1 which can provide reference for further study of automatic classification of disease. According to the analysis of the obtain result, the normalized Vegetation Indexes value has a inherent correlations with disease severity level acquired by ground survey, and with the development and application of more VIs, there is still a broader room for the development of this method.

Keywords： Rice；False Smut；Hyperspectral remote sensing；Vegetation index normalization; Regression Analysis

1 INTRODUCTION

Crop disease infection is one of the main variable threatening agriculture production and sustainability, so monitoring crop health timely is a very important part of effective agricultural field management(Zhan-yu LIU,2008). For example, rice as one of the significant food crops, the security of rice yield is essential for the stability and sustainable development of national economy. False smut of rice is a common fungous disease. In recent years, due to several reasons, rice false smut is becoming more and more serious and showing the characteristics of a wide range of disease, high frequency of disease and serious yield loss. So it is significant to explore timely rice disease occurrence data over large areas so that preventive measures can be taken to improve rice yield and quality.

Remote sensing makes it conceivable to screen crop disease quickly on a large scale, which has the advantages of being timely, easy to use, extensive, nondestructive, and objective(Jin, Huang, Ren, Luo, Wu, Jing and Wang,2013). Remote sensing has been utilized to recognize, screen and evaluate an assortment of diseases in various crops. Comprehensive reviews on the use of remote sensing for the detection of plant diseases are available(Calderón, Navas-Cortés, Lucena and Zarco-Tejada,2013). Among different types of remote sensing techniques, hyperspectral remote sensing is one of the most effective approaches to discern features that are difficult to detect in the spectrum, due to its high spectral resolution (Zhang, yuan, Pu, Loraamm, Yang and Wang,2014). Hyperspectral technology is broadly and effectively applied to monitor various types of stresses on crops(Duveiller, Weiss, Baret and Defourny,2010, Zhang, Pu, Wang, Huang, Yuan and Luo,2012).

Traditional way of monitoring of rice false smut mainly relies on artificial self-testing, which is time-consuming and laborious, so it is impossible to accurately estimate the regional distribution and severity of large area of disease. In addition, the workers may have different ability to accurately detect plant disease, and with the influence of fatigue, inexperience, and bias among workers the subjectivity of visual assessment can be increased.

Based on remote sensing data, many forms of VIs have been used for crop disease detection, and empirical statistical models between VIs and disease severity levels can be established using discriminant analysis, linear regression analysis, support vector machine(SVM), and other statistical methods (Prabhakar, Prasad, Thirupathi, Sreedevi, Dharajothi and Venkateswarlu,2011, Zhang, yuan, Pu, Loraamm, Yang and Wang,2014). Vegetation indices (VIs) enable the evaluation and observation of changes in canopy biophysical properties, such as leaf area index (LAI), chlorophyll content, and photosynthetically active radiation (PAR) (Ahamed, Tian, Zhang and Ting,2011). It has been revealed that crop disease can severely influence the biophysical property values, which can explain why the severity level is significantly correlated to crop VI values(Yang, Greenberg, Everitt and Fernandez,2012, Zhang, Sun, Wu and Zhang,2015).

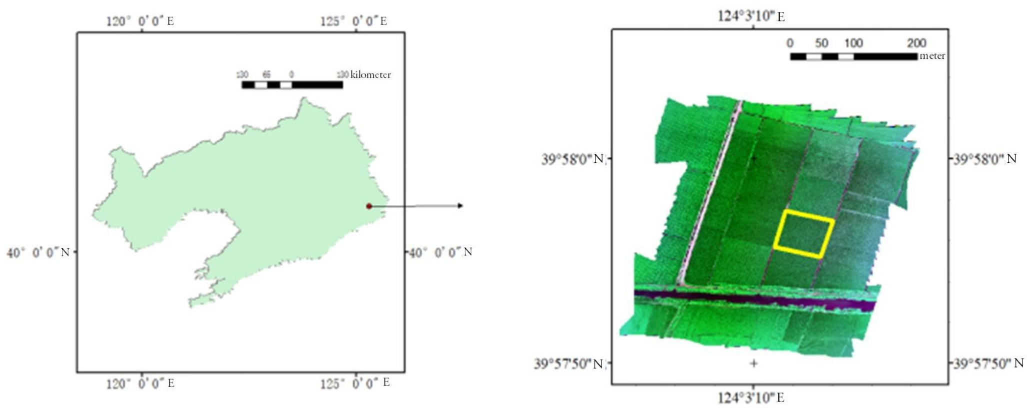
Since there is a significant correlation between vegetation index and crop disease. We can process regression analysis between crop disease level and VIs value. In addition,through the value ranges of VIs and the crop growth conditions can vary greatly from field to field, and this can influence the correlation between vegetation index and crop disease. We need to process normalization, as a simple and effective data-processing method, to eliminate the background differences for reflectance spectra and eliminate the field management difference and VI value ranges difference between different VIs(Oumar, Mutanga and Ismail,2013, Zhao, Yang, Guo, Zhang and Zhang,2020).

