RESEARCH ARTICLE

Prediction Dynamics in Cotton Aphid Using Unmanned Aerial Vehicle Multispectral Images and Vegetation Indices

PINGAN JIANG[1], XUELIN ZHOU [ID][1,6], **(Graduate Student Member, IEEE), TONGLAI LIU[2], XIAOHU GUO[3], DEYING MA[3], CONG ZHANG[4], YAN LI[5], AND SHUANGYIN LIU[2]**

[1]College of Grassland Sciences, Xinjiang Agricultural University, Urumqi 830052, China
[2]College of Information Science and Technology, Zhongkai University of Agriculture and Engineering, Guangzhou 510225, China
[3]College of Agricultural, Xinjiang Agricultural University, Urumqi 830052, China
[4]College of Resource and Environment, Xinjiang Agricultural University, Urumqi 830052, China
[5]Faculty of Land and Resources Engineering, Kunming University of Science and Technology, Kunming 650221, China
[6]Zhuhai Orbita Aerospace Science and Technology Company Ltd., Zhuhai 519080, China

Corresponding authors: Xuelin Zhou (zhouxuelin@live.cn) and Tonglai Liu (tonglailiu@zhku.edu.cn)

ABSTRACT Cotton harvest can be increased by having real-time information on the state of cotton aphid populations. However, traditional cotton aphid monitoring relies on ground sample methods supported by models such as linear regression, resulting in low forecast accuracy. Therefore, this paper purposes to enhance the precision of the remote sensing prediction model by investigating the cotton aphid prediction model construction approach. We explored the effectiveness of the XGBoost algorithm combined with the GWO algorithm and SVR method for cotton aphid prediction relying on vegetation indices derived from UAV multispectral photography. Originally, 12 indices related to cotton aphids were calculated by UAV multispectral reflectance. Additionally, the optimal index combination for pest prediction was determined utilizing analysis of correction and two-way ANOVA, combined with the XGBoost algorithm. Furthermore, a pest prevalence prediction model for cotton aphids was constructed via the SVR methodology associated with the optimal catalog combination, and the model was optimized using the GWO algorithm. Compared with the seven algorithms, experimental results demonstrate that the MSE and MAE of the XGBoost-GWO-SVR model are reduced by 90.20% and 70.36% (SVR), 90.14% and 70.26% (XGBoost-SVR), 7.47% and 0.14% (XGBoost-GA-SVR), 5.80% and 0.11% (XGBoost-PSO-SVR), 12.06% and 58.95% (LR), and 84.77% and 89.22% (BPNN), whereas the R^2 is increased by 22.5% (SVR and XGBoost-SVR), 0.3% (LR), and 12.51% (BPNN). The R^2 of the prediction model of XGBoost-SVR combined with GWO, PSO, and GA is not significantly different. Among these models, the XGBoost-GWO-SVR obtained the highest R^2 of 0.980 and the lowest MAE of 2.838.

INDEX TERMS Cotton aphid, support vector regression, unmanned aerial vehicle, XGBoost algorithm, integrated pest management, grey wolf optimizer.

The associate editor coordinating the review of this manuscript and approving it for publication was Ines Domingues [ID].

I. INTRODUCTION

Xinjiang is the largest cotton-producing region in the country. Pests inhibit cotton development throughout the year, resulting in 10-15% yield loss, and pest stress is continuously

P. Jiang et al.: Prediction Dynamics in Cotton Aphid Using Unmanned Aerial Vehicle Multispectral Images and Vegetation Indices

IEEE Access

growing due to global climate change and insufficient ecological and natural management of cotton fields [1]. Cotton aphids are common pests in the cotton industry. Aphid damage results in early leaf development, slow growth, and a delayed growth cycle, all of which affect the quality and yield of cotton [2]. Aphis gossypii Glover (cotton aphid), Aphis atrata Zhang, and Acyrthosiphon gossypii Mordviko are the three aphid species that cause damage to cotton plants in Xinjiang. Cotton aphids can cause the greatest damage and incidence, and they have a high capacity to migrate, resulting in rapid pest outbreaks [2], [3].

Cotton aphids are migratory, movable, with rapid outbreaks and rapid spread, susceptible to environmental changes, with unstable population development and random distribution in cotton fields. Furthermore, they are affected by many characteristics with extensive coupling relationships. Thus, it is challenging to construct a general model for reliable prediction. Existing assessments and forecasts of the degree of cotton aphid infestation rely on human ground surveys, and prediction models built from field survey data offer higher accuracy but are challenging to promote for large-scale applications [4]. Intelligent information processing techniques are crucial for accurately predicting the dynamics of cotton aphid occurrence, understanding the changing patterns of cotton aphids in cotton fields over time, recognizing abnormal changes in cotton aphids, and providing early warnings [5], [6], [7]. Remote sensing technology detects insect activities by detecting damage inflicted by the insects such as sooty deposits, plant defoliation, color changes, and geometric deformation of plant tissue shapes [8], [9], [10]. Field pests can be recognized, studied, and forecasted using precise spectral information from airborne aerial and satellite sensors, as well as vegetation indicators [4], [11], [12]. Several signal processing techniques, machine learning methods, and pattern recognition algorithms have been continuously applied to models used in continuous monitoring, as UAV data and pest monitoring have improved, and machine learning models have seen widespread usage in pest detection [13], [14], [15]. In recent years, several researchers have employed hybrid models generated from several methodologies to make predictions, indicating that merging different models may improve the prediction performance [16], [17]. Despite tremendous progress, certain unavoidable limitations remain, such as a weak model generalization capacity, local optima, and under- and over-learning. The use of UAV multispectral data, vegetation index data, and machine learning approaches to target cotton aphids is received relatively little attention in prior research. In this paper, we employ a support vector regression (SVR, table 1 presents a cross-reference of phrases and abbreviations used in the text) model to predict the cotton aphid population status, integrating the benefits of XGBoost and numerous different optimization techniques, which may address issues such as the low accuracy of traditional cotton aphid prediction.

TABLE 1. Phrase abbreviation cross-reference table.

Phrases	ABBREV
Unmanned aerial vehicle	UAV
Support vector regression	SVR
Extreme gradient boosting	XGBoost
Grey wolf optimizer	GWO
Analysis of variance	ANOVA
Linear regression	LR
Back propagation neural network	BPNN
Genetic algorithm	GA
Particle swarm optimization	PSO
The Global Positioning System	GPS
Mean square error	MSE
Mean absolute error	MAE
radial basis kernel function	RBF

SVR is a unique small-sample learning approach that is not based on chance or the law of large numbers. In comparison to other statistical approaches, SVR produces superior results for representative slice fitting, and it is better adapted to the current situation of a long time span and small sample size for cotton aphid pests [18]. Therefore, SVR has been widely used in many fields, such as in predicting vegetation coverage in desert areas [19], estimation capability of cotton leaf nitrogen [20], accurate temperature forecasting in the field of waterfowl breeding [21], predicting chlorophyll-a in water bodies [22], predicting surface temperature [23], estimating farmland surface soil moisture [24], and predicting lithium battery life [25]. However, the SVR model has a complex network and overfitting owing to the enormous variety of parameters, and the uncertainty of coefficient C and tolerance coefficient ϵ affect the model accuracy, which must be optimized. The XGBoost method is a data screening method that can effectively reduce overfitting and computational effort in wheat stripe rust prediction feature screening and parthenium weed classification [26]. The optimization algorithms of the SVR algorithm are the GA algorithm, PSO algorithm, and GWO algorithm, in which the GWO algorithm can quickly solve the optimal solution of coefficient C and tolerance coefficient ϵ. It has been applied to ship identification, data filtering, and path planning optimization [27], [28], [29].

