Prediction of Potassium Content in Rice Leaves Based on Spectral Features and Random Forests

Yue Yu, Haiye Yu, Xiaokai Li, Lei Zhang and Yuanyuan Sui *

School of Biological and Agricultural Engineering, Jilin University, Changchun 130022, China; yshi928ch@163.com (Y.Y.); haiye@jlu.edu.cn (H.Y.); lixk20@mails.jlu.edu.cn (X.L.); z_lei@jlu.edu.cn (L.Z.)
* Correspondence: suiyuan@jlu.edu.cn

Abstract: The information acquisition about potassium, which affects the quality and yield of crops, is of great significance for crop nutrient management and intelligent decision making in smart agriculture. This article proposes a method for predicting the rice leaf potassium content (LKC) using spectral characteristics and random forests (RF). The method screens spectral characteristic variables based on the linear correlation analysis results of rice LKC and four transformed spectra (original reflectance (R), first derivative reflectance (FDR), continuum-removed reflectance (CRR), and normalized reflectance (NR)) of leaves and the PCA dimensionality reduction results of vegetation indices. Following a second screening of the correlated single band and vegetation index variables of the four transformed spectra, the RF is used to obtain the mixed variable (MV), and regression models are developed to achieve an accurate prediction of rice LKC. Additionally, the effect of potassium spectral sensitivity bands, indices, spectral transformation form, and different modeling methods on rice LKC prediction accuracy is assessed. The results showed that the mixed variable obtained with the second screening using the random forest feature selection method could effectively improve the prediction accuracy of rice LKC. The regression models based on the single band variables (BV) and the vegetation index variables (IV), FDR–RF and IV–RF, with $R^2$ values of 0.62301 and 0.7387 and RMSE values of 0.24174 and 0.15045, respectively, are the best models. In comparison to the previous two models, the MV–RF validation had a higher $R^2$ and a lower RMSE, reaching 0.77817 and 0.14913, respectively. It can be seen that the RF has a better processing ability for the MV that contains vegetation indices and IV than for the BV. Furthermore, the results of different variable screening and regression analyses also revealed that the single band’s range of 1402–1428 nm and 1871–1907 nm, as well as the vegetation indices constituted of reflectance 1799–1881 nm and 2276–2350 nm, are of great significance for predicting rice LKC. This conclusion can provide a reference for establishing a universal vegetation index related to potassium.

Keywords: random forests; feature selection; transformed spectra; vegetation index; potassium content; rice

1. Introduction

Potassium (K) is the most important inorganic metal ion in plants and is involved in many physiological processes that support plant growth and development [1,2]. It plays an important role in the transport of photosynthetic products, intercellular osmotic regulation, and the synthesis of various enzymes [3–6]. Potassium deficiency can result in slow plant growth and development, susceptibility to lodging, and reduced yields. Excess potassium can also inhibit cation uptake, cause disease, waste resources, and pollute the environment [7,8]. Quantitative research on crop potassium is needed to improve agricultural production efficiency and achieve environmental sustainability.

The traditional quantitative research of biochemical concentration of plant nutrients mainly relies on destructive field sampling and complex laboratory analysis, which is very inefficient [9]. In the past few decades, the development of satellite remote sensing and...
UAV low-altitude remote sensing technology has relied on spectral technology, which has been proven to be a fast, real-time, high-resolution, and non-destructive detection method [10,11]. However, the application of remote sensing technology has a certain demand for an experimental scale, which is suitable for field experiments, and the image quality of remote sensing may not be usable due to weather factors, such as cloudy and rainy days [12]. The spectral analysis method based on the canopy and leaf is more suitable for small-scale experiments, such as plot and pot experiments. It is also more suitable for quantitative research on nutrient biochemical concentrations at the canopy and leaf scale of plants. The related principle is that the change in photosynthesis, cell structure, and chemical bond stretching vibration caused by nutrients affects the spectral reflectance of plants in the visible (VIS, about 380–780 nm), near-infrared (NIR, about 750–1100 nm), and short wave infrared (SWIR, about 1100–2500 nm) reflectance [13]. The spectral reflectance of plants were determined using their chemical and morphological characteristics. The VIS was mainly affected by various pigments, such as chlorophyll and anthocyanin. The reflectivity of the NIR was higher than other bands, which was caused by the inner cell structure of the plant. Two absorption peaks around 1450 and 1950 nm were consistent with the absorption characteristics of water content [14,15]. Changes in the potassium content would unavoidably have an impact on reflectance since there is a direct relationship between potassium content and the color, structure, and moisture content of plant leaves [5,16–19], which has been confirmed in many related studies. Pimstein [20] believed that reflectance at 1450 nm was significantly correlated with wheat LKC. Gómez Casero [21] found that potassium deficiency in olive trees can lead to changes in the NIR reflectance. Stein [22] concluded that potassium was closely related to the reflectance of the red and green regions of loblolly pine leaves. Yang [23] found that the reflectance of wheat leaves under low potassium conditions was higher than that of normal leaves at 1700 nm but lower than that of normal leaves at 2300 nm. Pacumbaba [24] held the view that potassium-deficient lettuce leaves showed an increase in reflectance at the red edge and the NIR range.

