Integrating Spectral Information and Meteorological Data to Monitor Wheat Yellow Rust at a Regional Scale: A Case Study

Qiong Zheng 1,*, Huichun Ye 2,3,*, Wenjiang Huang 2, Yingying Dong 2, Hao Jiang 1,*, Chongyang Wang 1,*, Dan Li 1, Li Wang 1 and Shuisen Chen 1

Abstract: Wheat yellow rust has a severe impact on wheat production and threatens food security in China; as such, an effective monitoring method is necessary at the regional scale. We propose a model for yellow rust monitoring based on Sentinel-2 multispectral images and a series of two-stage vegetation indices and meteorological data. Sensitive spectral vegetation indices (single- and two-stage indices) and meteorological features for wheat yellow rust discrimination were selected using the random forest method. Wheat yellow rust monitoring models were established using three different classification methods: linear discriminant analysis (LDA), support vector machine (SVM), and artificial neural network (ANN). The results show that models based on two-stage indices (i.e., those calculated using images from two different days) significantly outperform single-stage index models (i.e., those calculated using an image from a single day), the overall accuracy improved from 63.2% to 78.9%. The classification accuracies of models combining a vegetation index with meteorological feature are higher than those of pure vegetation index models. Among them, the model based on two-stage vegetation indices and meteorological features performs best, with a classification accuracy exceeding 73.7%. The SVM algorithm performed best for wheat yellow rust monitoring among the three algorithms; its classification accuracy (84.2%) was ~10.5% and 5.3% greater than those of LDA and ANN, respectively. Combined with crop growth and environmental information, our model has great potential for monitoring wheat yellow rust at a regional scale. Future work will focus on regional-scale monitoring and forecasting of crop disease.

Keywords: wheat yellow rust; vegetation indices; meteorological information; food security; regional remote sensing

1. Introduction

Wheat is the main grain crop for mankind [1]. Yellow rust (Puccinia striiformis f. sp. tritici Erikss) is a devastating disease in wheat planting that affects wheat growth, thus seriously affecting the quality and yield of wheat in China [1,2]. The average annual area of wheat yellow rust is 4 million hm², resulting in a reduction in wheat production of more than 1 billion kg per year [3]. Traditional methods of wheat yellow rust involve manual surveys that are time-consuming, laborious, and inefficient [4]. In recent decades, remote sensing technology has been proved to be an effective tool for monitoring of crop disease and pest, with advantages of large-scale and real time simultaneous monitoring [4,5]. Therefore, timely, effective, and accurate monitoring of wheat yellow rust based on remote
sensing technology and multi-source data of disease occurrence is essential to ensure food security and agricultural sustainability in China.

It is well-known that the phenology of vegetation is closely related to the seasonality of meteorological variables [6]. In addition, the integration of meteorological data with satellite hydrological models can improve irrigation scheduling management and help farmers to properly make rational irrigation plans [7]. For pathogens, the propagation, spread, and infection of pathogen spores require suitable environmental support; thus, the monitoring of crop pests and diseases is related to environmental conditions [4,8]. Meteorological data, such as temperature, humidity, sunshine, and rainfall are the key factors to determining the occurrence, development, and prevalence of crop diseases [4]. Zhang et al. showed that integrating meteorological data (precipitation, temperature, humidity, and solar radiation) and remote sensing features has significant potential to forecast the occurrence probability of wheat powdery mildew [9]. Papastamati et al. stated that the inoculation concentrations, temperature, and rainfall time were related to light leaf spot epidemics in winter oilseed rape disease and proposed a new model for predicting the disease [10]. Habitat monitoring has great potential for evaluating the occurrence and distribution of plant diseases on a regional scale [8].

In addition, biophysical parameters of the host have a certain effect on the infection of diseases [11]. In different infestation stages, plants exhibit specific host–pathogen interactions, such as a reduction in leaf area index, pigment content changes, and canopy structure morphology destruction [12]. These changes respond in the visible to near-infrared regions of the spectrum and can even be captured through multi-temporal observations [11–13]. Zhang et al. used single- and multi-temporal images from HJ satellite to monitor wheat powdery mildew, and proved that multi-stage images were superior to single-temporal images for wheat disease monitoring, with monitoring accuracy reaching 78% [14]. Ma et al. developed a multi-temporal vegetation index that indicated crop growth status and habitat characteristics to monitor wheat powdery mildew based on the k-nearest neighbor approach [15]. Therefore, the integration of habitat information and the host growth status related to disease and pest occurrence have great potential for crop stress monitoring.

Classifiers can learn the characteristics of target classes from training samples and apply this information to unclassified data. Methods, such as decision tree (DT), k-nearest neighbors (kNN) method, support vector machine (SVM), and artificial neural network (ANN), have been developed and applied in objects recognition and species classification in the remote sensing field [16–18]. Among these, the SVM and ANN methods are popular for crop classification and disease remote sensing monitoring [4,16].