In this study, based on VI normalization, we proposed a regression analysis between different disease level and VIs which has been normalized. Using hyperspectral data around several ground check point with varying levels of rice false smut. Six typical VIs and a typical regression method was used to conduct disease severity analysis, and the descriptive statistics were proposed to support other studies such as automatic disease classification.

2 EXPERIMENTAL MATERIALS AND METHODS

2.1 Study Area

The study area was located in a rice growing district in Donggang City, Liaoning Province, China (Figure 1a). There are several fields in the district. And Field A is selected as the experimental area (Figure 1b, false color image, with the band composition of red, red edge and green). Field A is 3263 square meters in area.

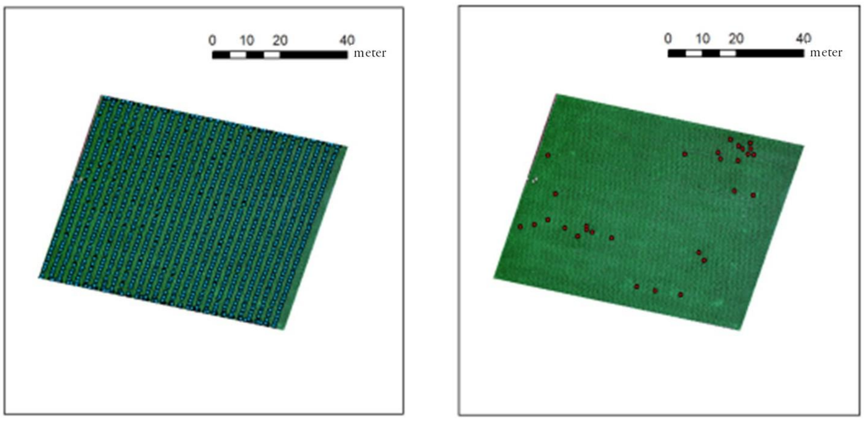


(a) **(b)**

Figure 1. Location of the study area:(a)Liaoning;(b)experiment fields(Field A in yellow).

2.2 Spectral and Ground Survey Data Acquisition

A McPin is installed on a UAV(Unmanned Aerial Vehicle) which flies near the ground to obtain ground hyperspectral data. The McPin is a hyperspectral remote sensor mounted on a UAV that can acquire imagery with a spectral resolution of 1nm and cover the spectral range of visible and near-infrared. In this experiment, the data was collected on August 22, 2019 with the flight height of 4m at the speed of 2m/s, and the field ridge was 2m in width. The wavelength range of the acquired hyperspectral data is 340-820nm and the distribution is shown in the figure (Figure 2a). At the same time, a manual ground survey was conducted on August 22 to check the occurrence of false smut in the rice field. The distribution of ground survey points is shown in the figure (Figure 2b).



(a) (b)

Figure 2.Data distribution:(a)Hyperspectral Data;(b)Ground survey Data.

2.3 Data Processing

2.3.1 Spectral Data Preprocessing

A buffer area with a diameter of 1.6m around the ground survey point is set. The average of the hyperspectral data in the buffer (Figure 3) is calculated as the hyperspectral data of the ground survey point (The overlapping points in the buffer area have similar disease severity level and will not affect the spectra of each level).