To accurately capture the response of the reflectance to nutrient changes, the choice of spectral characteristics is critical. A common and effective method is to construct a vegetation index (VI) [25]. Some vegetation indices based on fixed band combination calculations have been developed through a lot of research and verification, which have played a great role in the quantitative and inversion research of many biochemical indicators. Wen [26] had a good diagnostic effect on leaf nitrogen content during four growth stages of maize based on the VI of green or red combined with the red edge or near-infrared bands of the first derivative reflectance. Wang [27] found that the TCARI/OSAVI could effectively estimate nitrogen content in rice leaves and plant nitrogen accumulation, and the three-band index TBVI performed best in LNA estimation. Delegido [28] and Yu [29] discovered that the NAOC had good inversion capacity after using the index and its mathematical transformation form to invert the chlorophyll content of green vegetation, including sycamores and palms; Rodrigues [30] successfully used physiological parameters as bridges to predict the grain yield and the protein content of winter wheat using the NDVI and the RVI. Wang [31] confirmed through further research that the RVI had a greater advantage in predicting maize grain protein content.

Compared to the biochemical indicators above, the response of published VIs to changes in potassium content is weak, and research on VIs used to assess plant potassium nutrition has not yet reached a consistent conclusion [32,33]. The indices obtained from some studies are not universal; rather, they are valid for a single crop under specific circumstances. As a result, they still need to be tested in actual situations. Chen [34] and Peng [35] concluded that the indices based on the first derivative reflectance were highly sensitive to the potassium content of plant leaves, and the single-band first derivative index dR1686 and dLOG (1/R1337) could more accurately predict the potassium concentration in litchi leaves. Lu [36] applied these indices to the estimation of the potassium content in rice plants and found that the indices based on the first derivative, including dLOG
had less than optimal results. Thus, they developed a group of high prediction accuracy indices, which are the NDSI (FD $1505$, FD $805$), the NDSI (R $1210$, R $1105$), and the NDSI (LOG $1210$, LOG $1180$), using correlation analysis between dual band combinations and potassium content. Pimstein [20] found that the indices (R $870$ − R $1450$)/(R $870$ + R $1450$) and (R $1645$ − R $1715$)/(R $1645$ + R $1715$) could effectively predict potassium content and accumulation in wheat plants, which was confirmed in the research of Mahajan [13]. However, Yang [23] found in his research that these indices could not achieve good results in retrieving the potassium content and accumulation in narrow-leaved wheat plants. His modified ratio index (R $2275$ − R $445$)/(R $1875$ − R $445$) had higher accuracy and stability than the two-band indices and are suitable for quantitative estimation of potassium content in rice leaves; however, it was noted that the application of these indices under different environmental conditions still needs to be verified.