The main contributions of this paper are as follows:

- An improved XGBoost-GWO-SVR cotton aphid prediction model is proposed. It can estimate cotton aphid populations accurately and expand avenues for informative cotton aphid monitoring.

- Based on the information on vegetation spectral characteristics after the occurrence of pests and diseases, several new vegetation index algorithms are improved, such as the MSAVI-SAVI ratio index and the Simple ratio Green/Blue Vegetation Index.

- Compared with other algorithms, experimental results demonstrate that the XGBoost-GWO-SVR model can obtain better performance in prediction accuracy, and the combination of the GWO-SVR model and the XGBoost algorithm can enhance the prediction accuracy of cotton aphids.

IEEE Access

P. Jiang et al.: Prediction Dynamics in Cotton Aphid Using Unmanned Aerial Vehicle Multispectral Images and Vegetation Indices

TABLE 2. Field measurement activities in 2021.

Activity	Description
Seeding date	April 19, 2021
First detection of aphids' date	May 20, 2021
Cotton aphid measurements	5-day intervals
UAV imagery acquisition	50d, 66d, 73d, 88d from seeding date
Meteorological conditions during the monitoring period	Rainy in May; less rain in June and July; average temperatures in May [13 °C - 31 °C], June [14 °C - 32 °C], and July [18 °C - 35 °C]

TABLE 3. All vegetation indices for the sensitivity analysis.

Vegetation Index	ABBREV	Calculation Formulate	References
Atmospherically Resistant Vegetation Index	ARVI	$(\rho_{NIR} - 2 \cdot \rho_R + \rho_B)/(\rho_{NIR} + 2 \cdot \rho_R + \rho_B)$	[30]
Green-Blue Difference Vegetation Index	DVI	$\rho_G - \rho_B$	Customized
Simple ratio Green/Blue Vegetation Index	GBI	ρ_G/ρ_B	[31] Customized
Green Leaf Index	GLI	$(2 \cdot \rho_G - \rho_R - \rho_B)/(\rho_R + \rho_G + \rho_B)$	[32]
Transformed Chlorophyll Absorption Ratio	TCARI	$3 \cdot [(\rho_{RE} - \rho_R) - 0.2 \cdot (\rho_{RE} - \rho_G) \cdot (\rho_{RE}/\rho_R)]$	[33]
Soil Adjusted Vegetation Index	SAVI	$(1 + 0.16) \cdot (\rho_{NIR} - \rho_R)/(\rho_{NIR} + \rho_R + 0.16)$	[34]
Modified red edge simple ratio index	MSR_re	$(\rho_{NIR} - \rho_{RE} - 1)/((\sqrt{(\rho_{NIR} + \rho_{RE})} + 1)$	[35]
Green Normalized Difference Vegetation Index	GNDVI	$(\rho_{NIR} - \rho_G)/(\rho_{NIR} + \rho_G)$	[36]
Simple Ratio NIR/RedEdge Ratio Index	RRI	ρ_{RE}/ρ_{NIR}	[37]
MSAVI-SAVI ratio index	$\frac{MSAVI}{SAVI}$	$((2 \cdot \rho_{NIR} + 1 - ((2 \cdot \rho_{NIR} + 1)^2 - 8 \cdot (\rho_{NIR} - \rho_R))^{\frac{1}{2}})/2)/((1 + 0.16) \cdot (\rho_{NIR} - \rho_R)/(\rho_{NIR} + \rho_R + 0.16))$	[34], [38] Customized
Anthocyanin reflectance index	ARI	$\rho_{NIR} \cdot (1/\rho_G - 1/\rho_R)$	[39]
Structure Insensitive Pigment Index	SIPI	$(\rho_{NIR} - \rho_B)/(\rho_{NIR} - \rho_R)$	[40]

The rest of this paper is organized as follows. The materials collection and research methods are described in Section II. Experimental results are provided in Section III. The experimental results are discussed in Section IV. Finally, we conclude our work in Section V.

II. MATERIALS AND METHODS

A. FIELD PREPARATION

The study was anchored on a 63-acre research field (41°11′25.1091″ N, 82°51′37.9317″ E; elevation: 896 m), located in Gülbargh Town, Shayar County, Xinjiang Uygur Autonomous Region, China. Potassium, phosphorus, and plant growth regulators were not administered in the cotton fields to achieve uniform field management guidelines. Fig.1 shows that cotton was planted at approximately 16,700 seeds per acre on April 19, 2021. Twelve plots (≈ 68 m long × 36 m wide [12 columns]) were arranged in a randomized complete block design (three replicates per treatment) with resident populations of cotton aphids (Aphis gossypii Glover). Cotton aphids are perennial cotton pests in Gülbargh, and their populations develop normally in late May. The number of cotton aphids peaked during late June and early July (Table 2).

B. DATA COLLECTION

1) GROUND INVESTIGATION

The arthropod populations were monitored from May 20 to July 16, 2021. We monitored the cotton aphid population at 288 sample point locations in 12 plots starting with the introduction of cotton aphids into cotton fields, without the application of any chemical insecticides. Under natural conditions, enemies of aphids, such as ladybugs, lacewings, syrphid flies, and suppress aphid reproduction.

Ground-truthing data were collected by sampling aphids at 5-day intervals (Table 2). The GPS locations and aphid populations were recorded for all spots showing symptoms of aphid infection. Five cotton plant sites were randomly selected from each sample point, a sample of 10 leaves (third mainstem node leaf from the top of the plant) was counted, and individuals were counted directly under natural conditions (without destroying the cotton) using 4× magnification. The accumulated aphid days for the period after treatment was estimated using the aphid populations.