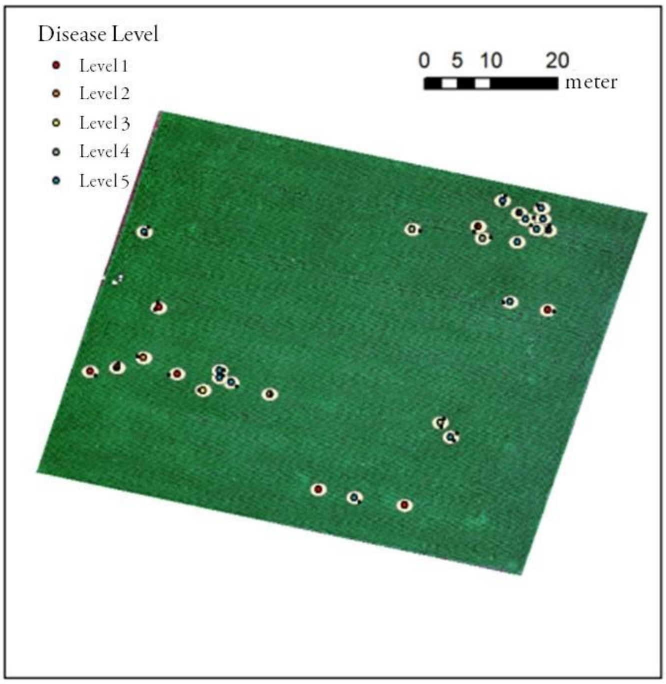


Figure 3. Spectral data point around ground survey point.

Collect the statistics of the spectra of each ground survey point, and calculate the average spectra for different disease severity level. And then compare the average spectra with the spectra of healthy plants (Figure 4). The result shows that the spectra of rice plant infected with false smut changes at the different stage of the disease development.

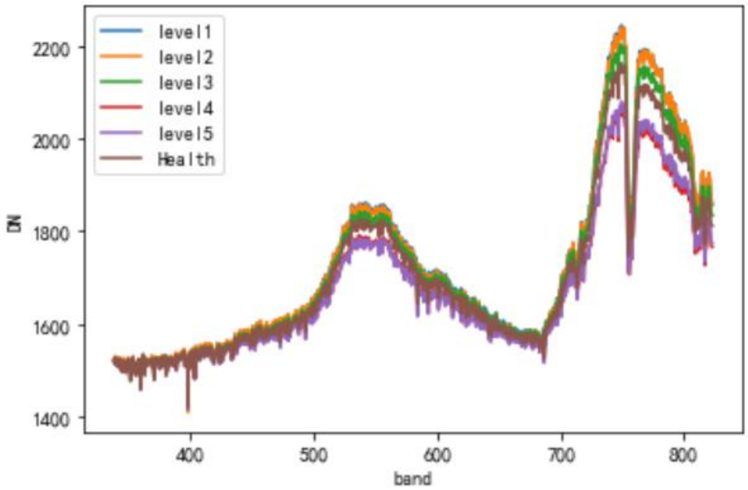


Figure 4. Mean spectral of different disease level.

2.3.2 Spectral Sensitivity Analysis

The spectral sensitivity is the ratio of the interpolation of the spectra of the infected plant leaf and the spectra of the healthy plant leaf to that of the healthy plant leaf. The band where the sensitivity peak is located can be used as a sensitive band which has a relatively high inversion accuracy. The sensitivity analysis result of the spectrum is shown in the figure (Figure 5).

