Development of a Multiscale XGBoost-based Model for Enhanced Detection of Potato Late Blight Using Sentinel-2, UAV, and Ground Data

Sheng Chang[1*], Zelong Chi[2], Hong Chen[3] and Tongle Hu[4]

[1]Key Laboratory of Remote Sensing and Digital Earth, Aerospace Information Research Institute, Chinese Academy of Sciences (AIRCAS), Beijing 100101, China

[2]Key Laboratory of Biodiversity and Environment on the Qinghai-Tibetan Plateau, Ministry of Education, Tibet University, Lhasa,850000, China

[3]China Aero Geophysical Survey and Remote Sensing Center for Nature Resources, Beijing, 100083, China

[4]College of Plant Protection, Hebei Agriculture University, Baoding 070001, China

*changsheng@aircas.ac.cn

Abstract: *Potatoes, a crucial staple crop, face significant threats from late blight, which pose serious risks to food security. Despite extensive research using ground and UAV hyperspectral data for crop disease monitoring, satellite-scale identification of diseases like Potato Late Blight (PLB) is limited. This study integrates Sentinel-2 data with UAV and ground spectral data for a multi-scale monitoring of PLB. The research found consistent spectral patterns across scales, with notable valley values in Blue and Red bands and peaks in Near Infrared bands, showing a decrease in reflectance with increasing disease severity. Furthermore, the study highlights scale-dependent spectral variations, significant differences in actual reflectance values were observed. Based on the developed Red Edge Index and Disease Stress Index with a suite of machine learning algorithms, we proposed a XGBoost-based model integrating spectral indices for PLB monitoring(PLB-SI-XGBoost). Notably, the proposed model demonstrated the highest average evaluation score of 0.88 and the lowest root mean square error (RMSE) of 13.50 during ground scale validation, outperforming other algorithms. At the UAV scale, the proposed model achieved a robust R-squared value of 0.74 and an RMSE of 18.27. Moreover, the application of Sentinel-2 data for disease detection at satellite scale yielded an accuracy of 70% in the model. The results of the study emphasize the importance of scale in disease monitoring models and illuminate the potential for satellite-scale surveillance of PLB. The exceptional performance of PLB-SI-XGBoost model in detecting PLB suggests its utility in enhancing agricultural decision-making with more accurate and reliable data support.*

Keywords: *Multiscale remote sensing, Potato late blight, XGBoost, Machine learning*

1. Introduction

In 2016, China introduced the "Potato Staple Food Strategy", aiming to industrialize and elevate the status of potatoes within the national diet. This initiative anticipates that half of the projected increase in China's grain production over the next two decades will be attributed to potatoes (B. Liu et al., 2021) . Potatoes play a crucial role in agricultural innovation and the shift towards sustainable Agriculture 4.0 practices. However, historical devastation such as the "Irish Potato Famine" in the 1840s and a severe potato blight in China's Sichuan and Chongqing regions in the 1940s, which led to an 80% reduction in yields, underscore the

significant threat posed by potato late blight (PLB). Recognized as one of the most severe plant diseases globally, the threat of PLB remains substantial (C. Gao et al., 2020). Traditional methods for monitoring PLB rely on field inspections conducted by agricultural practitioners or technicians.These methods labor-intensive and susceptible to misdiagnosis. The advent of remote sensing technology in the early 1930s (F. C. Bawden, 1933) introduced new possibilities for disease monitoring by exploiting the distinct spectral reflection characteristics of healthy and diseased plants across various bands (D. Moshou et al., 2005; E. C. Davis, 2010). Extended abstracts should be written according to underlying structure of the structured abstract and in a form that is shorter than the full text. Subtitles of the extended abstract should be composed from the sections of abstract and keywords, introduction (including background, research problem, objectives, scope, significance), method, findings and argument, conclusion and suggestions. Extended abstract of the full paper shall be written with Times New Roman font, 1.5 line spacing and 12 font size. Extended abstracts can contain figures, tables, formulations or images. Extended abstracts should be in a structure that explains the content of the declaration therefore the preparation phase of the said work. Page margins are formed regarding the A4 page size and are 2.5 cm wide from the right, left, top and bottom. Extended abstracts should be at least three pages, but not more than five pages in length including the references.

In the case of potato canopies, the reflectance within the 680-750nm wavelength range gradually increases, with near infrared (NIR) reflectance notably higher than in the visible spectrum. Late blight infection in potatoes leads to significant changes in reflectance between 750-1350nm, where the severity of the disease correlates with a varying decrease in reflectance (M. H. D. Franceschini et al., 2019). The variability in reflectance in the infrared region is more pronounced than in the visible region, attributed to the disease's impact on the plant's physiological responses, such as changes in chlorophyll content and alterations in the leaf's internal structure (M. Zhang et al., 2003). A decrease in chlorophyll content results in lower absorption in the red light spectrum, which in turn increases the reflectance of infected plants. Conversely, the reflectance in the NIR region decreases due to changes in the leaf's internal structure.

Extensive research has been dedicated to detecting PLB, including the use of hyperspectral imaging under laboratory conditions (D. Bienkowski et al., 2019), building the correlation between indoor and field models through hyperspectral databases (S. Appeltans et al., 2022). Field spectroscopic studies were performed using a handheld field spectrometer to examine spectral variability between healthy and late blight-infected potato canopies. Assessing which

spectral variables and at which time of late blight can be detected over potato crops by field spectroradiometer has been the focus of several studies (C. I. Fernandez et al., 2020). However, these studies are typically limited to assessments at leaf canopy-level, posing challenges for comprehensive monitoring at large region.

To address these challenges, researchers have increasingly utilized Unmanned Aerial Vehicles (UAVs) equipped with diverse camera technologies for rapid, high-resolution image acquisition. This approach includes techniques for assessing PLB severity through UAV-captured RGB images (C. Roy et al. 2020), as well as the deployment of several indices (J. Rodríguez et al., 2021) or deep learning models to automate PLB diagnosis from UAV hyperspectral imagery (M. H. D. Franceschini et al., 2019; Y. Shi et al., 2022). Despite these technological advances, UAV-based methods encounter constraints related to flight endurance and coverage areas, while the lack of freely available and reliable satellite data hampers the scope of satellite-scale monitoring capabilities. In recent years, satellite remote sensing has gained prominence as a feasible alternative for crop disease monitoring. Notable applications include the detection and monitoring of wheat diseases using Quickbird and SPOT-6 imagery (L. Yuan et al., 2014), alongside the effective employment of Sentinel-2 data for monitoring wheat powdery mildew in China, where monitoring accuracies have reached up to 78% (Z. Qiong et al., 2018).

Despite the advancements in remote sensing technology, satellite-scale identification of PLB remains underexplored, with much of the current research focus on UAV-scale data. This limitation restricts the potential for extensive, convenient monitoring of the disease. Machine learning techniques have increasingly been leveraged to tackle the challenge of PLB in recent years. In a pivotal study by Duarte-Carvajalino et al. (2018), a thorough assessment of late blight across 14 potato genotypes was performed using UAV-derived multispectral imagery in conjunction with machine learning approaches. These included multilayer perceptrons, convolutional neural networks, SVM, and RF. The findings indicated that machine learning algorithms are a viable alternative to traditional visual estimation methods, achieving an average absolute error of 11.72%, which is deemed acceptable for real applications. In another study, Sun et al. (2023) compared and developed models for monitoring the severity of PLB using SVM, RF, and KNN algorithms. The positive outcomes of these classifications have shed light on the potential of machine learning in addressing PLB issues.