Satellites with different spatial resolutions have been used in agricultural disease identification and monitoring [4]. Razz et al. and Yue et al. effectively utilized the high-resolution of PlantScope imagery (3 m) to detect soybean sudden death syndrome and rice diseases, respectively [19,20]. Calderón et al. used high spatial resolution and hyperspectral imagery for Verticillium wilt of olive for early detection [21]. Yuan et al. used high spatial resolution (Worldview 2) and medium-spatial resolution (Landsat 8) satellite images to monitor the distribution of wheat disease and pest, achieving a monitoring accuracy of more than 71.0% [8]. Therefore, satellite imagery has proven to be an efficient tool for monitoring crop disease and pest in farms on large regional scales. The Sentinel-2 multispectral satellite launched in 2015, has a 290 km swath width and a 10 days revisit cycle, although it can reach up to 5 days if two satellites work simultaneously [22]. Sentinel-2 multispectral images have 13 spectral bands with a spectral range that includes visible, near-infrared, and shortwave infrared regions; their spatial resolutions are 10, 20, and 60 m, respectively (Figure 1). The most innovative feature of the Sentinel-2 satellite is that it contains rich red-edge bands information with center wavelengths of 705, 740, and 783 nm, which provides abundant information for vegetation biophysical status monitoring and estimation [23,24]. In summary, Sentinel-2 has an unprecedented spatial, temporal resolution and revisit cycle, which is suitable for monitoring crop growth processes such as crop diseases or pest stress [25].
Yellow rust is an air-dispersed pandemic disease [2]; its occurrence is strongly related to habitat conditions, such as humidity, sunshine, and temperature. However, most of the existing remote sensing identification methods for wheat yellow rust depend on spectral data; few studies have considered the habitat characteristics of yellow rust disease [8]. In this study, we considered the characteristics of spectral changes and meteorological factors in the occurrence period of wheat yellow rust; and developed a large-scale and high-precision monitoring method for wheat yellow rust on a regional scale that integrates environment conditions and growth status. The primary aims of this research are: (1) To present a series of two-temporal vegetation indices for monitoring wheat yellow rust using two-stages satellite remote sensing images on a regional scale; (2) to explore the feasibility of combining remote sensing and meteorological information for yellow rust monitoring; and (3) to develop an optimal classification method for monitoring wheat yellow rust utilizing both spectral information and meteorological data on a regional scale.

2. Materials and Methods

2.1. Field Survey and Data Collection

2.1.1. Field Survey Area

Field surveys of wheat yellow rust were conducted in Ningqiang county (118°35'19.5" E, 37°35'51.75" N), Shaanxi province, China from 11 May 2018 to 14 May 2018 (i.e., the filling stage, Figure 2). In Ningqiang County, wheat is a major crop, and yellow rust is the dominant wheat disease; a severe infestation occurred in 2018. The average annual temperature and precipitation of this area are 13 °C, and 1812.2 mm, respectively; low temperature and high humidity environment conditions are conducive to the occurrence, spread, and epidemic of wheat yellow rust [23]. Shaanxi Province is considered to be an important spring epidemic area and winter breeding area of wheat yellow rust in China [26]. Therefore, there is an urgent demand to monitor wheat yellow rust disease in this region.

Three data types, including field survey data of yellow rust disease, multispectral satellite images, and meteorological information, were collected to develop a wheat yellow rust monitoring model on a regional scale.

![Figure 1. Parameters and information of Sentinel-2 multispectral satellite.](image-url)
Three data types, including field survey data of yellow rust disease, multispectral data carried 13 bands that include three different spatial resolution (Figure 1). For subsequent analysis, the spatial resolutions of the 13 bands were resampled to 10 m using the resampling tool in the software. The large-scale crop disease monitoring was based on the extraction of wheat planting area; therefore, we used the decision tree and multi-temporal phenological information methods, as proposed by Zhang et al. and Xu et al., to extract the planting area of wheat [14,27]. Field survey points were used to verify the accuracy of the extracted wheat area, which reached 94%. This result meets the demand of subsequent remote sensing monitoring of crop diseases.

2.1.2. Disease Survey Data Collection

In this field survey, 58 samples (22 healthy, 36 infected) were investigated from 11 May 2018 to 14 May 2018. Considering the pixels size of remote sensing images, uniformly growing wheat samples were randomly selected with a continuous area of 10 × 10 m, and the severity of disease was surveyed. We selected five representative 1 × 1 m plots (located at the four corners and centers of the 10 × 10 m plots), and the average severity of yellow rust for the five plots was used to represent the disease degree for one sample [8]. The center coordinate of each sample was recorded by a differential global positioning system (GPS) sensor (Trimble GeoXH). The field severity survey and disease index calculation of wheat yellow rust referenced the rule of the National Rules for the Investigation and Forecasting of Crop Diseases (GB/T 15795-1995) [12]. The information of wheat growth, disease incidence, and location are recorded in Appendix A Table A1.

2.1.3. Remote Sensing Data Collection and Wheat Planting Area Extraction

Sentinel-2A remote sensing images (processing level 1C) on 2 April 2018 (early onset of disease) and 12 May 2018 (disease outbreak stage) were downloaded from the European Space Agency Sentinels Scientific Date Hub (https://scihub.copernicus.eu/) for the study region [23]. The preprocessing of Sentinel-2A images included atmospheric correction, and clipping. Atmospheric correction was performed using the Sen2cor module (version 2.2.1) within the Sentinel-2 Toolbox, and image mosaic and cropping were implemented in the Sentinel Application Platform software (SNAP, 4.0.2) [23]. In addition, the Sentinel-2A multispectral data carried 13 bands that include three different spatial resolution (Figure 1). For subsequent analysis, the spatial resolutions of the 13 bands were resampled to 10 m using the resampling tool in the software. The large-scale crop disease monitoring was based on the extraction of wheat planting area; therefore, we used the decision tree and multi-temporal phenological information methods, as proposed by Zhang et al. and Xu et al., to extract the planting area of wheat [14,27]. Field survey points were used to verify the accuracy of the extracted wheat area, which reached 94%. This result meets the demand of subsequent remote sensing monitoring of crop diseases.