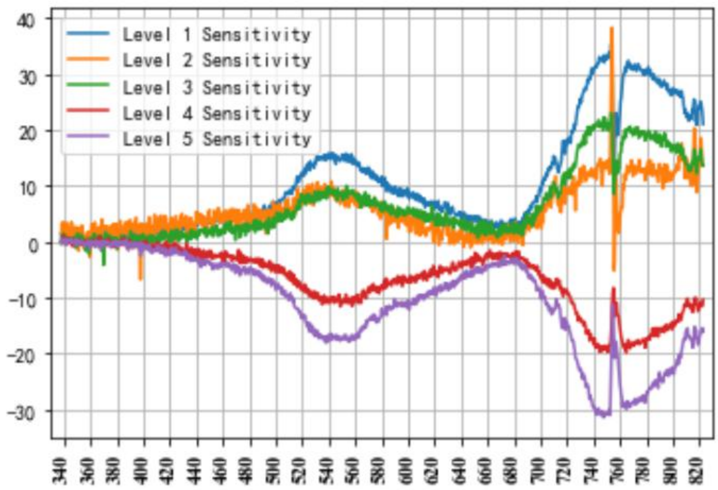


Figure 5 Spectral sensitivity of different disease level

It can be seen from the figure that when the disease severity level is generally low at the early stage, the sensitivity is positive and the value at level 1> level 2> level 3, indicating that the reflectance from the crops just infected with false smut, especially in the green and near infrared bands, is higher than that of healthy ones. But the reflectance will gradually decrease as the disease develops. When the disease is severe, the sensitivity is negative and the sensitivity at level 4>level 5 in the green band, level 5>level 4 in the near-infrared band, which proves that when the disease turns severe, the reflectance of the crop, especially in green and near-infrared band, will be lower than that of healthy plants. The 540nm and 740nm bands where the absolute peak value of sensitivity is located are selected as the sensitive bands for infected plants to build the 740 and band 540 band index, which are used for the following analysis.

2.3.3 Normalization of Vegetation Indices

Calculate the following five typical vegetation indices (Table 1) and the index of 740 and 540 bands mentioned above by using the average spectra of each ground survey point.

Table 1. Vegetation indices applied in this study.

Vegetation Index	Abbrev.	Formula	Application
Normalized difference vegetation index	NDVI	(R800 − R670)/(R800 + R670)	Biomass
Green normalized difference vegetation index	GNDVI	(R800 − R550)/(R800 + R550)	PAR, vegetation cover
Enhanced vegetation index	EVI	2.5(R800 − R670)/(R800 + 6R670 − 7.5R450 + 1)	Biomass, LAI
Soil adjusted vegetation index	SAVI	(1 + L)(R800 − R670)/(R800 + R670 + L)	Biomass, soil background
Red edge 2	RedEdge	(R710 − R680)/(R710 + R680)	Chlorophyll content

By performing normalization, the value range of different vegetation indices can be changed to a range between 0 and 1. In this equation, X is the original data, Xmax is the maximum and Xmin is the minimum, and then the normalized data Xnor can be derived as shown:

$$X_{nor} = 1 - \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{1}$$

3 DATA ANALYSIS AND RESULIT

3.1 Data Analysis

3.1.1 Disease Severity Level-Normalized Vegetation Indices Maps

A scatter diagram (Figure 6) is made with the disease severity level obtained from the ground survey as the X axis, and the normalized vegetation indices as the y axis.

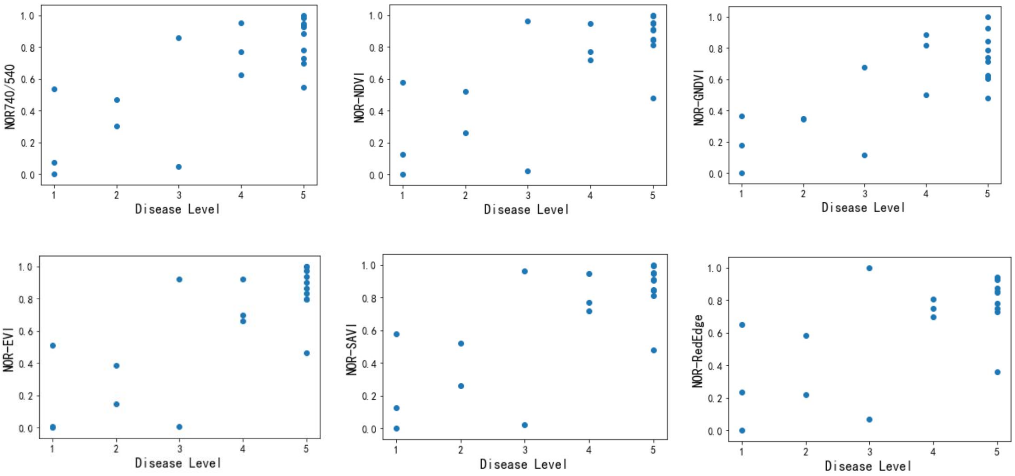


Figure 6. Disease Level-NOR VIs scatter

It can be seen from the figure that the normalized 740/540 index, NDVI, GNDVI, SAVI,EVI and RedEdge index all show a certain positive correlation with the disease severity level, namely vegetation index is on the rise after normalization as the disease severity level increases. And the correlation coefficient between them is shown in the following table (Table 2), which also verifies the discussion above.