The current study introduces a method for detecting PLB through the analysis of multiscale imagery. Several disease indices are developed and seven coupled machine learning models are trained on high-quality ground-based PLB spectral data and disease metrics. These

models are then employed for classification and regression on new datasets, followed by monitoring and validation at ground, UAV and satellite scales. To evaluate the effectiveness of integrating machine learning algorithms with remote sensing data, a comparative analysis was conducted using ground-based datasets. The aim of this comparative analysis is to explore the feasibility of satellite-scale monitoring of PLB using spectral information. The proposed model is anticipated to contribute to the development of more rational and effective decision-making systems for PLB management.

2. Materials and Methods

All materials and methods that have been used in the work must be stated clearly and subtitles should be used when necessary. The materials and methods section describe in detail all the materials that have been used to conduct a study as well as the procedures that are undertaken. As research writing should be orderly and organized therefore the materials in each of its sub-section should be presented in a logical manner. In each section of the materials and methods start with the most important procedure and go on to the least important. Also provide headings and subheading to make each sub part of the materials and methods section clear and understandable.

2.1 Study Area

Considering the availability of historical data and the ease of acquiring meteorological and ground-level data, the study area includes the southern part of Hebei province and the northern part of the Inner Mongolia Autonomous Region. These regions are recognized as principal potato production areas in northern China (Fig. 1). The potato cultivation cycle in these regions typically begins with the sowing and emergence stages from April to May, progresses through a rapid growth phase from June to August, and culminates in the harvesting season from late August to September. Three counties are situated in a temperate continental climate zone, characterized by rising summer temperatures, ample rainfall, and high humidity —conditions that are favorable for the development of PLB (I. Iglesias et al., 2010). The climatic factors, coupled with the region's critical role in potato production, make it as a ideal study area for research on modeling and experimental validation of the PLB monitoring.

2.2 Ground Data

This study conducted experiments (S1 and S2) at locations in Sirenwa Township (41.684003N, 115.746497E) and Shandianhe Township (41.700667 N, 115.795664E) within Guyuan County, as shown in Fig. 1, with the red triangles indicating the locations where ground data were collected. The first small-area experiment comprised 9 plots, while the

second involved 16 plots, each plot measuring 1m × 1m. In the first small-area experiment, two potato varieties 'Yizhangshu No.12' and 'Shishu No.1' were planted. Two control groups and four infection groups were established, with seedlings infected starting on May 13, 2020, at a spore concentration of 9mg/100mL.

The second small-area experiment comprised 6 control groups and 30 infection groups, with the infection process beginning on May 14, 2020, using the same spore concentration as in the first experiment. Field spectral data for two areas were collected by the SVC HR-1024i full-band spectral radiometer under cloudless conditions between 10:00 and 14:00 on August 16, 2020. The initial hyperspectral measurements obtained from the radiometer were corrected and registered based on predefined marked positions. The raw data were preprocessed using SVC HR-1024i PC-side software, facilitating the export of full-band dataset. Simultaneously, PLB severity were observed in each 0.2m× 0.2m sample area using the methods (C. Gao et al., 2020; R. Sugiura et al., 2016). In total, 623 pairs of high-quality ground spectral and disease data were collected, which are essential for capturing micro-environmental conditions at the plot level and constructing the PLB monitoring model.

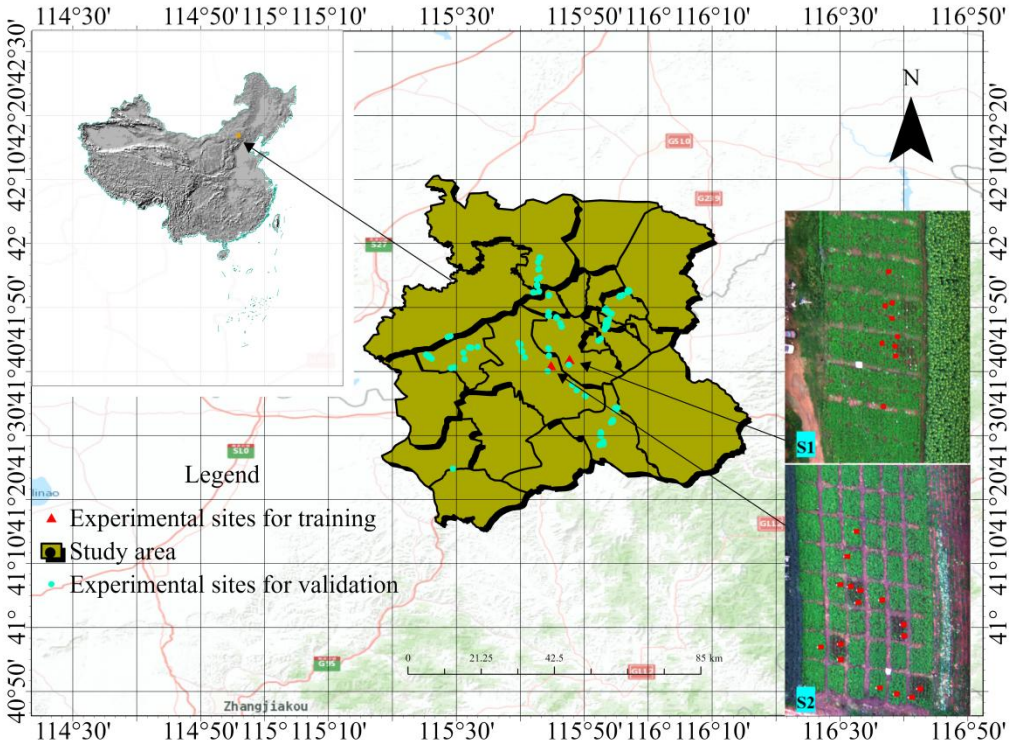

Figure 1: Presents a detailed geographic profile of the study area, illustrating the spatial distribution of the sampling points and plots.

2.3 Remote Sensing Data

The Sentinel-2 satellite constellation, developed and launched by the European Space Agency (ESA), provides high-resolution multispectral imaging. It comprises two satellites, Sentinel-2A and Sentinel-2B, each with a revisit 10-day revisit period， which combine to offer a 5-day revisit period. Equipped with a Multispectral Instrument (MSI), Sentinel-2 provides imagery across 13 spectral bands,including four with a 10-meter resolution (blue, green, red, and near-infrared), six with a 20-meter resolution (three red-edge bands and one near-infrared band), and three with a 60-meter resolution(as depicted in Table Ⅱ). In this study, a total of 10 bands spanning from the visible light to the shortwave infrared spectrum were utilized, encompassing both 10-meter and 20-meter resolutions. The Sentinel-2 data products were acquired from Google Earth Engine (GEE) (N. Gorelick et al., 2017), with a cloud masking algorithm applied to improve image quality.