2) UAV IMAGERY ACQUISITION

The low-altitude UAV camera system platform consisted of an M300 multi-rotor aircraft (DJI, Shenzhen, China) and a RedEdge-MX airborne multispectral imager (MicaSense, Seattle, WA, USA). The camera had a spatial resolution of 8 cm@120 m altitude hover-scan imaging, and five channels in the 400-900 nm band range were selected: blue (475 nm, 40 nm bandwidth), green (560 nm, 20 nm bandwidth), red (668 nm, 10 nm bandwidth), red edge (717 nm, 10 nm bandwidth), and near-infrared (840 nm, 40 nm bandwidth). The horizontal field of view was 47.2°. Flight data were collected during simultaneous ground collection efforts from June 8 to July 16, 2021. Table 2 describes the cotton seeding date, and the first detection of aphids date, and also records information on the time of the drone data collection on days 50, 66, 73, and 88 from seeding. Flight data were used to correspond to a range of arthropod populations and cotton growth statuses. Images were acquired under clear and cloud-free conditions, and the sensor was calibrated using calibration plates before and after the flight. During the flights, the altitude of the UAV was maintained at 70 m over the ground, with an image resolution of 5 cm. The flight data were preprocessed using the Pix4DMapper software, and the resulting ortho-images were georeferenced and indexed using ArcGIS Pro software before being produced as single-band TIFF grayscale images.
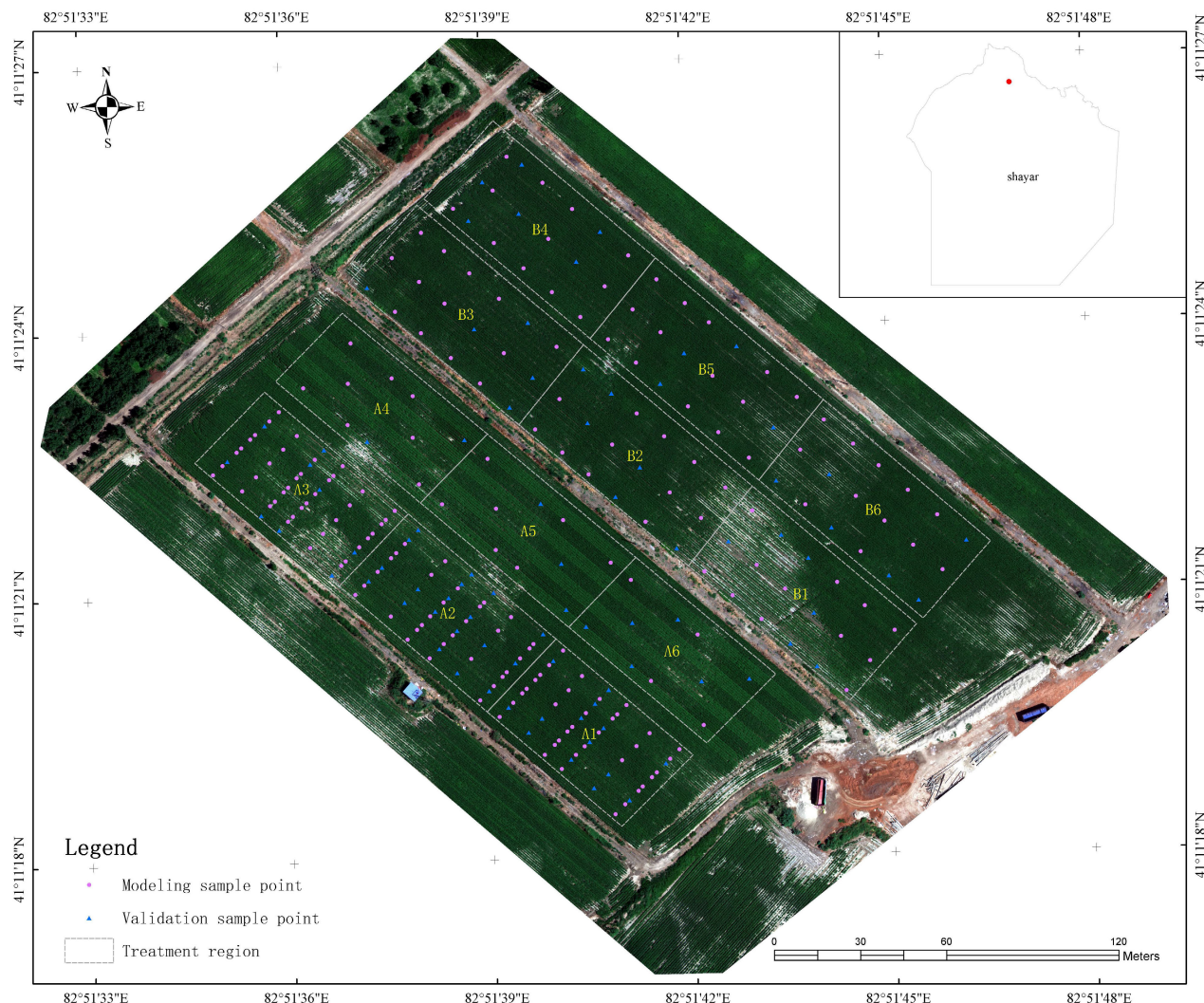
P. Jiang et al.: Prediction Dynamics in Cotton Aphid Using Unmanned Aerial Vehicle Multispectral Images and Vegetation Indices

IEEE *Access*

FIGURE 1. The research area is located in the county of shayar. Note: 70% of the sampled data is utilized for modeling, whereas 30% is employed for validation.

In addition, a linear resampling method was utilized to handle the UAV multispectral and cotton aphid survey data, and the resampled data were used for subsequent analyses.

C. VEGETATION INDEX EXTRACTION

Vegetation indices computed with help of spectrally susceptible bands can facilitate the prediction of crop pests. In this paper, 12 vegetation indices sensitive to cotton aphids were computed as initial features of the model aiming to the application of those indices in cotton aphid prediction on the basis of the features of cotton aphid damage [4]. Table 3 lists the indicator names, formulae, and references.

The reflectances of green (G), blue (B), red (R), near-infrared (NIR), and Red Edge (RE) were chosen to create the vegetation index based on the features of the Red Edge-MX sensor. Because the proposed indices, such as ARI, TCARI, and SAVI, are based on hyperspectral or MERIS remote

sensing data, and the reflectance of the required central wavelengths, such as 700 nm and 800 nm, are not available in UAV multispectral images, the nearest neighboring wavelengths are used to generate the updated indices instead.

D. METHODS

1) THEORETICAL BASIS OF SUPPORT VECTOR REGRESSION

SVR is a support vector machine regression problem extension [41]. Compared with other machine learning approaches, SVR offers the benefits of using fewer samples, obtaining better global optimum solutions, and producing better outcomes when handling multidimensional nonlinear problems [27], [42].

For the nonlinear case, for a given aphid training sample $D = [(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)], x \in \mathbb{R}_n$ are the input parameters and $y \in \mathbb{R}$ are the output parameters. The functional expression of the SVR corresponding to the

optimization problem is described following:

$$\min_{w,b,\xi_n,\xi_n^*} \frac{1}{2}||w||^2 + C \cdot \sum_{i=1}^{m}(\xi_n, \xi_n^*) \tag{1}$$

$$s.t. \begin{cases} f(x_n) - y_n & \leq \epsilon + \xi_n \\ y_n - f(x_n) & \leq \epsilon + \xi_n^* \\ \xi_n, \xi_n^* & \geq 0 \end{cases} \tag{2}$$

where w and b are the weights and biases, respectively, and the penalty factor $C > 0$ and the non-negative relaxation variables ξ_n, ξ_n^*, ϵ are insensitive loss function parameters. To improve the generalization ability using formula (1), and reduce the error using formula (2), by introducing the Lagrange function, we can obtain formula (3) as follows:

$$f(x) = \sum_{n=1}^{m}(a_n - a_n^*)K(x_n, x) + b^* \tag{3}$$

where $K(x_n, x) = \psi(x_n) \cdot \psi(x)$ is the kernel function, $a_n \geq 0$ and $a_n^* \geq 0$ are Lagrangian function multipliers. The goal of the kernel function is to compute the inner product of the original low-dimensional space vector in a high-dimensional space following a particular transformation. The polynomial, radial basis, and sigmoid kernels are examples of common kernel functions.

2) IMPROVED XGBOOST-GWO-SVR MODEL CONSTRUCTION AND EVALUATION

We utilized the SVR method for cotton aphid data with a long-time span, small sample size, remarkable randomness, and nonlinearity, which is more suitable for cotton aphid data analysis and prediction because of its small sample size, high accuracy, high generalization capacity, and good resilience. However, the SVR model is hampered by plenty of features in the dataset, which leads to overfitting or low fitting accuracy. Consequently, we utilized an appropriate optimization approach to address SVR's weaknesses of SVR.

First, the classic SVR technique is not ideal for large datasets because of its poor performance when the number of features per data stage exceeds the number of training data specimens, or when the dataset contains more noise. The natural habitat in which the cotton aphids were found in this study contained a variety of information. If all gathered data are entered into the SVR method, the precision of the model is compromised. Therefore, we employed the XGBoost method for input data screening to minimize the total of features in the training data samples and to increase the accuracy and execution efficiency of the SVR model. XGBoost utilizes a decision-tree-based integration technique that employs a gradient boosting algorithm to minimize the loss of previously created decision trees and generate new trees to construct the model, thereby ensuring the dependability of the final decision [43], [44]. The XGBoost algorithm provides the emphasis score of each feature in each iteration to improve the performance of a training process to generate a new tree,

indicating the significance of each feature in the model being trained and providing a rationale for constructing a new tree by its orientation in the following iteration [45].