Table 2. Correlation between disease level and NOR VIs

	NOR740/540	NOR-NDVI	NOR-GNDVI	NOR-EVI	NOR-SAVI	NOR-RedEdge
Disease Level	0.7754	0.7520	0.7741	0.7864	0.7520	0.6577

3.1.2 Linear Regression Analysis

Linear regression analysis is performed on the EVI with the highest correlation, which verifies the conclusions of previous study, and the 740 and 540 index with the second highest correlation. The classic least squares method is used first. The data of 18 of the 20 ground survey points is used as the training data and data of 2 points as test data for regression. The regression equation is shown in Figure 7.

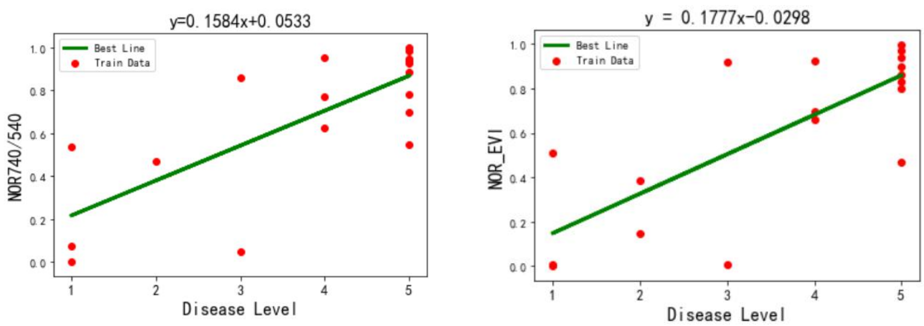


Figure 7. Linear regression result

Due to the small amount of data, input all the data in the training and test set when verifying the validity of the regression equation to perform cross validation for five times (Table 3) and calculate the mean square error(MSE)(Equation 2). In the equation, y_i is the predicted result of the regression equation, $\hat{y}$ is the corresponding true value, and m is the number of samples.

$$MSE = \frac{\sum_{1}^{m}(y_i - \hat{y}_i)^2}{\mathrm{m}} \tag{2}$$

Table 3. MSE of the linear regression model

Cross-validation	1	2	3	4	5	Mean
740/540	0.0293	0.0293	0.1157	0.0134	0.0626	0.0501
EVI	0.0487	0.0029	0.1185	0.0466	0.1005	0.0635

3.1.3 Statistical Analysis

Statistics of the mean and variance of the normalized vegetation indices is obtained as shown in the Table 4 and Table 5 respectively.

Table 4. Mean value of NOR VIs between different disease level

Disease level	Mean NOR740/540	Mean nor-NDVI	Mean NOR-GNDVI	Mean NOR-EVI	Mean NOR-SAVI	Mean NOR-RedEdge
1	0.2045	0.2342	0.1804	0.1715	0.2342	0.2950
2	0.3873	0.3908	0.3492	0.2667	0.3908	0.4017
3	0.4537	0.4920	0.3951	0.4625	0.4920	0.5350
4	0.7818	0.8113	0.7344	0.7600	0.8113	0.7502
5	0.8438	0.8688	0.7327	0.8561	0.8688	0.8004

Table 5. Variance value of NOR VIs between different disease level

Disease level	Variance NOR740/540	Variance nor-NDVI	Variance NOR-GNDVI	Variance NOR-EVI	Variance NOR-SAVI	Variance NOR-RedEdge
1	0.0849	0.0916	0.0332	0.0859	0.0916	0.1085
2	0.0140	0.0347	0.0000	0.0287	0.0347	0.0672
3	0.3250	0.4399	0.1563	0.4193	0.4399	0.4325
4	0.0269	0.0144	0.0431	0.0200	0.0144	0.0031
5	0.0218	0.0228	0.0255	0.0248	0.0228	0.0297

It can be seen that the mean value varies in a regular pattern as the disease severity level changes. And in the disease severity level with a relatively large number of samples, the variance is also small with much of it below 0.1. This can provide a reference for the subsequent automatic disease classification.