2.1.4. Meteorological Data

Meteorological data are the basis for analyzing and describing climate characteristics and their laws of change [28]. The occurrence and prevalence of wheat yellow rust depend on the interaction among wheat varieties, amount of yellow rust disease, and environmental conditions [1,4,8]. When both pathogen and host have the potential for an epidemic, environmental conditions, specifically meteorological conditions, become the dominant factor in a wheat yellow rust epidemic [4,29].

Figure 2. Map of Ningqiang county, Shaanxi province in China. Green crosses denote field survey locations of healthy wheat, red denote field survey locations of yellow rust-infested wheat.
Considering the influence of climate conditions on the wheat infection of yellow rust pathogens, five types of meteorological data were collected for March to May 2018 from the National Meteorological Information Center in 37 sampling sites around Ningqiang county, including average temperature (TEM), precipitation (PRE), sunshine hours (SSD), wind speed (WIN), and relative humidity (RHU). From each of these, we calculated a monthly mean value for March, April, and May. Therefore, a total of 15 meteorological features were calculated in this study.

### 2.2. Vegetation Indices for Plant Diseases Discrimination

Crop under the stress of pests and diseases often undergo changes that impact their spectral properties, including pigmentation, moisture, and biomass. The sensitive spectral bands were combined to construct vegetation indices (VIs) in the relevant mathematical forms. VIs related to plant growth status, vegetation coverage, and pigmentation content were used to capture the physiological and biochemical changes caused by wheat yellow rust infection (Table 1).

#### Table 1. Multispectral vegetation indices for wheat yellow rust discrimination.

<table>
<thead>
<tr>
<th>Vegetation Indices</th>
<th>Name</th>
<th>Formula</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDVI</td>
<td>Normalized difference vegetation index</td>
<td>((R_{NIR} - R_R)/(R_{NIR} + R_R))</td>
<td>[30]</td>
</tr>
<tr>
<td>GNDVI</td>
<td>Green normalized difference vegetation index</td>
<td>((R_{NIR} - R_G)/(R_{NIR} + R_G))</td>
<td>[31]</td>
</tr>
<tr>
<td>SAVI</td>
<td>Soil adjusted vegetation index</td>
<td>(((R_{NIR} - R_R) + 1.5)/(R_{NIR} + R_R + 0.5))</td>
<td>[32]</td>
</tr>
<tr>
<td>SIPI</td>
<td>Structural independent pigment index</td>
<td>((R_{NIR} - R_B)/(R_{NIR} - R_R))</td>
<td>[33]</td>
</tr>
<tr>
<td>EVI</td>
<td>Enhanced vegetation index</td>
<td>(2.5(R_{NIR} - R_R)/(R_{NIR} + 6R_R - 0.5R_B + 1))</td>
<td>[34]</td>
</tr>
<tr>
<td>RDVI</td>
<td>Normalized difference vegetation index</td>
<td>((R_{NIR} - R_R)/\sqrt{(R_{NIR} + R_R)})</td>
<td>[35]</td>
</tr>
<tr>
<td>RGR</td>
<td>Ration of red and green</td>
<td>(R_R/R_G)</td>
<td>[36]</td>
</tr>
<tr>
<td>VARIgreen</td>
<td>Visible atmospherically resistant index</td>
<td>((R_G - R_R)/(R_G + R_R))</td>
<td>[37]</td>
</tr>
<tr>
<td>NDVIred1</td>
<td>Normalized difference vegetation index</td>
<td>((R_{NIR} - R_{Re1})/(R_{NIR} + R_{Re1}))</td>
<td>[30]</td>
</tr>
<tr>
<td>NREDI1</td>
<td>Normalized red—edge1 index</td>
<td>((R_{Re2} - R_{Re1})/(R_{Re2} + R_{Re1}))</td>
<td>[38]</td>
</tr>
<tr>
<td>NREDI2</td>
<td>Normalized red—edge2 index</td>
<td>((R_{Re3} - R_{Re1})/(R_{Re3} + R_{Re1}))</td>
<td>[38]</td>
</tr>
<tr>
<td>NREDI3</td>
<td>Normalized red—edge3 index</td>
<td>((R_{Re3} - R_{Re2})/(R_{Re3} + R_{Re2}))</td>
<td>[38]</td>
</tr>
<tr>
<td>PSRI1</td>
<td>Plant senescence reflectance index</td>
<td>((R_G - R_R)/R_{Re1})</td>
<td>[24]</td>
</tr>
<tr>
<td>REDI</td>
<td>Red—edge disease stress index</td>
<td>(((705 - 665)(R_{Re3} - R_R) - (783 - 665)(R_{Re1} - R_R))/(2R_R))</td>
<td>[23]</td>
</tr>
</tbody>
</table>

#### 2.3. Two-Stage Vegetation Index for Wheat Yellow Rust Monitoring

The study area belongs to the wheat region of southwest China. Generally, the initial stage of wheat yellow rust disease in this region is from the end of March to the beginning of April, the yellow rust outbreak occurred in mid-May in 2018. Accordingly, we selected Sentinel-2 images acquired on 2 April and 12 May 2018. Based on the commonly used vegetation indices listed in Table 1, we calculated the change in magnitude from 2 April to 12 May using the normalization quantification formula:

\[
nVIs = \frac{VI_{12\text{May}} - VI_{2\text{April}}}{VI_{12\text{May}} + VI_{2\text{April}}}\tag{1}
\]
where nVIs represents the change of vegetation index features between two-stages; VI\textsubscript{2April} and VI\textsubscript{12May} indicate the values of the vegetation index extracted from the images at the time of the first occurrence (2 April 2018) of yellow rust and at the large outbreak of yellow rust (12 May 2018), respectively.