Therefore, this paper is inspired by the idea in the previous work with the RBF [46]. However, the RBF penalty coefficient C and tolerance coefficient ϵ can significantly affect the classification accuracy. The GWO method is a cluster intelligence optimization approach that creates optimum solutions without the need for any input parameters by utilizing multiple built-in functions [47]. The optimal penalty coefficient C and tolerance coefficient ϵ obtained from the GWO algorithm were substituted into the SVR for the construction of the XGBoost-GWO-SVR prediction model.

The fundamental idea of the cotton aphid prediction model, XGBoost, is utilized as a precursor data processing system for SVR, and the target dataset (UAV multispectral reflectance data, vegetation index data, and cotton aphid survey data) was screened for features, and the feature data were normalized by feature screening. Second, the SVR penalty coefficient C and tolerance coefficient ϵ are adjusted using GWO, and the tuning parameters are fed into the SVR algorithm. Fig.2 shows the construction process of The XGBoost-GWO-SVR prediction model.

To better evaluate the advantages of the XGBoost-GWO-SVR algorithm in the prediction model, two classical algorithms (LR and BPNN) and four SVR algorithms (SVR, XGBoost-GA-SVR, XGBoost-PSO-SVR, and XGBoost-SVR) were compared with the XGBoost-GWO-SVR algorithm. Seven prediction models for cotton aphids were constructed by combining two sets of features with seven combinations of the methods. A fold cross-validation method is performed to assess the accuracy and stability of the predictive models. In this methodology, the cotton aphid samples are subdivided into 10 subsets of similar size that are repelled by each other. Three data subsets are assigned as the test set and the remaining seven subsets are dedicated as the training set. Finally, the prediction model's reliability was assessed by using MSE, MAE, and coefficient of determination (R^2) of prediction results [48].

$$MSE = \frac{1}{n} \cdot \sum_{i=1}^{n}[y_i - g(x_i)]^2 \tag{4}$$

$$MAE = \frac{1}{n} \cdot |y_i - g(x_i)| \tag{5}$$

$$R^2 = 1 - \frac{\sum[y_i - g(x_i0]^2}{\sum[y_i - \bar{y}]^2} \tag{6}$$

where y_i is the actual value, $g(x_i)$ is the predicted value, $\bar{y}$ is the mean value, and n is the number of samples.

According to the preceding equation, the MSE represents the average square of the difference between the anticipated and actual values. The MAE represents the real situation of the inaccuracy of the predicted values, and R^2 is the criterion for assessing the goodness of the model.
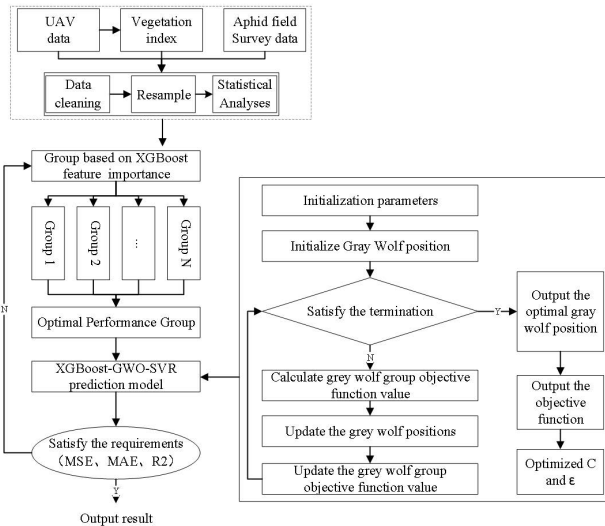
P. Jiang et al.: Prediction Dynamics in Cotton Aphid Using Unmanned Aerial Vehicle Multispectral Images and Vegetation Indices

IEEE *Access*

FIGURE 2. Workflow diagram of cotton aphid prediction model based on XGBoost-GWO-SVR.

III. RESULTS

A. OPTIMAL FEATURE SELECTION AND ANALYSIS

Previous research depended on the relevance of vegetation indicators in identifying the presence of vegetation pests, Fig.3 illustrates the determination squares of the correlation coefficients of vegetation index and multispectral reflectance data with cotton aphids. To establish confidence levels (p) for the healthy and sick subjects in the 12 vegetation indicators and five band reflectances, a t-test for each sample and two-way ANOVA were utilized. Cotton aphids exhibited a considerable potential in the p corresponding to each of the vegetative indicators throughout the figure ($p < 0.001$). The data in Fig.3 show that the R_Blue band has the lowest coefficient of determination value, indicating the significant relationship between R_Blue and the prevalence of cotton aphids, and the vegetation indices GNDVI and ARI have the ability to indicate aphid occurrence. Relying on the initial values of vegetation indices and UAV multispectral reflectance data, 12 vegetation indices and five reflectance data were discovered to better reflect cotton aphid pest incidence ($p < 0.001$), and these indicators were utilized as the shortlisting factors for later studies.

As shown in Fig.4, aphid data, UAV multispectral reflectance data, and vegetation indices have a random distribution; therefore, we employ the XGBoost method to compute data error values in order to categorize the relevance of the original feature variables. The better the element priority level, the further significant the feature [49]. Fig.4 shows that different indicators define aphids differently, with ARI having the most fantastic relevance score among the 19 indicators, owing to the aphids' ability to drain nutrients from leaves, resulting in unusually high anthocyanin concentrations [39]. Considering that selecting too many or too few features affects the model's accuracy and stability, we test various values to figure out the optimal number of features
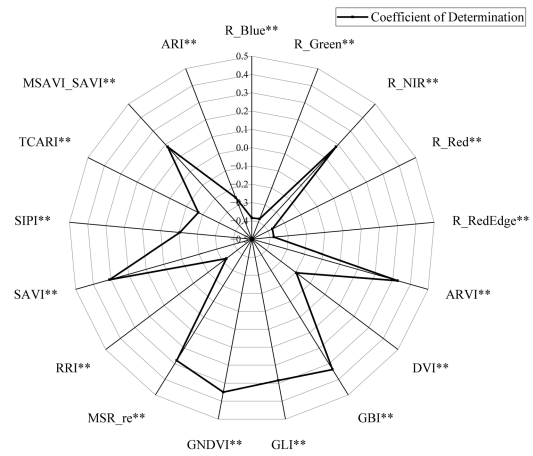

FIGURE 3. Determination coefficients of selected vegetation indices and cotton aphid. Note: ** indicates confidence levels $p < 0.001$.

(N) based on the results of the information gain listing, and we decide on the final selected value through comparing the consequences of various attributes on the model's MAE value. Table 4 illustrates the results of assessing the importance of the features and combining them with MAE as the evaluation index, demonstrating that the influence of MAE worsens when $N > 12$. This study determines that the best combinations of environmental variables were ARI, GBI, GLI, TCARI, R_NIR, RRI, MSR_re, MSAVI_SAVI, ARVI, GNDVI, and SAVI.
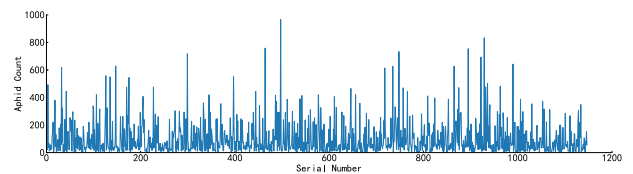

FIGURE 4. Depicts the random distribution of aphid data.

B. MODEL ANALYSIS AND VALIDATION

Table 5 exhibits the findings of each model, including the MSE and MAE of various types of cotton aphid prediction models that combine the feature selection methods. According to the statistics in Table 5, XGBoost-SVR reduces MSE and MAE by 0.078% and 0.35%, respectively, when compared to a single SVR, and the model's efficiency is improved by reducing the number of features from 20 to 12. It can be observed that XGBoost can increase the efficiency of the algorithm by lowering the quantity of data supplied to the SVR to a certain level, and it can lessen the influence of irrelevant data on SVR accuracy.