In previous studies, the types of characteristic variables used for quantitative analysis of crop potassium nutrition were different. These variables existed in the form of single-band, dual-band, or three-band combinations of VIs or their transformations, and the predictive ability and universality of the characteristic variables required further research. Artificially intelligent models have emerged in recent years as useful tools for addressing the nonlinear relationships between independent and dependent variables. Examples included artificial neural networks (ANN), deep learning (DL), support vector machines (SVM), and random forests (RF) [38–40]. As an undifferentiated sampling method, the RF might randomly sample any type of feature, determine each feature’s value individually, and rank them, which could assist us in deliberately choosing more significant features. The mechanism of RF included the verification process of out-of-bag data, so there was no need to adopt cross-validation, and it could effectively prevent the model from over-fitting. It has been widely used in ecological research and has achieved good performance in the prediction of plant physiological and biochemical indicators and screening of spectral characteristic variables [41,42]. Majumdar [43] ranked the environmental factors affecting rice irrigation water demand according to the importance results of random forest characteristics and successfully predicted the optimal water demand unique to all rice growth stages based on the optimal characteristics; Li [44] predicted the chlorophyll content in shallow lakes by combining the mRMR and RF feature selection methods with the regression model and found that the RF method had a better predictive effect. In remote sensing research, Jiang [45] confirmed the finding that VIs were typically more significant than single bands by examining the importance of features based on single bands and VIs from multispectral images. Pranga [46] developed an RF method based on structural information and VI variables, which improved the estimation effect of ryegrass biomass yield and was superior to “structure only” and “indices only” models. Viljanen [47] extracted grassland canopy RGB information, near-infrared distance intensity value, VIs, and remote sensing sensitive bands to estimate grassland canopy height and biomass. Compared with simple linear regression, the use of an RF that integrates multiple features improved the results. Karunaratne [48] compared the “structural only” information from movement, the “spectral only” information from VIs, and the best performance of their fusion. In every prediction outcome, the fusion-based RF outperformed the independent information model. When Johansen [49] used an RF to sort the phenotypic data from many sources, including tomato shape features, VIs, and entropy texture, he discovered that shape characteristics had the most important impact in predicting the average biomass and yield of tomatoes. Lu [36] increased the accuracy of the potassium content estimation model by incorporating climatic elements into the feature screening of the environmental characteristics in rice. The research has demonstrated that random forests have a potent filtering capability for variables of complicated and various sorts.

In addition, many research results indicate that spectral transformation is also of extreme importance for the quantitative study of potassium. As one of the most used forms
of transformation, a first derivative transformation has shown a good ability to predict potassium elements in the studies of Chen [34], Peng [35], and Azadnia [50]. Among other spectral forms of transformation, Mutanga [51] successfully used continuum-removed absorption characteristics to predict and analyze potassium content in forages. Ferwerda [52] found that the best prediction results were also obtained utilizing continuum-removed reflectance in predicting sodium, potassium, and calcium content in several woody plants. Wu [53], Verger [54], and Yang [23] concluded that in the process of obtaining plant leaf spectra, the absolute value of the reflectance was often underestimated due to unstable lighting conditions or narrow leaves that could not fully cover the window, leading to a decrease in the accuracy of plant potassium content prediction. Therefore, spectral normalized methods were applied to eliminate the effects of differences in incident radiation and leaf width. In this study, we evaluated the consequences of several spectral transformations.

Based on the issues mentioned above, this study makes use of VIs and sensitive bands of transformed spectra from earlier potassium quantitative studies under the assumption that there is no general VI that applies to potassium. To obtain the best modeling variable set and improve the rice LKC’s prediction accuracy, the two variables and their mixed variables are second screened using the random forest feature selection technique. This work could be a guide for the future development of potassium-related VI.