2.4 Alternative Machine Learning Methods

To effectively construct the proposed PLB monitoring model, seven machine learning algorithms were employed. Below (Table 1) is a concise list explaining the machine learning algorithms evaluated in the study.

Table 1: Alternative Algorithms and Introductions

Algorithm	Description
Random Forest (RF)	An ensemble of decision trees that operate on the majority-vote principle for classification or average predictions for regression(L. Breiman, 2001).
Gradient Boosting (GB)	Sequentially trains weak learners, focusing on the residuals of the previous model, refining predictions through gradient descent (R. Lawrence et al., 2004).
K-Nearest Neighbors (KNN)	A non-parametric algorithm that uses the 'K' closest neighbors in the training set for classification or prediction (Y.-H. Lee et al., 2012).
Neural Network (NN)	Inspired by biological neural networks, capable of capturing complex, nonlinear relationships through interconnected layers of neurons (K. M. He et al., 2016).
Support Vector Machine (SVM)	Identifies an optimal hyperplane for class separation in feature space, using kernel functions for nonlinear data (V. N. Vapnik, 1995).
eXtreme Gradient	A variant of gradient boosting that introduces regularization

Boosting (XGBoost)	and parallel computing to improve efficiency and performance (T. Chen and C. Guestrin, 2016).
Categorical Boosting (CatBoost)	Specifically designed for classification problems, with technology for automatic transformation of categorical features and robust generalization (L. Prokhorenkova et al., 2017).

2.5 Multi-scale disease severity-spectral pattern and sensitivity analysis methods

Spearman's rank correlation coefficient Spearman's correlation coefficient is a nonparametric statistical method used to measure the correlation between the rank (or non-rank) of two variables without requiring the data to satisfy the requirements of a normal distribution or other specific distribution. This correlation coefficient was used to analyze the correlation between disease severity and spectral patterns.

To investigate the relationship between disease severity and spectral patterns across various scales, and to compare them for these scales in the context of late blight progression, we synthesized data from three distinct scales. This synthesis included alternative indices ─ excluding the UAV band's missing data ─ and was based on the reflectance values of Sentinel-2 satellite bands. Additionally, we incorporated the aggregate of six disease indices (excluding the UAV band's missing segment). Additionally, we incorporated the aggregate of six disease indices, again excluding the missing UAV band data for UAV scale. We then analyzed the sensitivity and correlation of each index to changes in disease severity using the correlation coefficient. During the visualization process, variables were ordered according to the eigenvalues of the matrix. These eigenvalues signify the extent of variation within the matrix, and by sorting them, we were able to cluster variables with higher correlations, thereby enhancing the clarity of the correlational structure.

2.6 Construction of PLB monitoring model

This study is dedicated to establishing a monitoring model for PLB , achieved through a series of meticulously designed steps. Initially, spectral data from potato canopies across the visible and near-infrared spectrum were collected using ground spectroradiometers, UAV spectroradiometer, and satellite sensors. With this high-quality ground observation data, a machine learning model was trained to identify and learn the spectral characteristics of PLB. The model was then executed on new datasets to perform classification and regression tasks,

predicting disease occurrence. Results were validated at ground, UAV and satellite scales to ensure its accuracy and applicability.

To evaluate the performance of different algorithms, a comparative analysis was conducted based on ground data, testing seven types of machine learning models: RF, GB, KNN, NN, SVM, XGBoost, and CatBoost. The inputs to the models were spectral bands and indices from the ground measurement and spectral data, and the outputs were disease severity, quantified on a scale of 0 to 100 for regression analysis, or categorized into five classes based on 20-unit intervals for classification.

3. Results and Discussion

3.1 Multiscale sensitivity analysis of PLB to spectral characteristics

In the ground-scale correlation matrix depicted in Figure 2, a non-statistically significant positive correlation is observed between disease severity and bands B12 and B4. Conversely, a statistically significant negative correlation is evident with the remaining bands, particularly with the near-infrared, red-edge bands and derived indices. This relationship is characterized by a negative correlation coefficient exceeding -0.9. These findings suggest that at the ground scale, as the severity of PLB increases, there is a corresponding decrease in the reflectance values of the red edge, near-infrared bands, and derived indices.

In the UAV-scale correlation plot shown in Figure 3, a strong negative correlation is observed between disease severity and the three calculated indices, as well as between the red edge and near-infrared bands of the Sentinel-2 satellite image. These relationships confirm the spectral analysis results, indicating a significant correlation between PLB severity and spectral data. At the UAV scale, PLB was positively correlated with the B4 band, weakly correlated with B2, B3 and B5, and strongly negatively correlated with other bands. Statistical analysis reveals a significant negative correlation, with values ranging from -0.89 to -0.74. The positive correlation for the B4 band suggests that an increase in reflectivity in this band is associated with a decrease in disease severity, possibly linked to certain spectral signatures of the plant health .

The statistically significant relationship between disease severity and the red edge and near-infrared bands is observed to be lower at the satellite scale compared to the ground and UAV scales (Figure 4). Specifically, the negative correlation coefficient falls within the range of -0.4 to -0.5, indicating a weaker association. Furthermore, the negative correlations observed among some bands are even less pronounced and do not reach statistical significance. These findings align with the spectral analysis results, which also reveal a more subdued alteration in the spectral reflectance curves compared to the ground and UAV scales. This suggests that,

at the satellite scale, there is a slight decrease in the spectral reflectance values of the red edge and near-infrared bands as the PLB severity increases.

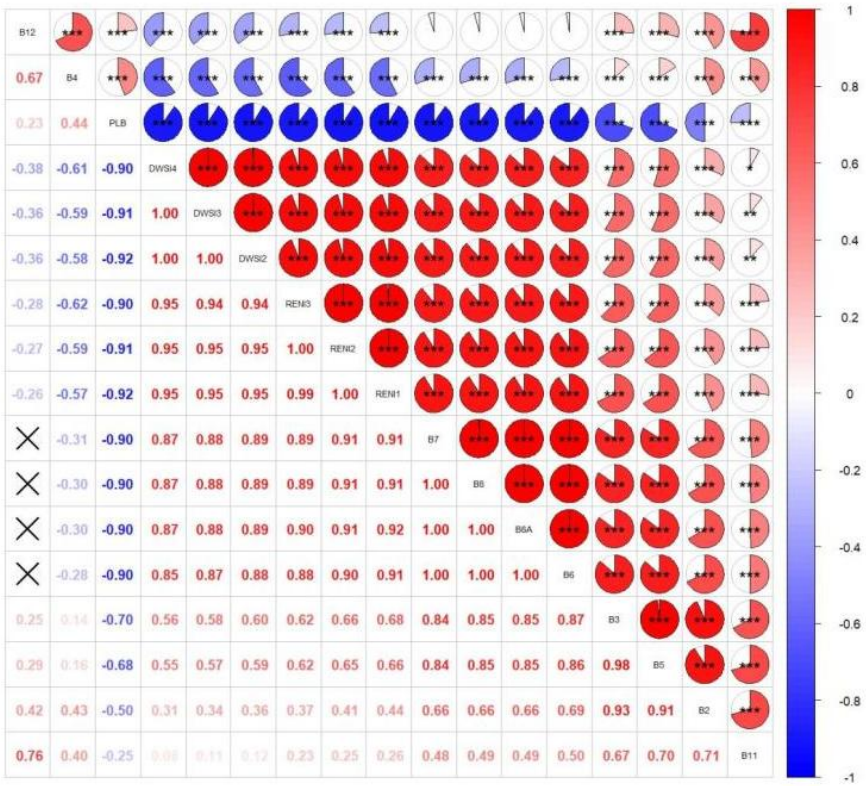

Figure 2: Correlations between PLB severity and spectrual bands/indices. Color bars indicate correlation strength: blue (negative) to red (positive). Asterisk denote significance levels: *p<0.05 ,**p<0.01,***p< 0.001.