Although XGBoost can improve the SVR efficiency to some extent, it cannot address the issue of insufficient SVR accuracy, which requires an optimization technique. The GWO technique is a heuristic optimization algorithm inspired by the prey hunting behavior of grey wolves, with a good

IEEE *Access*

P. Jiang et al.: Prediction Dynamics in Cotton Aphid Using Unmanned Aerial Vehicle Multispectral Images and Vegetation Indices

TABLE 4. Topn feature MSE.

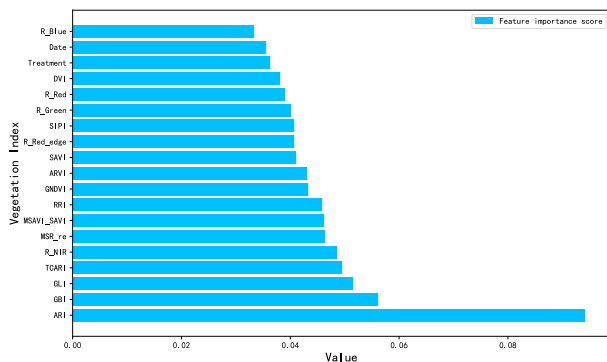| TOPN FEATURE	THRESH	MSE
1	0.130	9.858
2	0.094	9.734
3	0.056	9.680
4	0.051	9.633
5	0.049	9.605
6	0.048	9.589
7	0.046	9.576
8	0.045	9.567
9	0.043	9.558
10	0.043	9.552
11	0.041	9.546
12	0.040	9.542
13	0.040	9.543
14	0.040	9.546
15	0.039	9.553
16	0.038	9.562
17	0.036	9.565
18	0.035	9.566
19	0.035	9.570
20	0.030	9.576

FIGURE 5. XGBoost feature importance score result.

TABLE 5. The effect of XGBoost feature filtering on SVR model accuacy.

MODELS	MSE	MAE
SVR	3025.329	9.576
XGBoost	3022.943	9.542

TABLE 6. Implementation results of XGBoost-GWO-SVR.

Items	Description
Number of positions of grey wolfs	10
Number of initial populations for GA	10
Number of populations for particle	10
Cross over ratio	Dynamic
Mutation ratio	Dynamic
Number of GWO iterations	50
Number of GA iterations	50
Number of PSO iterations	50

TABLE 7. Results of optimum scale values using XGBoost-GWO-SVR.

Type of model	C	ϵ
XGBoost-GA-SVR	780.63160533	1.29867322
XGBoost-PSO-SVR	606.33847888	1.23270312
XGBoost-GWO-SVR	408.87708232	1.04715462

convergence capability, limited parameters, and a simple implementation [47]. The GWO method was utilized to optimize the penalty coefficient C and tolerance coefficient ϵ of SVR to increase the prediction accuracy and prediction speed of SVR by selecting the global optimal parameters for a total of 50 rounds with 10 rounds each; the addressing ranges of the penalty coefficient C and tolerance coefficient ϵ are (0,1000] and (0,10], respectively. The implementation results of XGBoost-GWO-SVR are given in Table 6. The optimum scale values using XGBoost-GWO-SVR are listed in Table 7. The above-mentioned data analysis and algorithm programming were conducted in Python 3.7.

Table 8 displays the outcomes of each model, including the performance assessment indices for the proposed XGBoost-GWO-SVR model and the six comparator models. The cotton aphid prediction model employing the XGBoost algorithm

in conjunction with the GWO algorithm combined with the SVR method had the lowest MSE and MAE as well as the highest R^2. The findings revealed that the XGBoost screening features as input variables paired with the GWO-SVR method could successfully forecast the incidence of cotton aphids; however, the SVR prediction model had the lowest accuracy and stability, and the MAE of BPNN was the poorest. When compared to the LR method, the combined XGBoost-GWO-SVR prediction model exhibits a 12.06% and 58.93% reduction in MSE and MAE, respectively, and improved R^2 by 0.3%. Compared with the BPNN model, the MSE and MAE were reduced by 84.76% and 89.21%, respectively, and R^2 was increased by 12.5%. Moreover, for the same variables, the XGBoost-GWO-SVR model had the best accuracy and correctness, the LR model had the second highest accuracy and correctness, and the BPNN model had the lowest accuracy and correctness. Therefore, the strategy used to construct a prediction model has a substantial impact on the accuracy of cotton aphid prediction. In comparison to LR and BPNN, XGBoost-GWO-SVR is based on a radial basis kernel function that addresses the problems of small sample size, randomness, and nonlinearity in cotton aphid monitoring. Furthermore, the accuracy of the prediction model was significantly improved by optimizing the penalty coefficient C and tolerance coefficient ε using the GWO algorithm. Overall, the cotton aphid prediction models developed using the LR, BPNN, SVR, XGBoost-SVR, XGBoost-PSO-SVR, and XGBoost-GA-SVR algorithms outperformed the XGBoost-GWO-SVR model.

Furthermore, the performance of GWO, GA, and PSO in optimizing the cotton aphid prediction SVR algorithm after XGBoost-filtered features were compared in this work. The model accuracy and efficiency of the three optimization algorithms are significantly improved when compared with those of the SVR and XGBoost-SVR algorithms. The R^2 of the models obtained by all three optimization algorithms reached 0.980, while the MSE and MAE of the models were significantly reduced, indicating that the optimization algorithms
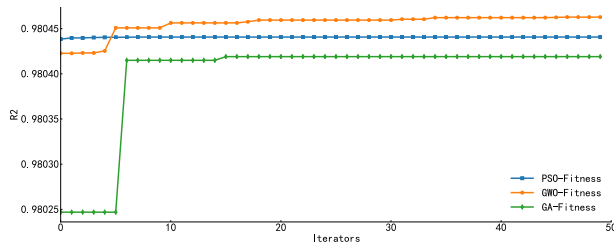
P. Jiang et al.: Prediction Dynamics in Cotton Aphid Using Unmanned Aerial Vehicle Multispectral Images and Vegetation Indices

IEEE Access

FIGURE 6. Effect of different optimization algorithms on SVR model performance.

TABLE 8. Table comparing the performance of several cotton aphid prediction models.

MODELS	MSE	MAE	R^2
LR	337.248	6.913	0.977
SVR	3025.329	9.576	0.800
BPNN	1946.838	26.323	0.871
XGBoost-SVR	3022.943	9.542	0.800
XGBoost-GWO-SVR	296.567	2.838	0.980
XGBoost-GA-SVR	296.982	3.067	0.980
XGBoost-PSO-SVR	296.899	3.013	0.980

can improve the model accuracy of SVR. Compared to the XGBoost-SVR method, the MSE and MAE of the XGBoost-GWO-SVR model decreased by 90.18% and 70.25%, respectively, while its R^2 was improved by 22.5%. Table 8 shows that the MSE and MAE of the GWO method outperformed the other two optimization procedures, indicating that the GWO algorithm was more suitable for optimizing the SVR cotton aphid prediction model than the other two optimization strategies. Meanwhile, Fig.6 illustrates that the GA is less stable than the GWO and PSO methods.