2.4. Spectral VIs and Meteorological Features Importance Ranking

There are many features including vegetation indices and meteorological parameters that are potentially relevant to crop diseases monitoring, however the sensitivity of these features varies substantially. It is necessary to describe the degree that how much a feature will impact on the model predictions. In this study, random forest (RF) was applied for classification and feature importance analysis and was first described by Breimen et al. [39]. It is an ensemble approach for building decision trees for predictions. The feature importance in RF is computed as the average contribution of each feature on each tree in the RF [39]. We used the out-of-bag (OOB) data to calculate the error (errOOB\textsubscript{t}) for each tree in the RF algorithm. Subsequently, we compared the difference in the OOB error of each feature before (errOOB\textsubscript{b}) and after adding noise (errOOB\textsubscript{i}) to calculate the importance of the feature (X), where Ntree denotes the number of trees in RF [22]. Finally, the importance of feature X\textsubscript{i} was defined as:

\[
V(X^i) = \frac{1}{N} \sum_t (\text{errOOB}_t - \text{errOOB}_t^i)
\]  

In addition, we used the analysis of variance methods to test the significance of the selected features [40]. The statistical significance expressed by the ρ value reflects the suitability of the feature [9,40]. Finally, we selected the features that were highly important and significant as the optimal features for yellow rust detection.

2.5. Monitoring Methods

The main purpose of this research is to explore the feasibility of remote sensing monitoring of crop diseases by meteorological data information. In addition, because of the small sample size in this experiment, three commonly used methods of liner discriminant analysis (LDA), SVM, and ANN and regular parameter settings were selected to construct wheat yellow rust monitoring models.

LDA is a dimensionality reduction method based on the best classification effect [41], usually by finding a set of linear feature combinations to classify two or more targets. The primary idea is to find a linear combination of variables to maximize in-between variance and minimize within-class variance [41]. The LDA model was implemented using the Statistical Package for the Social Sciences (SPSS 20.0). The parameters were set as default value.

In the SVM classification algorithm, the primary idea is to determine an optimal decision boundary and maximize the distance of the closest samples in two categories as much possible across the boundary [42]. Using the radial basis function (RBF) as the kernel function for SVM classification exhibited superior performance in the case of inseparable linearity [17]. The key parameters of SVM are shown in Table 2. The model is trained and tested in the Matlab R2016 software.

ANNs can be described as parallel and complex computing systems composed of large numbers of interconnected simple processors (neurons, also called nodes) [43]. As an important data mining tool, ANNs have comprehensive mathematical mechanisms and have been applied in various fields of remote sensing, such as ground objects identification and change detection [18,43]. In this study, the vegetation indices and meteorological data were the ANN input parameters. The transfer functions of logarithm sigmoid (logsig) transfer and linear (purelin) were used to activate the hidden layers and weighted output layers, respectively [44,45]. The learning rule takes the approach of a gradient descent backpropagation (traingd) training function. The key parameters of ANN are shown in Table 2. We used the MATLAB R2016 software to run the ANN models.
Once sensitive spectral features and meteorological features were identified, the three classification algorithms (LDA, SVM, and ANN) were used to establish a classification model for wheat yellow rust. According to the three methods, wheat yellow rust classification was conducted using the following datasets: case 1: spectral vegetation indices (containing single temporal VIs and two-temporal nVIs); case 2: a combination of spectral vegetation indices and meteorological information.

### 2.6. Classification Accuracy Assessment

Considering the number of samples (n = 58, of which healthy samples = 22, and disease samples = 36), the samples were divided into three parts, two of which (n = 39) were used as training samples for model training; with the remaining part (n = 19) used as a validation dataset to verify the accuracy of the model. Accuracy evaluation was determined by the overall classification accuracy (OA), kappa coefficient, user’s accuracy (U.a), and producer’s accuracy (P.a) from confusion matrices [46].

### 3. Results

#### 3.1. Vegetation Index Response for Yellow Rust Disease

The responses of the 14 VIs and corresponding two-stage vegetation indices (nVIs) are shown in Figure 3, where their mean and difference values are compared for healthy and yellow rust-infected wheat. The values of the re-normalized difference vegetation index (RDVI), and red-edge disease stress index (REDSI) were reduced by 100 times to maintain the same magnitude as that of the other indices. On the single-stage indices, according to the difference information of each vegetation index in healthy and diseased wheat, RDVI, REDSI, enhance vegetation index (EVI), soil adjusted vegetation index (SAVI), and normalized red-edge3 index (NREDI3) were most suitable for discriminating healthy and yellow rust-infected wheat (Figure 3a). Among them, the difference value of RDVI for healthy and disease samples was the largest, reaching 5.5. However, the magnitude of the difference value in the entire single-stage VIs was relatively small, except for RDVI and REDSI.

![Figure 3. Mean and difference values for healthy wheat and yellow rust infected wheat. (a) Single-stage vegetation indices and (b) normalized two-stage vegetation indices.](image-url)
Regarding the normalized two-stage indices, the difference between health samples and yellow rust infection samples was evident, specifically for the normalized REDSI (nREDSI), normalized visible atmospherically resistant index (nVARIgreen), normalized difference vegetation index reg-edge 1 (nNDVIre1), and normalized plant senescence reflectance index (nPSRI1). Among them, the difference between healthy and diseased samples in nREDSI was up to 1.1 (Figure 3b). The two-stage nVIs (using images from both 2 April and 12 May 2018) exhibited a greater difference between healthy and yellow rust-infested wheat compared with the corresponding single-stage VIs (using the images from 12 May). This confirms that nVIs are closely related to the pathological progress of the crop, which can more clearly reflect leaf wilting, leaf tissue death, and canopy structure changes caused by the yellow rust pathogen.