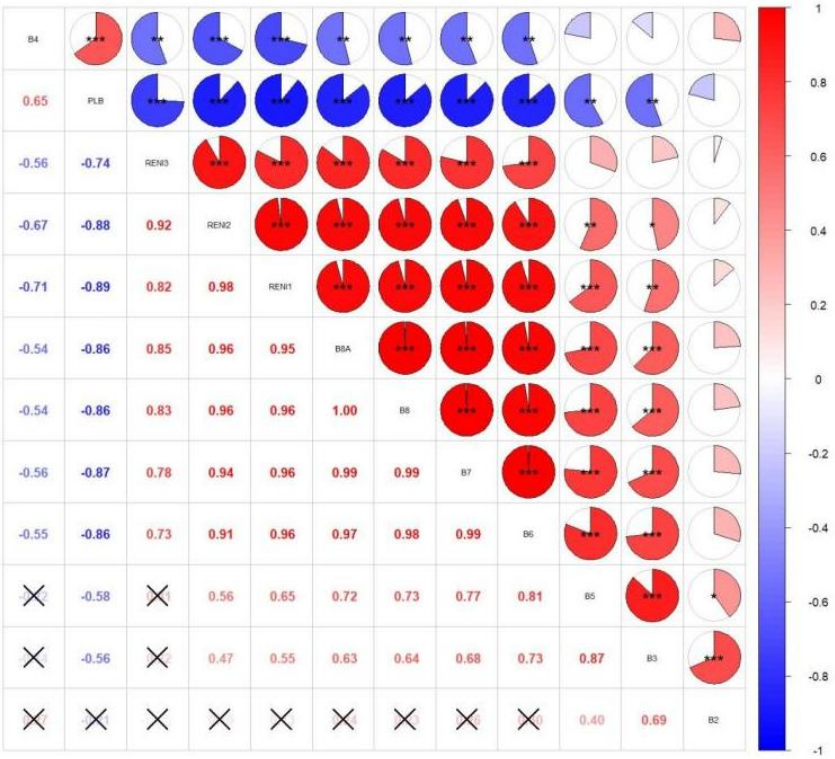

Figure 3: Correlations between PLB severity and UAV spectral bands/indices. Color bars indicate correlation strength: blue (negative) to red (positive). Asterisk denote significance levels: *p<0.05 ,**p<0.01,***p< 0.001.

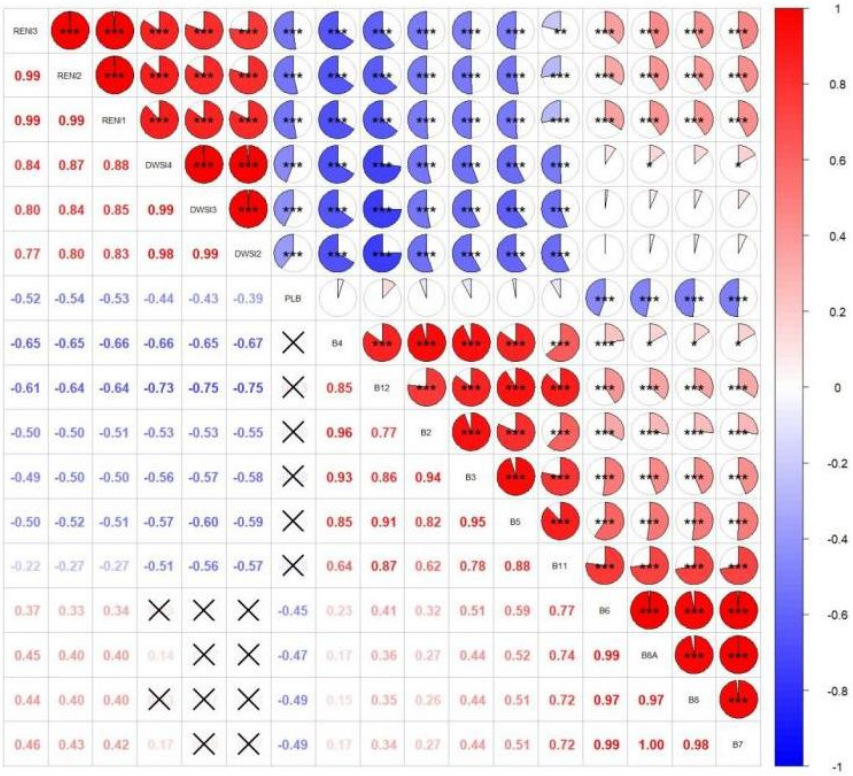

Figure 4: Correlations between PLB severity and Sentinel-2 spectrual bands/indices. Color

bars indicate correlation strength: blue (negative) to red (positive). Asterisk denote significance levels: *p<0.05 ,**p<0.01,***p< 0.001.

3.2 Analysis of disease occurrence law of inter-scale spectroscopy

This study identified a correlation between ground data and the incidence of PLB, as well as between remote sensing data and the incidence of PLB across varying scales, as illustrated in Figure 5. Generally this correlation weaken with the expansion of the scale. At the ground scale, a correlation was detected between disease severity and 13 spectral bands or indices. In contrast, at the UAV and satellite scales, the number of correlated bands or indices was reduced to 9 and 4, respectively. This reduction underscores the variability in correlation intensity across different scales. Additionally, the study observed a negative correlation between the PLB severity and the spectral responses in bands B6, B7, B8, and B8A, with the correlation weakening as the scale increases. These insights highlight the relationship between spectral signatures and the occurrence of PLB, as well as the distinctive application traits of remote sensing data at each scale.

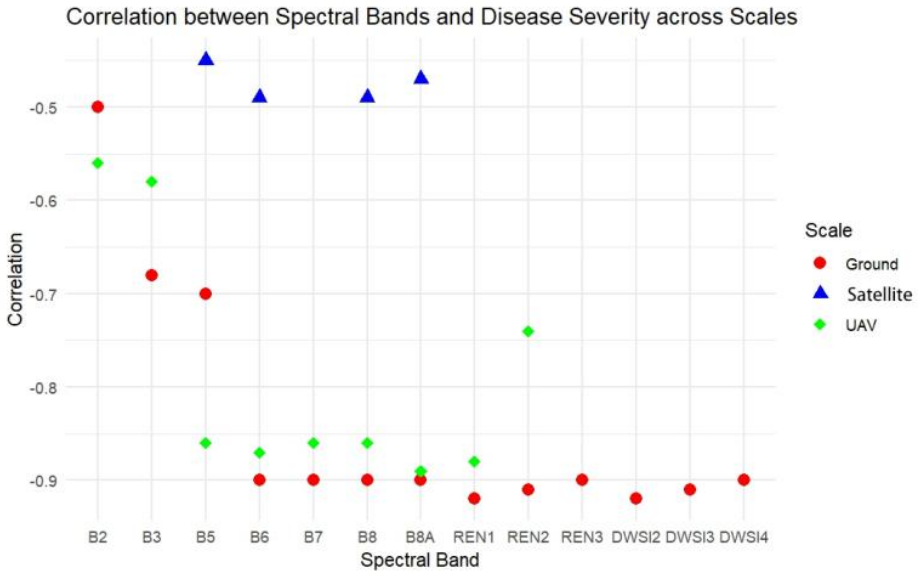

Figure 5: Differential performance of the correlation between spectral data and PLB severity at different scales (ground, UAV, satellite) with a significance level of 0.001.