The combined aphid population prediction model of XGBoost-GWO-SVR is compared with LR, BPNN, SVR, XGBoost-SVR, XGBoost-GA-SVR, and XGBoost-PSO-SVR to confirm the performance of the model described in this study. A collapsed cross-validation procedure was utilized in this study to evaluate the precision and robustness of the prediction models. Cotton aphid samples were split into ten mutually exclusive subgroups of identical size using this procedure. A training set of seven subsets of data and a test set of three subsets were employed. Finally, the cotton aphid prediction model's performance has been evaluated utilizing the results of MSE, MAE, and R^2. Fig.7 illustrates the results of the experiments; the accuracies of the XGBoost-GWO-SVR model outperforms that of the LR and BPNN strategies, while SVR and XGBoost-SVR had the lowest accuracy, and the curve of the XGBoost-GWO-SVR model was closest to the true value. As a result, the XGBoost-GWO-SVR model has an excellent prediction effect and can reliably estimate the aphid population based on the given parameters.

In conclusion, we discovered that the enhanced XGBoost-GWO-SVR approach may overcome the shortcomings of the standard support vector regression methodology, such as overfitting, low accuracy, and slow response.

The experimental findings demonstrate that the 11 index factors screened by XGBoost increase SVR efficiency to some amount, but cannot solve the problem of insufficient SVR accuracy. As a result, optimization procedures are required to increase the accuracy and resilience of the SVR model. The GWO method has the best stability among the three optimization algorithms utilized in this study, whereas the GA algorithm has the poorest. When compared to previous prediction models, the XGBoost-GWO-SVR model has excellent stability and high accuracy, and it can predict the dynamic distribution of cotton aphid populations in near real-time using UAV multispectral data.

IV. DISCUSSION

The goal of this research is to investigate a generic technique for cotton aphid infestation prediction that is independent of ground surveys to take advantage of the benefits of remote sensing in monitoring broad regions while also reducing human and financial investment in ground surveys. In this study, several improved vegetation indices are proposed based on UAV multispectral data and the characteristics of cotton aphids. First of all, suitable remote sensing index feature factors are screened using feature screening methods, the SVR model is algorithmically optimized, and finally, the XGBoost-GWO-SVR model applicable to cotton aphid prediction is obtained. The experimental results show that the 11 index factors screened by XGBoost improve the efficiency of SVR to a certain extent, but they cannot solve the problem of insufficient SVR accuracy. Optimization algorithms can significantly increase the accuracy and robustness of the SVR model, and among the three optimization algorithms used in this study, the stability of GWO and PSO is approximated. However, the stability of the GA model is weaker. Compared with the LR and BPNN methods, the proposed XGBoost-GWO-SVR model has strong stability and high accuracy; thus, it is more suitable for prediction analysis of small samples and nonlinear data, such as cotton aphid infestation. The cotton aphid prediction model XGBoost-GWO-SVR, which combines the benefits of the XGBoost, GWO, and SVR algorithms, has strong cotton aphid prediction ability and may be used in cotton field production management.

The vast majority of vegetation indices employed in this study are broadband vegetation indices, which are less constrained by atmospheric conditions and loads than narrowband vegetation indices, and may be acquired from multispectral aerial or satellite imagery. However, the small-scale UAV multispectral data utilized in this study only include sensors from visible to NIR range of wavelength, with the SWIR band omitted, making identifying water-stress signs in cotton problematic. Vegetation indices, including those in the short-wave infrared band, could be considered in future studies to obtain better results. Furthermore, the adaptability of large-scale remote sensing data and its indexes in the XGBoost-GWO-SVR model must be investigated in future research.
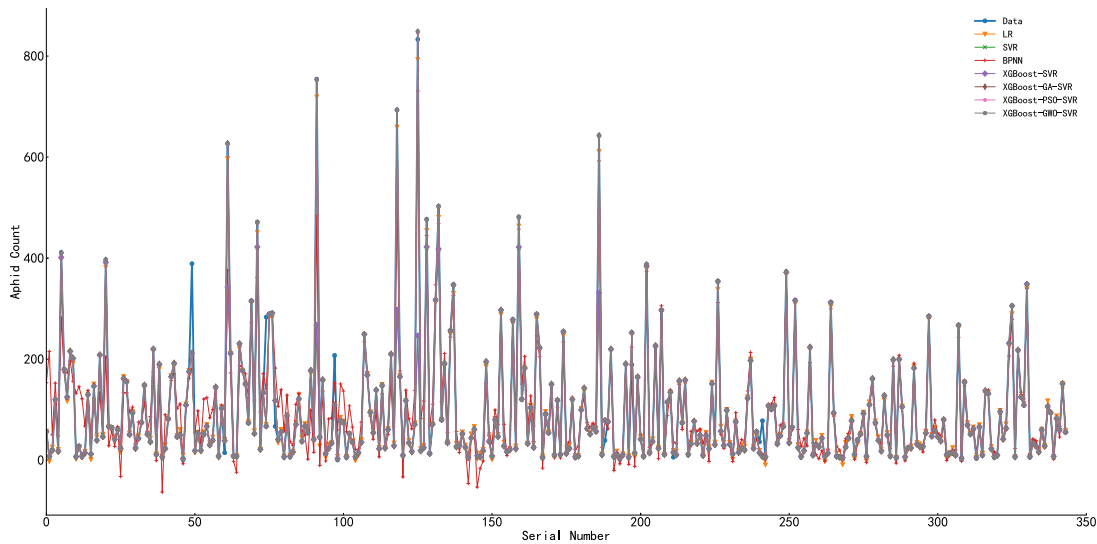
IEEE *Access*

P. Jiang et al.: Prediction Dynamics in Cotton Aphid Using Unmanned Aerial Vehicle Multispectral Images and Vegetation Indices

FIGURE 7. Comparison of the different cotton aphid prediction models.

V. CONCLUSION

In this work, the XGBoost-GWO-SVR model on the basis of UAV multispectral data was implemented to predict the occurrence level of cotton aphids. The model was combined with GWO and SVR algorithms to develop a predictive model via feature selection of vegetation indicators employing the XGBoost feature selection algorithm. We can obtain the following conclusions.

1) The precision of the model obtained by the XGBoost features extraction method outperforms the models without the selection algorithm, demonstrating that the combination of XGBoost algorithms can provide better prediction of pests.

2) XGBoost was employed as the method of selecting features, the prediction accuracies of the GWO-SVR algorithm and the MSE, MAE, and R^2 of cotton aphids reached 196.567, 2.838, and 0.980, respectively, which are better than those of SVR, LR, BPNN, XGBoost-SVR, XGBoost-GA-SVR, and XGBoost-PSO-SVR for predicting cotton aphids. This observation shows that the combination of the GWO-SVR model and XGBoost algorithm can enhance the prediction accuracy of cotton aphids effectively, and it can provide a reference of methodology and technology for the treatment of early cotton aphids.

Furthermore, when comparing the three distinct SVR model optimization approaches, the GWO algorithm-optimized prediction model has the highest accuracy, and the GA algorithm has the minimum accuracy. The findings show that the XGBoost-GWO-SVR model has excellent accuracy and performance, which can be useful in aphid prediction situations that require high precision and efficiency, as well as in cotton pest treatment.

ACKNOWLEDGMENT

REFERENCES

[1] J.-J. Cui, H.-Y. Chen, X.-H. Zhao, and J.-Y. Luo, "Research course of the cotton IPM and its prospect," *Cotton Sci.*, vol. 19, no. 5, pp. 385–390, 2007.

[2] R. Song, Q. He, and X. Liu, "Characteristics of major pests and diseases in cotton cultivation in Xinjiang and control techniques," *Appl. Eng. Technol.*, vol. 40, no. 26, pp. 44–45, Sep. 2020.

[3] Y. Lu, J. Li, X. Wei, H. Wang, and P. Wang, "A technical regulation for integrated control of cotton aphids in Xinjiang," *China Cotton*, vol. 49, no. 1, pp. 38–41, Jan. 2022.

[4] H. Zhao, C. Yang, W. Guo, L. Zhang, and D. Zhang, "Automatic estimation of crop disease severity levels based on vegetation index normalization," *Remote Sens.*, vol. 12, no. 12, p. 1930, Jun. 2020.