3.2. Meteorological Data Processing and Selection

Each type of meteorological data was averaged by month to obtain its monthly average value. Then, the meteorological factors from March to May 2018 were spatially interpolated using an inverse distance weighted method in the ArcGIS software for subsequent continuous spatial pixel-scale analysis [9]. All meteorological factors were interpolated at a spatial resolution of 10 m, which matches the resolution of the Sentinel-2 satellite imagery. Figure 4 presents the results of meteorological data spatial resolution in May Finally, based on continuous meteorological data, meteorological characteristics were extracted for wheat yellow rust habitat monitoring.

Figure 4. Spatial interpolation of meteorological data in Ningqiang county in May 2018. (a) Average temperature (TEM; 0.1 °C), (b) average relative humidity (RHU; %), (c) average sunshine hours (SSD; 0.1 h), (d) average wind speed (WIN; 0.1 m/s), (e) average of precipitation (PRE; 0.1 mm).
3.3. VIs and Meteorological Features Sensitivity of Yellow Rust Monitoring

We observed a strong correlation between VIs as well as wheat physiological and, biochemical parameters caused by the development of yellow rust; however, correlations and multiple collinearities among different VIs limit the extraction of sensitive information for wheat yellow rust discrimination. Therefore, we selected the single- and two-stage VIs most sensitive to reflect the state of the crop after being stressed by disease using the important criterion in the RF method. The selected single-stage vegetation index, normalized two-stage vegetation index, and meteorological data were used to determine the important features for yellow rust detection (Figure 5).

![Figure 5](image)

**Figure 5.** Variable importance of vegetation indices and meteorological factors in identifying yellow rust-infected wheat. (a) Vegetation indices (VIs) and normalized VIs (nVIs) based on a single- and two-stage imagery; (b) meteorological data from March to May 2018.

The relative importance of the three features for wheat yellow rust discrimination were analyzed using the RF method (Figure 5). According to the importance ranking of spectral VIs (Figure 5a), we selected the features with variable importance greater than 0.05 for subsequent analysis. In terms of single-stage VIs, the RDVI, REDSI, VARIgreen, NREDI3, RGR, PSRI1, NDVIre1, SAVI, and EVI were selected; for nVIs, the nREDSI, nVARIgreen, nPSRI1, nNDVIre1, nNREDI1, nNREDI2, nSAVI, and nNREDI3 were selected. For meteorological data, the SSD_03, PRE_05, RHU_04, PRE_04, RHU_05, WIN_03, WIN_04, TEM_03 and SSD_05 were selected. To avoid information redundancy of selected features, we used the analysis of variance (ANOVA) method to optimize important features (Table 3). For VIs, the EVI, NREDIre1, PSRI1, and SAVI showed no significant differences ($p > 0.05$); and for nVIs, the differences of nNREDI3, nNREDI2, and nSAVI with other nVIs were insignificant ($p > 0.05$).

The variable importance values of the vegetation indices based on a single image and two images taken at different times in the discrimination of wheat yellow ruts differed. The five most important vegetation indices were selected for subsequent analysis. For single-stage imagery, the RDVI, REDSI, NREDI3, RGR, and VARIgreen were most sensitive to wheat yellow rust (Figure 5a, indigo histogram); for two-stage imagery, nREDSI, nVARIgreen, nPSRI1, nNREDI1, and nNDVIre1 were most sensitive to wheat yellow rust (Figure 5a, magenta histogram). This is generally consistent with the results shown in Figure 3b and allows us to distinguish between healthy wheat and yellow rust infection. Similarly, the meteorological features of WIN_03, WIN_04, TEM_03, and SSD_05 were excluded using ANOVA.

Figure 5b demonstrated the importance ranking of meteorological features. Due to the strong correlation between the same type of meteorological data, we selected the features above the average value (0.07) of all variable importances for yellow rust identification. Accordingly, five of the most important meteorological features were selected for monitoring wheat yellow rust on a regional scale: average sunshine hours in March (SSD_03),

...
average relative humidity (RHU_04, RHU_05) in April and May, and average precipitation (PRE_04, PRE_05) in April and May.

Table 3. ANOVA for vegetation indices feature sets (VIs and nVIs).