3.3 PLB monitoring model based on machine learning and spectral indices

To detect PLB, a suite of machine learning algorithms(introduced in Section II) were employed to construct the PLB monitoring model, with a thorough performance evaluation. After normalizing the predictive indicators, the comparative results of seven machine learning models are visually presented in Figure 6. RF and CatBoost perform well in regression, exhibiting lower RMSE and higher R-square values. Their ability to capture

complex relationships within the data suggests a complex relationship between PLB severity and spectral features. XGBoost performs well in classification, demonstrating higher Accuracy, Precision, Recall, and F1 score. While CatBoost is perform well in regression, exhibiting lower RMSE and higher R-square values. Their ability to capture complex relationships within the data suggests a complex relationship between PLB severity and spectral features. XGBoost performs well in classification, demonstrating higher Accuracy, Precision, Recall, and F1 score. While CatBoost is comparable in these metrics, XGBoost slightly outperformed in Precision and Recall. Overall, XGBoost demonstrates superior performance on the dataset, particularly in classification, indicating its higher feasibility and practicality for satellite-scale monitoring compared to regression methods.

The XGBoost model's feature importance evaluation results indicates that the importance score of the DWSI2 feature is significantly higher than of other features (Figure 7), indicating its crucial role in predicting the target variable. Compared to previous studies, our research further emphasizes the importance of DWSI2, providing new insights for relevant work in the field. This finding enhances the understanding of machine learning methods for PLB monitoring and offers valuable information for satellite-scale monitoring. The computation of the DWSI2 index provides information about vegetation chlorophyll, moisture status, and land characteristics. Specifically, the numerator's red edge index minus the green band represents the specific reflection characteristics of the potato canopy, while the denominator involves combinations of SWIR1 and red bands. This comprehensive approach captures the reflection characteristics of both the potato canopy and land surface. DWSI2 is particularly effective in capturing features relevant to PLB monitoring, exhibiting high importance among various bands used in XGBoost machine learning modeling.

Thus, the PLB monitoring model integration of spectral indices and XGBoost algorithm (PLB-SI-XGBoost) was proposed for the ground-scale PLB prediction. Firstly, 596 high-quality ground data from the data source were divided into training and testing sets with a ratio of 0.8 to 0.2. To find the optimal model configuration,a grid search was conducted in the hyperparameter space. The hyperparameters such as learning rate, number of trees, maximum depth, subsample ratio, and column sampling ratio were all considered. The optimal combination was determined with a learning rate of 0.2, 50 trees, maximum depth of 3, subsample ratio of 0.8, and a column sampling ratio of 1.0.

To evaluate the PLB-SI-XGBoost model's generalization ability, the model was then applied to the testing set and calculated the RMSE and five-fold cross-validated R-square value. The five-fold cross-validation results (Table 2) show excellent performance, with an

RMSE of approximately 12.09 and an R-square of 0.98 on the training set. On the validation set, an MSE of 184.18, an R^2 of 0.88, and an RMSE of approximately 13.50 were observed. Overall, the proposed model performed well on the training set and exhibited relatively good performance on the validation set, demonstrating strong prediction performance and generalization ability for PLB.

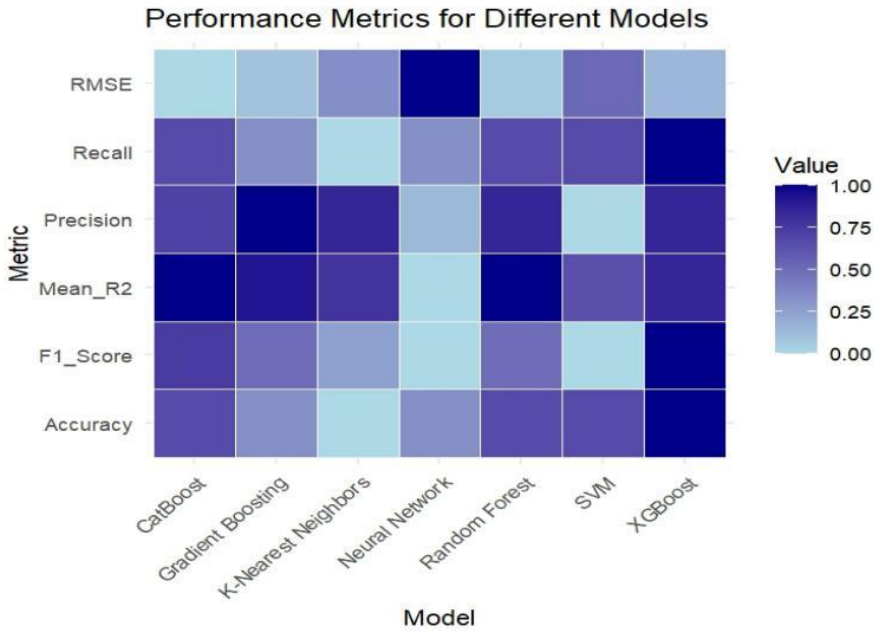

Figure 6: Comparative analysis of seven models integrating spectral indices and machine learning algorithms

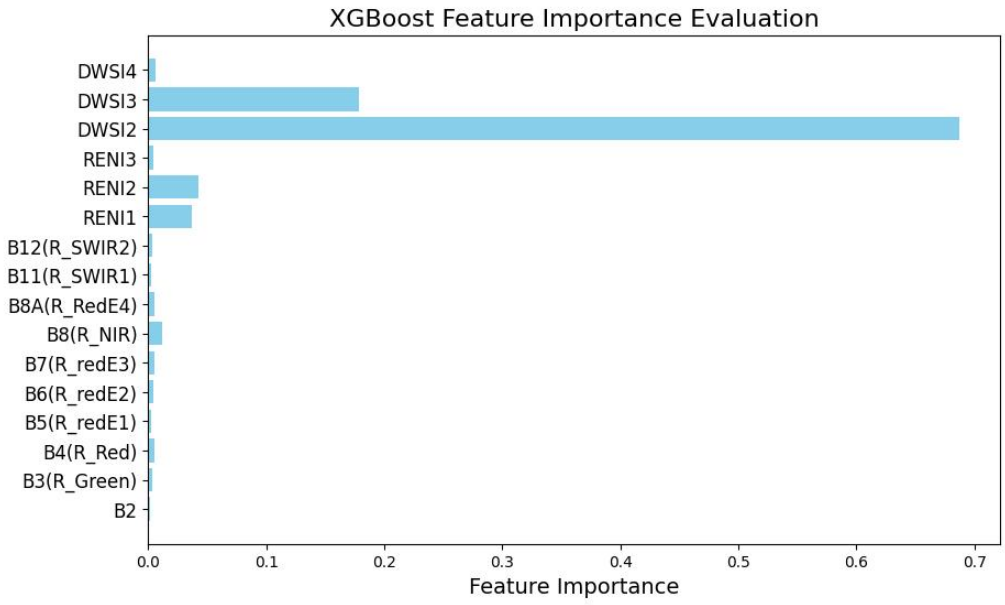

Figure7: Feature importance of the PLB-SI-XGBoost model based on the ground obersvation.