[5] N. Peralta, Y. Assefa, J. Du, C. Barden, and I. Ciampitti, "Mid-season high-resolution satellite imagery for forecasting site-specific corn yield," *Remote Sens.*, vol. 8, no. 10, p. 848, Oct. 2016.

[6] J. Su, C. Liu, M. Coombes, X. Hu, C. Wang, X. Xu, Q. Li, L. Guo, and W.-H. Chen, "Wheat yellow rust monitoring by learning from multispectral UAV aerial imagery," *Comput. Electron. Agricult.*, vol. 155, pp. 157–166, Dec. 2018.

[7] P. Chen, W. Xu, Y. Zhan, W. Yang, J. Wang, and Y. Lan, "Evaluation of cotton defoliation rate and establishment of spray prescription map using remote sensing imagery," *Remote Sens.*, vol. 14, no. 17, p. 4206, Aug. 2022.

[8] Y. Yan, C.-C. Feng, and K. Chang, "Towards enhancing integrated pest management based on volunteered geographic information," *ISPRS Int. J. Geo-Inf.*, vol. 6, no. 7, p. 224, Jul. 2017.

[9] F. Vanegas, D. Bratanov, K. Powell, J. Weiss, and F. Gonzalez, "A novel methodology for improving plant pest surveillance in vineyards and crops using UAV-based hyperspectral and spatial data," *Sensors*, vol. 18, no. 1, p. 260, Jan. 2018.

[10] H. Wenjiang, S. Yue, D. Yingying, Y. Huichun, W. Mingquan, C. Bei, and L. Linyi, "Progress and prospects of crop diseases and pests monitoring by remote sensing," *Smart Agricult.*, vol. 1, no. 4, pp. 1–11, Jul. 2019.

[11] E. R. Hunt and S. I. Rondon, "Detection of potato beetle damage using remote sensing from small unmanned aircraft systems," *J. Appl. Remote Sens.*, vol. 11, no. 2, p. 1, May 2017.

[12] Z. A. Pretorius, C. X. Lan, R. Prins, V. Knight, N. W. McLaren, R. P. Singh, C. M. Bender, and F. J. Kloppers, "Application of remote sensing to identify adult plant resistance loci to stripe rust in two bread wheat mapping populations," *Precis. Agricult.*, vol. 18, no. 4, pp. 411–428, Aug. 2017.

[13] M. P. Pound, J. A. Atkinson, A. J. Townsend, M. H. Wilson, M. Griffiths, A. S. Jackson, A. Bulat, G. Tzimiropoulos, D. M. Wells, E. H. Murchie, T. P. Pridmore, and A. P. French, "Deep machine learning provides state-of-the-art performance in image-based plant phenotyping," *GigaScience*, vol. 6, no. 10, pp. 1–10, Oct. 2017.

P. Jiang et al.: Prediction Dynamics in Cotton Aphid Using Unmanned Aerial Vehicle Multispectral Images and Vegetation Indices

IEEE *Access*

[14] S. Aich and I. Stavness, "Leaf counting with deep convolutional and deconvolutional networks," in *Proc. IEEE Int. Conf. Comput. Vis. Workshops (ICCVW)*, Oct. 2017, pp. 2080–2089.

[15] M. Zhu, S. Liu, Z. Xia, G. Wang, Y. Hu, and Z. Liu, "Crop growth stage GPP-driven spectral model for evaluation of cultivated land quality using GA-BPNN," *Agriculture*, vol. 10, no. 8, p. 318, Aug. 2020.

[16] X. Wang, C. Zhang, S. Zhang, and Y. Zhu, "Forecasting of cotton diseases and pests based on adaptive discriminant deep belief network," *Trans. Chin. Soc. Agric. Eng.*, vol. 34, no. 14, pp. 157–164, Jan. 2018.

[17] C. Zhao, J. Zhang, N. Su, and Y. Yan, "Multipath convolutional neural network algorithm for hyperspectral change detection," *J. Harbin Eng. Univ.*, vol. 41, no. 9, pp. 1398–1404, Jan. 2018.

[18] S. Ren, J. Pan, X. Li, G. Xu, and G. Gong, "Novel machinery fault diagnosis approach via ELMD and improved SMSVM," *J. Nanjing Univ. Aeronaut. Astronaut.*, vol. 51, no. 5, pp. 693–703, Jan. 2019.

[19] Y. Miao, R. Zhang, J. Guo, S. Yi, B. Meng, and J. Liu, "Vegetation coverage in the desert area of the Junggar basin of Xinjiang, China, based on unmanned aerial vehicle technology and multisource data," *Remote Sens.*, vol. 14, no. 20, p. 5146, Oct. 2022.

[20] I. J. Marang, P. Filippi, T. B. Weaver, B. J. Evans, B. M. Whelan, T. F. A. Bishop, M. O. F. Murad, D. Al-Shammari, and G. Roth, "Machine learning optimised hyperspectral remote sensing retrieves cotton nitrogen status," *Remote Sens.*, vol. 13, no. 8, p. 1428, Apr. 2021.

[21] Z. Xue, Y. Zhang, C. Cheng, and G. Ma, "Remaining useful life prediction of lithium-ion batteries with adaptive unscented Kalman filter and optimized support vector regression," *Neurocomputing*, vol. 376, pp. 95–102, Feb. 2020.

[22] S. Xu, S. Li, Z. Tao, K. Song, Z. Wen, Y. Li, and F. Chen, "Remote sensing of chlorophyll-A in Xinkai lake using machine learning and GF-6 WFV images," *Remote Sens.*, vol. 14, no. 20, p. 5136, Oct. 2022.

[23] A. Mathew, S. Sreekumar, S. Khandelwal, and R. Kumar, "Prediction of land surface temperatures for surface urban heat island assessment over Chandigarh city using support vector regression model," *Sol. Energy*, vol. 186, pp. 404–415, Jul. 2019.

[24] J. Zhao, C. Zhang, L. Min, Z. Guo, and N. Li, "Retrieval of farmland surface soil moisture based on feature optimization and machine learning," *Remote Sens.*, vol. 14, no. 20, p. 5102, Oct. 2022.

[25] Z. Yang, Y. Wang, and C. Kong, "Remaining useful life prediction of lithium-ion batteries based on a mixture of ensemble empirical mode decomposition and GWO-SVR model," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–11, 2021.

[26] B. Costello, O. O. Osunkoya, J. Sandino, W. Marinic, P. Trotter, B. Shi, F. Gonzalez, and K. Dhileepan, "Detection of parthenium weed (*Parthenium hysterophorus* L.) and its growth stages using artificial intelligence," *Agriculture*, vol. 12, no. 11, p. 1838, Nov. 2022.

[27] Y. Meng, X. Zhang, Z. Liu, and X. Wang, "Support vector regression ship motion identification modeling based on grey wolf optimizer," in *Proc. 40th Chin. Control Conf. (CCC)*, Jul. 2021, pp. 1172–1176.

[28] A. Alzaqebah, I. Aljarah, O. Al-Kadi, and R. Damaševičius, "A modified grey wolf optimization algorithm for an intrusion detection system," *Mathematics*, vol. 10, no. 6, p. 999, Mar. 2022.

[29] M. S. Shaikh, C. Hua, M. A. Jatoi, M. M. Ansari, and A. A. Qader, "Application of grey wolf optimisation algorithm in parameter calculation of overhead transmission line system," *IET Sci., Meas. Technol.*, vol. 15, no. 2, pp. 218–231, Mar. 2021.