<table>
<thead>
<tr>
<th>VIs</th>
<th>RDVI</th>
<th>NDVIRE1</th>
<th>NREDI3</th>
<th>VARIg</th>
<th>EVI</th>
<th>REDSI</th>
<th>RGR</th>
<th>SAVI</th>
<th>PSRI1</th>
</tr>
</thead>
<tbody>
<tr>
<td>RDVI</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NDVIRE1</td>
<td>0.0180</td>
<td>a</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NREDI3</td>
<td>0.0047</td>
<td>b</td>
<td>0.31</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VARIg</td>
<td>0.0087</td>
<td>b</td>
<td>0.066</td>
<td>0.043</td>
<td>a</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EVI</td>
<td>0.1250</td>
<td></td>
<td>0.058</td>
<td>0.032</td>
<td>a</td>
<td>0.04</td>
<td>a</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>REDSI</td>
<td>0.0016</td>
<td>b</td>
<td>0.132</td>
<td>0.0007</td>
<td>b</td>
<td>0.002</td>
<td>b</td>
<td>0.15</td>
<td>-</td>
</tr>
<tr>
<td>RGR</td>
<td>0.0003</td>
<td>c</td>
<td>0.027</td>
<td>0.0035</td>
<td>b</td>
<td>0.012</td>
<td>a</td>
<td>0.03</td>
<td>a</td>
</tr>
<tr>
<td>SAVI</td>
<td>0.0910</td>
<td></td>
<td>0.132</td>
<td>0.001</td>
<td>b</td>
<td>0.019</td>
<td>a</td>
<td>0.16</td>
<td>0.001b</td>
</tr>
<tr>
<td>PSRI1</td>
<td>0.0370</td>
<td>a</td>
<td>0.224</td>
<td>0.0009</td>
<td>c</td>
<td>0.32</td>
<td>0.06</td>
<td>0.129</td>
<td>0.009b</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>nVIs</th>
<th>nREDSI1</th>
<th>nNREDI3</th>
<th>nVARIg</th>
<th>nNDVIRE1</th>
<th>nNREDI1</th>
<th>nPSRI1</th>
<th>nNREDI2</th>
<th>nSAVI</th>
</tr>
</thead>
<tbody>
<tr>
<td>nREDSI1</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nNREDI3</td>
<td>0.14</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nVARIg</td>
<td>0.0001</td>
<td>c</td>
<td>0.14</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nNDVIRE1</td>
<td>0.072</td>
<td></td>
<td>0.06</td>
<td>0.019</td>
<td>a</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nNREDI1</td>
<td>0.002</td>
<td>b</td>
<td>0.27</td>
<td>0.0004</td>
<td>c</td>
<td>0.0002</td>
<td>c</td>
<td>-</td>
</tr>
<tr>
<td>nPSRI1</td>
<td>0.0005</td>
<td>c</td>
<td>0.38</td>
<td>0.0027</td>
<td>b</td>
<td>0.0004</td>
<td>c</td>
<td>0.0002</td>
</tr>
<tr>
<td>nNREDI2</td>
<td>0.35</td>
<td></td>
<td>0.06</td>
<td>0.0003</td>
<td>a</td>
<td>0.026</td>
<td>a</td>
<td>0.0004</td>
</tr>
<tr>
<td>nSAVI</td>
<td>0.0004</td>
<td>c</td>
<td>0.074</td>
<td>0.15</td>
<td>0.0001</td>
<td>c</td>
<td>0.11</td>
<td>0.0009</td>
</tr>
</tbody>
</table>

Note: "a" indicates the difference is significant at the 0.95 confidence level, "b" indicates the difference is significant at the 0.99 confidence level, and "c" indicates the difference is significant at the 0.999 confidence level; "VARIg" = VARIgreen, "nVARIg" = nVARIgreen,” and "mete" = meteorological.

3.4. Wheat Yellow Rust Monitoring Based on Spectral Vegetation Indices

Monitoring models for wheat yellow rust were built using the LDA, SVM, and ANN algorithms. The RDVI, REDSI, NREDI3, RGR, and VARIg were selected for use in single-stage monitoring models; nREDSI, nVARIg, nPSRI1, nNREDI1, and nNDVIRE1 were selected for use in two-stage monitoring models. Table 4 presents the classification results of the three algorithms using the different VIs.

Table 4. Classification accuracies of three classification algorithms using single- and two-stage vegetation indices (VIs and nVIs, respectively; n = 19).

<table>
<thead>
<tr>
<th>Classifier</th>
<th>VIs</th>
<th>nVIs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LDA</td>
<td>Healthy</td>
</tr>
<tr>
<td>Healthy</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>YR</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>U.a (%)</td>
<td>50.0</td>
<td>69.2</td>
</tr>
<tr>
<td>SVM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Healthy</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>YR</td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>U.a (%)</td>
<td>66.7</td>
<td>76.9</td>
</tr>
<tr>
<td>ANN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Healthy</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>YR</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>U.a (%)</td>
<td>50.0</td>
<td>72.7</td>
</tr>
</tbody>
</table>

Note: Pa = producer’s accuracy, U.a = user’s accuracy, OA = overall classification accuracy.

For the single-stage vegetation index model, the overall classification accuracy and kappa coefficient were 63.2% and 0.18 for the LDA algorithm, respectively; 73.7% and 0.42 for the SVM algorithm, respectively; and 63.2% and 0.23 for the ANN algorithm,
respectively. For the nVIs models, the overall classification accuracy and kappa coefficient were 68.4% and 0.32 for the LDA algorithm, respectively; 78.9% and 0.55 for the SVM algorithm, respectively; and 68.4% and 0.32 for the ANN algorithm, respectively. Based on these results, the classification accuracy of wheat yellow rust monitoring models using nVIS as the input features is better than that of models using VIS; the overall accuracy is improved by 5.2%. Compared with the VIS model, the Pa of healthy wheat identification and yellow rust wheat exceeded 57.1% and 83.3%, respectively, for nVIs. Among the algorithms, SVM performed the best.

### 3.5. Wheat Yellow Rust Monitoring Based on Meteorological Data and Spectral Information

The classification results of the three algorithms based on both vegetation indices (VISs and nVISs) and meteorological data are shown in Table 5. For the single-stage vegetation index model, the overall classification accuracy and kappa coefficient were 68.4% and 0.32 for the LDA algorithm, respectively; 78.9% and 0.55 for the SVM algorithm, respectively; and 73.7% and 0.45 for the ANN algorithm, respectively. For the two-stage models, the overall classification accuracy and kappa coefficient were 73.7% and 0.42 for the LDA algorithm, respectively; 84.2% and 0.65 for the SVM algorithm, respectively; and 78.9% and 0.55 for the ANN algorithm, respectively.