Table 2: the Different Performances of the Model Based on Ground-Scale Training Set and the Validation set

	Training set	Validation set
(RMSE)	12.09	13.50
(R^2)	0.98	0.88

To further validate the effectiveness of the model, the relationship between the PLB-SI-XGBoost model predictions and the actual observations is shown in Figure 8. Most of the scatter points were distributed near the diagonal, indicating that the model's predictions were in good agreement with reality. This demonstrates that the model's effectiveness in capturing overall trends and changes in the data, and in most cases accurately predicting the target variables. However, some discrete points were also observed, which could be outliers or instances of poor model performance. Overall, the comparison chart between prediction and reality reflects the model's performance, providing an important clue for evaluating the accuracy and reliability of the model.This indicates a generally favorable modeling effect.

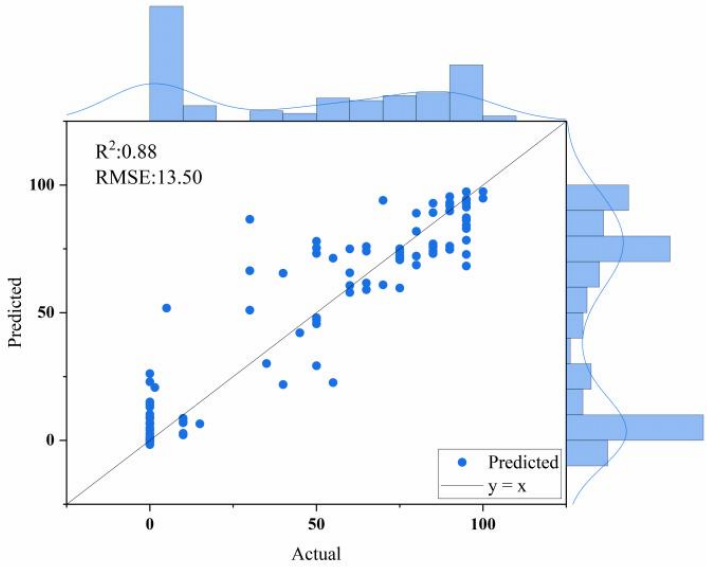

Figure 8: Comparison of the actual value with the prediction by the PLB-SI-XGBoost model trained at the ground-scale.

3.4 Verification of the model at the UAV and satellite scale

The PLB-SI-XGBoost model developed at ground scale, was then validated at the UAV and

satellite scale, respectively.

The validation results show an R-square of up to 0.74, and an RMSE of 18.27. The model performs well in UAV-scale regression and demonstrates good generalization ability.

The predicted values and the corresponding actual observed values for each sample are shown in Fig. 9. Most of the scatter points are distributed near the diagonal, indicating that the model's predictions are in good agreement with reality and accurately predict the target variable in most cases. However, some discrete points, especially those with low levels of disease, were observed. These discrete values could be outliers or samples where the model does not perform well in low PLB levels.

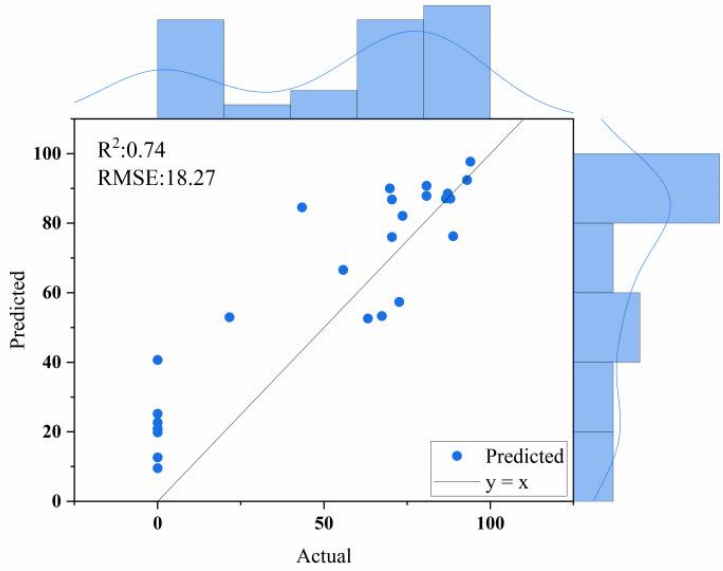

Figure 9: Comparison of the actual value and the prediction by the PLB-SI-XGBoost model trained at the UAV scale.

The accuracy of the PLB-SI-XGBoost model trained at ground scale was verified by Sentinel-2 data and field observation data in different regions. As shown in Figure 10, the 70% accuracy in the prediction of PLB indicates a good overall classification performance. The model demonstrated a precision of 65%, meaning that 65% of the samples predicted as infected were indeed infected, indicating a low misjudgment rate. Additionally, the model showed a recall rate of 70%, effectively capturing real infection samples and reducing the possibility of missed diagnoses, which is beneficial for disease control. Considering both accuracy and recall, the F1 score reaches 66%, highlighting the comprehensive performance evaluation of the model and providing useful guidance for balancing the accuracy and comprehensiveness of the predictions.

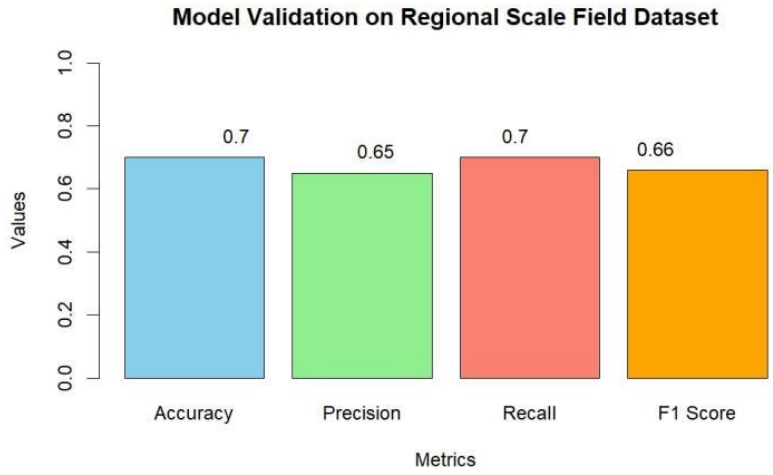

Figure 10: The validation results of the PLB-SI-XGBoost model at the satellite scale.