[30] A. Huete, C. Justice, and H. Liu, "Development of vegetation and soil indices for MODIS-EOS," *Remote Sens. Environ.*, vol. 49, no. 3, pp. 224–234, Sep. 1994.

[31] H. Santhanam and R. Majumdar, "Quantification of green-blue ratios, impervious surface area and pace of urbanisation for sustainable management of urban lake—Land zones in India—A case study from Bengaluru city," *J. Urban Manage.*, vol. 11, no. 3, pp. 310–320, Sep. 2022, doi: 10.1016/j.jum.2022.03.001.

[32] J.-L. Widlowski, M. M. Verstraete, B. Pinty, and N. Gobron, "Advanced vegetation indices optimized for up-coming sensors: Design, performance, and applications," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 6, Nov. 2000.

[33] D. Haboudane, J. R. Miller, N. Tremblay, P. J. Zarco-Tejada, and L. Dextraze, "Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture," *Remote Sens. Environ.*, vol. 81, nos. 2–3, pp. 416–426, Aug. 2002.

[34] A. R. Huete, "A soil-adjusted vegetation index (SAVI)," *Remote Sens. Environ.*, vol. 25, no. 3, pp. 295–309, Aug. 1988.

[35] C. Wu, Z. Niu, Q. Tang, and W. Huang, "Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation," *Agricult. Forest Meteorol.*, vol. 148, nos. 8–9, pp. 1230–1241, Jul. 2008.

[36] A. A. Gitelson, Y. J. Kaufman, and M. N. Merzlyak, "Use of a green channel in remote sensing of global vegetation from EOS-MODIS," *Remote Sens. Environ.*, vol. 58, no. 3, pp. 289–298, Dec. 1996.

[37] D. A. Sims and J. A. Gamon, "Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages," *Remote Sens. Environ.*, vol. 81, nos. 2–3, pp. 337–354, Aug. 2002.

[38] J. Qi, A. Chehbouni, A. R. Huete, Y. H. Kerr, and S. Sorooshian, "A modified soil adjusted vegetation index," *Remote Sens. Environ.*, vol. 48, no. 2, pp. 119–126, May 1994.

[39] A. A. Gitelson, G. P. Keydan, and M. N. Merzlyak, "Three-band model for noninvasive estimation of chlorophyll, carotenoids, and anthocyanin contents in higher plant leaves," *Geophys. Res. Lett.*, vol. 33, no. 11, Jun. 2006, Art. no. L11402.

[40] G. A. Blackburn, "Spectral indices for estimating photosynthetic pigment concentrations: A test using senescent tree leaves," *Int. J. Remote Sens.*, vol. 19, no. 4, pp. 657–675, Jan. 1998.

[41] V. Vapnik, S. E. Golowich, and A. Smola, "Support vector method for function approximation, regression estimation and signal processing," in *Proc. 9th Int. Conf. Neural Inf. Process. Syst.*, 1996, pp. 281–287.

[42] J. Cai, J. King, C. Yu, J. Liu, and L. Sun, "Support vector regression-based behavioral modeling technique for RF power transistors," *IEEE Microw. Compon. Lett.*, vol. 28, no. 5, pp. 428–430, May 2018.

[43] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2016, pp. 785–794.

[44] Y. Jiang, G. Tong, H. Yin, and N. Xiong, "A pedestrian detection method based on genetic algorithm for optimize XGBoost training parameters," *IEEE Access*, vol. 7, pp. 118310–118321, 2019.

[45] X. Zhao, N. Xia, Y. Xu, X. Huang, and M. Li, "Mapping population distribution based on XGBoost using multisource data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 11567–11580, 2021.

[46] L. Huang, K. Wu, W. Huang, Y. Dong, H. Ma, Y. Liu, and L. Liu, "Detection of fusarium head blight in wheat ears using continuous wavelet analysis and PSO-SVM," *Agriculture*, vol. 11, no. 10, p. 998, Oct. 2021.

[47] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Adv. Eng. Softw.*, vol. 69, pp. 46–61, Mar. 2014.

[48] D. J. Brus, "Sampling for digital soil mapping: A tutorial supported by R scripts," *Geoderma*, vol. 338, pp. 464–480, Mar. 2019.

[49] L. Huang, Y. Liu, W. Huang, Y. Dong, H. Ma, K. Wu, and A. Guo, "Combining random forest and XGBoost methods in detecting early and mid-term winter wheat stripe rust using canopy level hyperspectral measurements," *Agriculture*, vol. 12, no. 1, p. 74, Jan. 2022.

PINGAN JIANG was born in Sichuan, China, in 1965. He received the Ph.D. degree in soil science from the College of Resource and Environment, China Agricultural University, Beijing, China. His research interests include precision agriculture technology applied research and demonstration, large-scale contemporary agricultural digital technologies applied research, cotton pest monitoring, and early warning information systems.

XUELIN ZHOU (Graduate Student Member, IEEE) was born in Anqing, Anhui, China, in 1986. He received the M.Sc. degree from the Jiangxi University of Science and Technology, in 2012. He is currently pursuing the Ph.D. degree in agronomy with Xinjiang Agricultural University. He is working in the field of agricultural remote sensing and pest and disease remote sensing monitoring for a long time. His research interests include cotton pest and disease remote sensing monitoring technology and vegetation chemosensory ecological effects.

IEEE *Access*

P. Jiang et al.: Prediction Dynamics in Cotton Aphid Using Unmanned Aerial Vehicle Multispectral Images and Vegetation Indices

TONGLAI LIU was born in 1982. He received the B.E. and M.E. degrees from the Guilin University of Electronic Technology, China, in 2007 and 2010, respectively, and the Ph.D. degree from the Guangdong University of Technology, in 2021. He joined at the Guilin University of Electronic Technology. He is currently an Associate Professor with the Zhongkai University of Agriculture and Engineering. His current research interests include data mining, blockchain technology, edge computing, and database.

CONG ZHANG was born in Zhangjiakou, Hebei, China, in 1995. He received the Bachelor of Engineering degree in information management and information systems from the Taiyuan University of Science and Technology in 2018. He is currently pursuing the master's degree in agricultural resources and environment with Xinjiang Agricultural University. His research interests include agricultural ecology, environmental protection, and remote sensing.

XIAOHU GUO was born in Fukang, Xinjiang, China, in 1997. He received the Bachelor of Agriculture degree in forest protection from Xinjiang Agricultural University in 2019, where he is currently pursuing the Ph.D. degree in crop science. His research interest includes integrated pest management for cotton.

YAN LI was born in Liulang, Yunnan, China, in 1987. She received the M.Sc. degree from Guangzhou University, in 2011. She is currently pursuing the doctor's degree with the Kunming University of Science and Technology. Her main research interests include photogrammetry and remote sensing.

DEYING MA was born in Xinjiang, China, in 1968. She received the Ph.D. degree in pesticide science from the College of Plant Protection, Shandong Agricultural University, in 2006. Her main research interests include the occurrence of agricultural and forestry pests, the conservation and use of natural enemies, the research and application of environmentally friendly pesticides, and the management of pest resistance. She is a member of the Weed Branch of the Chinese Society of Plant Protection, a Council Member of the Xinjiang Agricultural Society, a Council Member of the Xinjiang Society of Plant Protection, a Reviewer for Insect Science, a Senior Member of the Royal Entomological Society, and a member of the International (European) Network for the Study of Sooty Mites. She is a member of the editorial board of the *Journal of Biosafety and Biosecurity*.

SHUANGYIN LIU received the Ph.D. degree from the College of Information and Electrical Engineering, China Agricultural University, in 2014. He is currently a Professor with the College of Information Science and Technology, Zhongkai University of Agriculture and Engineering. His current research interests include intelligent information systems for agriculture, artificial intelligence, big data, software engineering, and computational intelligence.

• • •