<table>
<thead>
<tr>
<th>Classifier</th>
<th>VISs_Meteorological Data</th>
<th>nVISs_Meteorological Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Healthy</td>
<td>YR</td>
</tr>
<tr>
<td>LDA</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>57.1</td>
<td>75.0</td>
</tr>
<tr>
<td>SVM</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>71.4</td>
<td>83.3</td>
</tr>
<tr>
<td>ANN</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>71.4</td>
<td>83.3</td>
</tr>
</tbody>
</table>

According to these results, the accuracies of wheat yellow rust monitoring models with the nVISs and meteorological data as the input features are higher than those based on VISs and meteorological data. This is consistent with the results based on the pure VISs model (see Section 3.4). Among the three algorithms, SVM again had the highest classification accuracy. Moreover, the Ua of healthy and yellow rust wheat identification was 83.3% and 84.6%, respectively; the Pa of yellow rust reached 91.7% in the nVISs and meteorological model (nVISs_meteorological data). The results confirm that the inclusion of meteorological data improves model accuracy and offers the potential for crop disease monitoring on a regional scale.

### 3.6. Mapping Wheat Yellow Rust Using the Optimal Monitoring Model

Figure 6 shows a map of wheat yellow rust in Ningqiang county, Shaanxi Province during the filling period based on the optimal model (SVM algorithm using two-stage spectral vegetation indices and meteorological data). The wheat yellow rust infected region is highly consistent with the field observation, which verifies the feasibility of the model for crop disease monitoring. This remote sensing method has the potential for effective, rapid (near real-time), and a spatially continuous regional monitoring of crop disease, offering substantial labor-, time-, and cost-savings.
4. Discussion

Remote sensing data has the characteristics of spatial continuity and rich information, which facilitates the acquisition of crop growth and environmental information, and provided a basis for crop pest monitoring [8,47]. This study explored the potential of spectral VIs and meteorological information related to disease occurrence to monitor wheat yellow rust infestation on a regional-scale.

4.1. Performance of Spectral Vegetation Indices in Wheat Yellow Rust Discrimination

VIs can reflect the biophysical and biochemical change of crops, and can be used for detection and identification of plant diseases [37,48]. Yellow rust primarily infects wheat leaves, causing green fading and deformation of leaf tissue, thereby significantly changing the chlorophyll content and biomass [23]. We selected VIs with highly sensitive yellow rust discrimination during the wheat milking stage based on single- and two-stage remote sensing images. The Sentinel-2 satellite has rich red-edge information that are significant for crop growth status and stress monitoring [22,23]. In particular, REDSI consists of red edges and bands and was proposed by Zheng for monitoring wheat yellow rust, particularly during the filling stage [12,23]. In this study, REDSI and VARIgreen were more important for wheat yellow rust discrimination among the single-stage indices; PSRI1, NREDI1, and...
NDVIre1 were most sensitive wheat yellow rust among the two-stage indices. This is primarily related to the destruction of the tissue structure of leaf cells and the decrease of leaves chlorophyll content under yellow rust disease stress, resulting in the shift of the spectrum on the red edge \[22,49\]. PSRI1 can be used to assess the crop pigment content and status \[10\]. Moreover, the band combination of nPSRI1, nNREDI1, and nNDVIre1 contains red-edge information that can capture changes in physiological and biochemical parameter, and better eliminate the effects of growth factors compared with single-stage vegetation index models (classification accuracy is 5.2% higher, Table 4) \[23\]. Here, the optimal model (i.e., that using the two-stage vegetation indices) captured changes caused by yellow rust disease with a classification accuracy of 78.9%.

4.2. Performance of Meteorological Data in Wheat Yellow Rust Discrimination

The propagation, spread, and infection of pathogen spores require suitable environmental conditions (such as, precipitation, humidity, and temperature). Wheat yellow rust disease occurs in high humidity and low-temperature environments. Favorable climate conditions such as warm winters and heavy rainfall in early spring are external causes of wheat yellow rust occurrence and epidemics in Shaanxi Province \[26\]. In this study, the average relative humidity (RHU_04, RHU_05) and average precipitation (PRE_04, PRE_05) were sensitive to yellow rust discrimination (Figure 5b). Moreover, the TEM in Shaanxi Province reaches a suitable range for the incidence of wheat yellow rust in April and May. The study area belongs to the winter breeding region of wheat yellow rust in China \[50\]. That is, the yellow rust pathogen in this area infects wheat during the winter, making WIN less important in the monitoring of wheat yellow rust than RHU and PRE \[26\]. However, as wheat yellow rust disease is an air-borne bacterium, WIN can provide important information for forecasting. In summary, our results confirm that meteorological information can provide crop disease monitoring, which is consistent with the conclusions of Yuan et al. \[8\].

4.3. Performance of Wheat Yellow Rust Monitoring Classification Algorithms

Among LDA, SVM, and ANN, the SVM algorithm exhibited the best performance for distinguishing healthy and yellow rust-infested wheat, with a classification accuracy of 73.7–84.2%. The SVM classifier is based on the threshold discriminant rule and maps the samples to appropriate feature space. Some researchers have also shown that the SVM is superior to LDA in remote sensing classification or extraction in plants \[40,41,51\]. For example, Yue et al. reported that SVM-based models achieve higher classification accuracies than those using LDA in wheat yellow rust monitoring on leaf scale \[40\].