The confusion matrix (Figure 11) shows that 116 samples in the first category (disease severity 0-25) were accurately classified, 1 sample was incorrectly classified as the second category, and 9 were misclassified as the third category. For the second category (disease severity 25-50), only 1 out of 20 samples were correctly classified, 1sample was misclassified as the first category, and 2 were misclassified as the third category. For the third category (disease severity 50-100), 23 out of the 51 samples were accurately classified, 4 were misclassified as the first category, and 24 were misclassified as the second category. In general, the model demonstrates a good classification performance for samples with a disease degree of 0-25, and there is a certain degree of misclassification for samples with severity of 25-50 and 50-100. The results indicate that PLB can be better detected at an early stage when the disease severity is low, which can support intelligent and efficient decision-making in agricultural production , thereby reducing losses and costs.

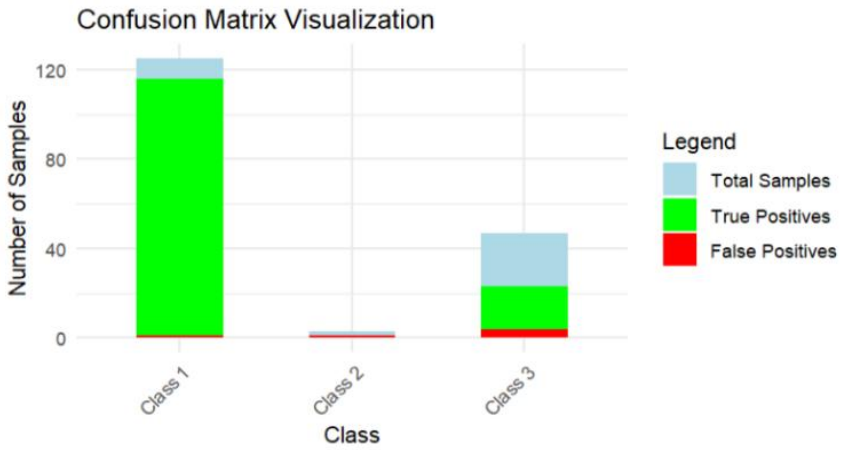

Figure 11: Confusion matrix results of the PLB-SI-XGBoost model for the severity classification of PLB in study area. (Class 1:0-25; Class 2:25-50; Class 3:50-100).

In addition,the spatial distribution of errors in the PLB-SI-XGBoost model was visualized (Figure 12). Overall, the model performs well in predicting the severity of PLB, enabling early-stage PLB monitoring, and providing decision support to mitigate further time loss.

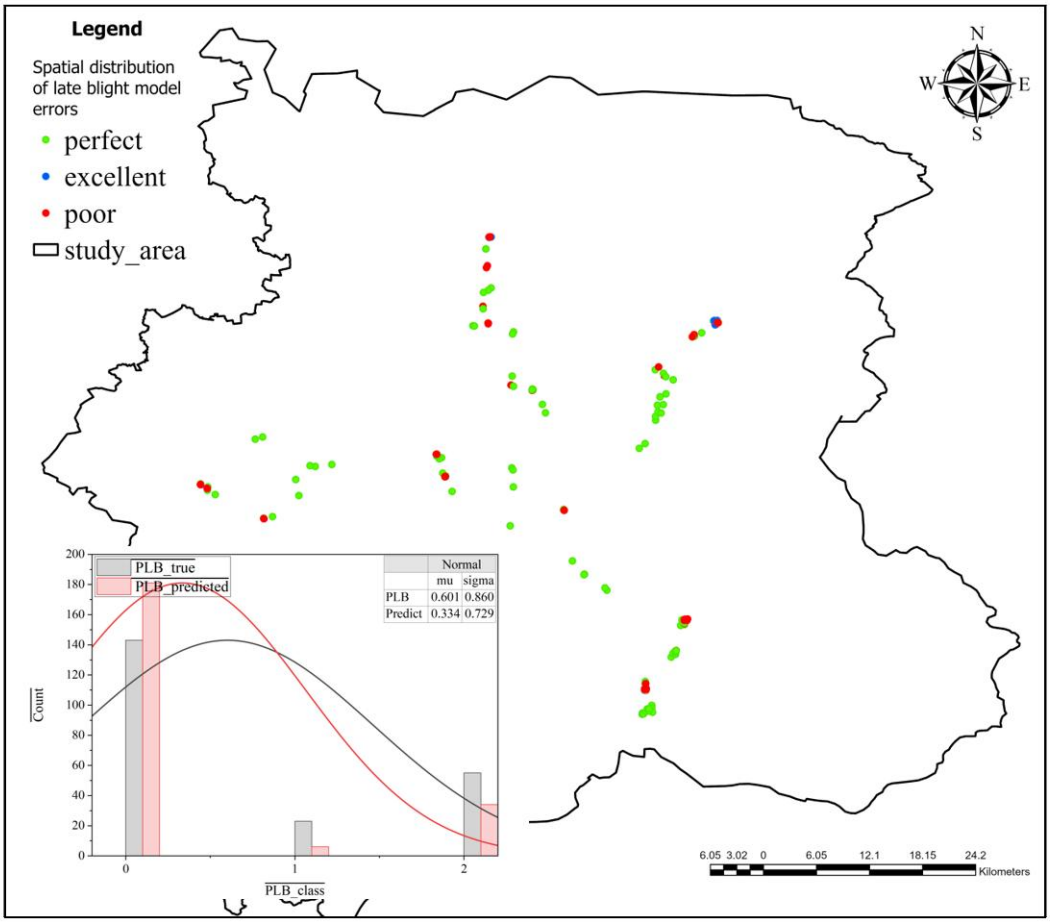

Figure 12: Spatial distribution of the error of the PLB-SI-XGBoost model on August 12, 2021. The error is calculated as the absolute difference between the predicted value and the field data: green represents perfect model performance, blue represents excellent performance, and red represents poor performance.

4. Conclusion and Recommendation

By integrating technological solutions, this study fills the gap in satellite-scale PLB monitoring, enhancing the agricultural decision-making process. The multi-scale data analysis underscores the effectiveness of machine learning algorithms and remote sensing technologies in detecting PLB over large areas. This research employs a synergistic approach, combining ground experiments, UAV spectral data, and satellite imagery, to overcome the limitations inherent in traditional monitoring methods.

This study is the first to emphasize the consistency in spectral reflectance trends across

ground, UAV, and satellite scales. These trends include specific band valleys and peaks, and the overall shape of the spectral reflectance curves. The red edge and near-infrared bands exhibit universal sensitivity to PLB detection across scales, highlighting the importance of these bands in spectral analysis. As disease severity increases, there is a general decrease in reflectance for specific bands (B6, B7, B8, B8A), with varying degrees of reduction corresponding to different levels of disease severity.

The study also explores scale-dependent spectral variations, finding that as PLB severity increases, the spectral values at the ground scale, especially in bands B6, B7, B8, and B8A, exhibit more pronounced changes compared to those observed at UAV and satellite scales. UAV-derived spectral curves show a narrower bandwidth and some gaps, possibly due to limitations in the spatial resolution and data capture continuity of UAV sensors. Compared to satellite systems, the spectral values at ground and UAV scales are similar, while the satellite scale shows wider upper and lower limits. Generally, it notes that scale effects may influence the generalizability of predictive models. While the ground scale provides the highest accuracy, UAV and satellite scales are equally effective for disease monitoring, offering a robust framework for multi-scale disease detection.