4 CONCLUSION

In this study, we normalized several typical VIs. Two significant vegetation indexes were used for regression analysis, and good accuracy was obtained. The comprehensive mean MSE was below 0.0635 after five cross-validation. There are various kinds of vegetation index and still in development, through exploration of more VIs in the future the vegetation index normalization method will have more extensive precision and generalization ability and multiple linear regression can be used to improve the accuracy. Furthermore, description statistic was performed on the normalized VIs. The results showed that the mean value had a good rule with the disease grade, and the variance of the disease grade with a large number of samples was also very small, most of which were less than 0.1 which can provide reference for further automatic classification of disease.

Acknowledgements:

This research received the support from the “Algorithm Developer Community of Digital Agriculture” of Mcfly Technology.

References

T. Ahamed, L. Tian, Y. Zhang and K. C. Ting, 2011. A review of remote sensing methods for biomass feedstock production. Biomass and Bioenergy,35.

R. Calderón, J. A. Navas-Cortés, C. Lucena and P. J. Zarco-Tejada, 2013. High-resolution airborne hyperspectral and thermal imagery for early detection of Verticillium wilt of olive using fluorescence, temperature and narrow-band spectral indices. Remote Sensing of Environment, 39.

G. Duveiller, M. Weiss, F. Baret and P. Defourny, 2010. Retrieving wheat Green Area Index during the growing season from optical time series measurements based on neural network radiative transfer inversion. Remote Sensing of Environment,115.

N. Jin, W. Huang, Y. Ren, J. Luo, Y. Wu, Y. Jing and D. Wang, 2013. Hyperspectral identification of cotton verticillium disease severity. Optik - International Journal for Light and Electron Optics, 124.

Z. Oumar, O. Mutanga and R. Ismail, 2013. Predicting Thaumastocoris peregrinus damage using narrow band normalized indices and hyperspectral indices using field spectra resampled to the Hyperion sensor. International Journal of Applied Earth Observations and Geoinformation, 21.

M. Prabhakar, Y. G. Prasad, M. Thirupathi, G. Sreedevi, B. Dharajothi and B. Venkateswarlu, 2011. Use of ground based hyperspectral remote sensing for detection of stress in cotton caused

by leafhopper (Hemiptera: Cicadellidae). Computers and Electronics in Agriculture, 79.
C. Yang, S. M. Greenberg, J. H. Everitt and C. J. Fernandez, 2012. Assessing defoliation, regrowth control and root rot infection using remote sensing technology. Inventi Impact Agro Tech, 2012.
J.-f. H. Zhan-yu LIU, Rong-xiang TAO, 2008. Characterizing and Estimating Fungal Disease Severity of Rice Brown Spot with Hyperspectral Reflectance Data. Rice Science, 232-242.
J.-C. Zhang, R.-l. Pu, J.-h. Wang, W.-j. Huang, L. Yuan and J.-h. Luo, 2012. Detecting powdery mildew of winter wheat using leaf level hyperspectral measurements. Computers and Electronics in Agriculture, 85.
J. Zhang, L. yuan, R. Pu, R. W. Loraamm, G. Yang and J. Wang, 2014. Comparison between wavelet spectral features and conventional spectral features in detecting yellow rust for winter wheat. Computers and Electronics in Agriculture, 100.
L. Zhang, X. Sun, T. Wu and H. Zhang, 2015. An Analysis of Shadow Effects on Spectral Vegetation Indexes Using a Ground-Based Imaging Spectrometer. IEEE Geoscience and Remote Sensing Letters, 12, 2188-2192.
H. Zhao, C. Yang, W. Guo, L. Zhang and D. Zhang, 2020. Automatic Estimation of Crop Disease Severity Levels Based on Vegetation Index Normalization. Remote Sensing, 12.