In terms of classification accuracy, SVM outperformed the ANN classifiers by 5.2–10.5% in different feature spaces. These results differ from those of Raczkó et al. who found that ANN performed better than the SVM model \[18\]. However, ANNs are more difficult to use and optimize, and require many parameters. The number of samples in this study was limited, and ANN requires a large number of parameters to set the initial value of the network topology, weights, and thresholds, thereby making it difficult to optimize the model \[18\]. Compared with ANNs, the SVM algorithm can solve classification problems for nonlinear and small sample situations, and avoid the neural network structure selection and local minima problem \[52\]. Overall, considering that monitoring and positioning crop disease on a regional scale are more complicated and challenging than at the canopy and on leaf scales, the classification accuracy achieved in this study (73.7–84.2% based on SVM classifier) is acceptable.

Although the current classification accuracy is lower than that obtained based on airborne hyperspectral images (for example, Zhang et al. used deep convolutional neural network to identify wheat yellow rust based on airborne hyperspectral images with an accuracy of 85.0% \[53\]), it meets the practical demands of disease monitoring and management. It is majorly based on airborne hyperspectral images, for which fine spectral resolution enables more abundant spectral information to be extracted and analyzed, which may lead to a certain improvement in the accuracy of disease mapping. Many researchers have
used medium- and high-resolution satellite images to monitor crop disease. Chemura et al. used spectral indices to identify coffee leaf rust infection based on Sentinel-2 satellite data, with the discrimination accuracy of 82.5% [22]. Yuan et al. used the crop growth index (GNDVI and VARIred-edge) and environmental characteristics to monitor crop disease and pests based on the Worldview2 and Landsat 8 satellite, and proved that the accuracy (82.0%) of models combining VIs and environmental characteristics are better than those of traditional monitoring models that only rely on spectral information [8]. However, despite the significant potential for crop disease monitoring, we should optimize the parameters of the methods to build more robust and reasonable models under the condition of enough samples, and improve the accuracy of crop disease monitoring for practical applications.

In this study, adding monthly average meteorological data to the remote sensing monitoring of crop diseases, we established an effective remote sensing monitoring model for wheat yellow rust. However, crop disease occurrence is also the result of environmental factors, the amount of pathogen, crop planting landscape patterns, and farmland management [4,29]. Therefore, future work should integrate more multi-source data (remote sensing and non-remote sensing data) with well-characterized mechanisms and high stability to further improve crop disease monitoring and forecasting. Moreover, due to the influence of weather and manpower, the sample size of wheat yellow rust in this study was small. In the future, we will collect wheat yellow rust data from large areas in different years to verify and improve the wheat yellow rust monitoring model. Furthermore, we will attempt to effectively utilize the complementary features of meteorological data (for example, various types of meteorological data and 10-day average meteorological data), terrain features, and remote sensing data to establish a collaborative scheme for forecasting crop disease at an early stage.

The rapid and large-scale monitoring of crop disease and pests relieves huge pressure on plant protection personnel. It is a weapon to prevent and control disease, promote healthy development of agriculture, and achieve the goal of sustainable agricultural development. In addition, this will contribute to eradicating hunger, achieving food security, improving nutrition and promoting sustainable agriculture as outlined in the United Nations Sustainable Development Goals.

5. Conclusions

In this study, multispectral satellite imagery (Sentinel-2A) and meteorological data were used to monitor wheat yellow rust disease based on three classification methods (linear discriminant analysis, a support vector machine, and an artificial neural network) on a regional scale. Five meteorological features (sunshine hours in March (SSD_03), average relative humidity in April and May (RHU_04, RHU_05), and average precipitation in April and May (PRE_04, PRE_05)) combined with two-stage vegetation indices using the SVM algorithm were found to be optimal for wheat yellow rust monitoring. In addition, the model for yellow rust monitoring base of two-stage vegetation indices significantly outperformed single-stage vegetation index models, with the overall classification accuracy increasing from 63.2% to 78.9%. Moreover, the addition of meteorological data, which is closely related to yellow rust occurrence, increased the accuracy of the two-stage index SVM model to 84.2%. The proposed model is suitable for rapid, large-scale monitoring and forecasting of biotic (bacterial and fungal disease) stress in crops and offers an effective approach for reducing the impacts of crop disease, including the implications for global food security. In the future, we will consider information from multiple sources to develop further comprehensive and reliable crop disease forecasting models.

Author Contributions: Q.Z.: field survey, methodology, writing—original draft. H.Y. and W.H.: conceived and designed the experiments; Y.D.: data collection; H.J.: software and processed the data; C.W., D.L., and L.W.: provided advices to improve manuscript; S.C.: modified the structure of the paper and grammar. All authors have read and agreed to the published version of the manuscript.
Funding: This research was funded by GDAS’ Project of Science and Technology Development (2020GDASYL-20200103004), Guangdong Province Agricultural Science and Technology Innovation and Promotion Project (No.2020KJ102), Guangzhou Basic Research Project (202002020076), National special support program for high-level personnel recruitment (Wenjiang Huang).

Institutional Review Board Statement: This study not involving humans.

Informed Consent Statement: This study not involving humans.

Data Availability Statement: Data sharing is not application to this article.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Field survey record form for wheat yellow rust.

<table>
<thead>
<tr>
<th>Contents</th>
<th>1</th>
<th>2</th>
<th>…</th>
</tr>
</thead>
<tbody>
<tr>
<td>Survey date</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Location (longitude, latitude)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Crop growth information

- Cultivated varieties
- Growth stage
- Plant height
- Planting density

Disease information

- Number of leaves investigated
- Number of diseased leaves
- Disease severity (%)
- Damage percentage (%)

Remarks (Altitude, irrigation information, etc.)

References