Furthermore, based on the developed Red Edge Index and Disease Stress Index with a suite of machine learning algorithms, a PLB-SI-XGBoost model was proposed. Notably, the proposed model demonstrated the highest average evaluation score of 0.88 and the lowest RMSE of 13.50 during ground scale validation, outperforming other algorithms. At the UAV scale, the proposed model achieved a robust R-squared value of 0.74 and an RMSE of 18.27. Moreover, the application of Sentinel-2 data for disease detection at satellite scale yielded an accuracy of 70% in the model. The PLB-SI-XGBoost model demonstrated its exceptional generalization capabilities.

The research indicates that although monitoring patterns exhibit a degree of universality and systematicity across scales, the strength of correlation may weaken as the scale expands. Satellite remote sensing emerges as a promising method for large scale monitoring, laying a solid foundation for the integration of machine learning with remote sensing techniques. Future studies should focus on refining model algorithms and enhancing the resolution and frequency of satellite data to improve the precision and applicability of PLB monitoring. In conclusion, this study pioneers new avenues for disease monitoring in satellite-scale agricultural production, paving the way for enhanced agricultural efficiency, minimized resource wastage, and the advancement towards Agriculture 4.0.

References

B. Liu et al., (2021). Promoting potato as staple food can reduce the carbon–land–water impacts of crops in China, *Nature Food, vol. 2, no. 8, pp. 570-577.*

C. Gao et al., (2020). Pathogen manipulation of chloroplast function triggers a light-dependent immune recognition, *Proceedings of the National Academy of Sciences, vol. 117, no. 17, p. 202002759.*

C. I. Fernandez, B. Leblon, (2020). A. Haddadi, J. F. Wang, and K. Wang, "Potato Late Blight Detection at the Leaf and Canopy Level Using Hyperspectral Data," *CANADIAN JOURNAL OF REMOTE SENSING, vol. 46, no. 4, pp. 390-413.*

C. Roy et al., (2020). Microbiome and ecology of a hot spring-microbialite system on the Trans-Himalayan Plateau, *Scientific Reports, vol. 10, no. 1, Apr 2020, Art no. 5917.*

D. Bienkowski, Matt J.Lees, Alison K.Gallagher, ChristopherNeilson, Roy, (2019). Detection and differentiation between potato (Solanum tuberosum) diseases using calibration models trained with non-imaging spectrometry data, *Computers and Electronics in Agriculture, vol. 167.*

D. Moshou et al., (2005). Plant disease detection based on data fusion of hyper-spectral and multi-spectral fluorescence imaging using Kohonen maps, *Real-Time Imaging, vol. 11, no. 2, pp. 75-83.*

E. C. Davis, (2010). A review of advanced techniques for detecting plant diseases, *Computers and Electronics in Agriculture.*

F. C. Bawden, (1933). Infra-Red Photography and Plant Virus Diseases, *Nature, vol. 132, no. 3326, pp. 168-168, 1933/07/01 1933.*

H. Sun et al., (2023). Potato late blight severity monitoring based on the relief-mRmR algorithm with dual-drone cooperation, Computers and Electronics in Agriculture, vol. 215, p. 108438.

I. Iglesias, O. Escuredo, C. Seijo, and J. Méndez, (2010). Phytophthora infestans Prediction for a Potato Crop, *American Journal of Potato Research, vol. 87, no. 1, pp. 32-40.*

J. Duarte-Carvajalino, D. Alzate, A. Ramirez, J. Santa-Sepulveda, A. Fajardo-Rojas, and M. Soto-Suárez, (2018). Evaluating Late Blight Severity in Potato Crops Using Unmanned Aerial Vehicles and Machine Learning Algorithms, *Remote Sensing, vol. 10, no. 10, pp. 1513.*

J. Rodríguez, I. Lizarazo, F. Prieto, and V. Angulo-Morales, (2021). Assessment of potato late blight from UAV-based multispectral imagery," *Computers and Electronics in Agriculture, vol. 184, p. 106061.*

K. M. He, X. Y. Zhang, S. Q. Ren, J. Sun, and Ieee, (2016). Deep Residual Learning for Image Recognition, *IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION (CVPR), 2016.*

L. Breiman, (2001). Random Forests, Machine Learning, vol. 45, no. 1, pp. 5-32.

L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, (2017). CatBoost: unbiased boosting with categorical features.

L. Yuan, J. Zhang, Y. Shi, C. Nie, L. Wei, and J. Wang, (2014). Damage Mapping of Powdery Mildew in Winter Wheat with High-Resolution Satellite Image, *Remote. Sens., vol. 6, pp. 3611-3623.*

M. H. D. Franceschini, H. Bartholomeus, D. F. Van Apeldoorn, J. Suomalainen, and L. Kooistra, (2019). Feasibility of Unmanned Aerial Vehicle Optical Imagery for Early Detection and Severity Assessment of Late Blight in Potato, *Remote Sensing, vol. 11, no. 3.*

M. Zhang, Z. Qin, X. Liu, and S. L. Ustin, (2003). Detection of stress in tomatoes induced by late blight disease in California, USA, using hyperspectral remote sensing, *International Journal of Applied Earth Observations & Geoinformation, vol. 4, no. 4, pp. 295-310.*

N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore, (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone, *Remote Sensing of Environment, vol. 202, pp. 18-277.*

R. Lawrence, A. Bunn, S. Powell, and M. Zambon, (2004). Classification of remotely sensed imagery using stochastic gradient boosting as a refinement of classification tree analysis, *Remote Sensing of Environment, vol. 90, no. 3, pp. 331-336.*

R. Sugiura et al., (2016). Field phenotyping system for the assessment of potato late blight resistance using RGB imagery from an unmanned aerial vehicle, *Biosystems Engineering, vol. 148, pp. 1-10.*

S. Appeltans, J. G. Pieters, and A. M. Mouazen, (2022). Potential of laboratory hyperspectral data for in-field detection of Phytophthora infestans on potato, P*recision Agriculture, no. 3, p. 23*.

T. Chen and C. Guestrin, (2016). XGBoost: A Scalable Tree Boosting System, *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, *2016*.

V. N. Vapnik, (1995). The Nature of Statistical Learning Theory, *Springer, 1995.*

Y. Shi, L. Han, A. Kleerekoper, S. Chang, and T. Hu, (2022). Novel CropdocNet Model for Automated Potato Late Blight Disease Detection from Unmanned Aerial Vehicle-Based Hyperspectral Imagery," *Remote Sensing, vol. 14, no. 2, p. 396*.

Y.-H. Lee, C.-P. Wei, T.-H. Cheng, and C.-T. Yang, (2012). Nearest-neighbor-based approach to time-series classification, *Decision Support Systems, vol. 53, no. 1, pp. 207-217.*

Z. Qiong, H. Wenjiang, C. Ximin, S. Yue, and L. Linyi, (2018). New Spectral Index for Detecting Wheat Yellow Rust Using Sentinel-2 Multispectral Imagery, S*ensors, vol. 18, no. 3, p. 868.*