2. Materials and Methods

2.1. Experimental Design

The experiment (Exp.) was carried out at the Agricultural Experiment Station of Jilin University in Changchun, Jilin Province (125°14′52″ E, 43°57′4″ N). The soil’s basic physical and chemical characteristics taken using a spectrophotometer (UNICO WFJ-2100, Shanghai, China) in the tillage layer (0–20 cm) were as follows: 27.8 g·kg⁻¹ of soil organic matter, 93.4 mg·kg⁻¹ of alkali-hydro nitrogen, 46.1 mg·kg⁻¹ of rapidly available phosphorus, and 166 mg·kg⁻¹ of rapidly available potassium. The specifics of the experiment are displayed in Table 1.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Time</th>
<th>Cultivars</th>
<th>Experimental Conditions</th>
<th>N Treatment (kg ha⁻¹)</th>
<th>K Treatment (kg ha⁻¹)</th>
<th>Transplanting Date</th>
<th>Sampling Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exp.1</td>
<td>2021</td>
<td>Jijing 88</td>
<td>Plot experiment</td>
<td>60, 120, 180</td>
<td>50, 100, 150</td>
<td>30 May</td>
<td>7/4, 7/16, 8/2</td>
</tr>
<tr>
<td>Exp.2</td>
<td>2022</td>
<td>Jijing 88</td>
<td>Pot experiment</td>
<td>100, 150, 200</td>
<td>50, 100, 150</td>
<td>2 June</td>
<td>7/8, 7/22, 8/5</td>
</tr>
<tr>
<td>Exp.3</td>
<td>2022</td>
<td>Jiudao 546</td>
<td>Pot experiment</td>
<td>100, 150, 200</td>
<td>50, 100, 150</td>
<td>2 June</td>
<td>7/22</td>
</tr>
</tbody>
</table>

A Jijing 88 type of rice was used in Exp.1 and Exp.2 and a Jiudao 546 type in Exp.3. Exp.1 was carried out in group plots where each had a size of 3 m × 3 m. To prevent water and fertilizer leaks, isolation zones between the plots were made using PVC boards. The total height of PVC boards was 100 cm, placed 40 cm above ground and 60 cm below ground. The rice was transplanted to each plot once it developed into seedlings, with four seedlings in each hole and three replications. Three nitrogen fertilizer gradients were set up, namely 60 kg·ha⁻¹ (N1), 120 kg·ha⁻¹ (N2), and 180 kg·ha⁻¹ (N3). The nitrogen fertilizer used was urea (Sinofert Holding Ltd., Beijing, China), which had a 46% nitrogen content. It was sprayed in the following proportions: 50% base fertilizer, 30% tiller fertilizer, and 20% ear fertilizer. Three potassium fertilizer gradients were established, namely 50 kg·ha⁻¹ (K1), 100 kg·ha⁻¹ (K2), and 150 kg·ha⁻¹ (K3). The potassium fertilizer was potassium chloride (K₂O content 60%, Sinofert Holding Ltd., Beijing, China), which was applied following the ratio of 50% base fertilizer and 50% ear fertilizer. Calcium superphosphate (P₂O₅ content 16%, Hubei Jiangshan Chemical Co., Jingmen, China) was used as the basis fertilizer in all treatments, with a uniform application standard of 100 kg·ha⁻¹. The soil for Exp.2 and Exp.3 were obtained from Jilin University’s Agricultural Experiment Station. The gradient of nitrogen fertilizer was established at 100 kg·ha⁻¹ for N1, 150 kg·ha⁻¹ for N2, and 200 kg·ha⁻¹ for N3. The potassium fertilization method used in Exp.1 was
also used here. Each treatment contained 10 pots of rice with four plants in each pot and three replications.

2.2. Data Collection

2.2.1. Measurement of Spectral Data
Because the chemical determination of potassium requires the destruction of rice leaves, the Agilent Cary 5000 (Agilent Technology, Santa Clara, CA, USA) equipped with an integrating sphere was used to measure the reflectance before measuring the content of potassium. Compared with the portable spectrometer, it has the advantages of a wider detection range (350–2500 nm), higher resolution (1 nm), and lower error [55]. When measuring, the target leaf was fixed on the bracket, with the light spot focused on the designated area. Then, the sample chamber was closed, and the reflectance measurement was completed under the condition of avoiding light.

2.2.2. Measurement of Agronomic Parameter
Following the measurement of reflectance, the leaf was cut, its weight was recorded using a 10,000-cent scale balance (Mettler Toledo ME104E, Zurich, Switzerland), and it was then placed in a cowhide bag and dried for 30 min in a 105 °C drum drying box before being kept dry at 70 °C to maintain a constant weight. The leaf was then ground up, sieved through a 0.25 mm mesh sieve, then crushed again. Once all the target leaves had been processed, a sample for testing was taken and the process was repeated. The samples were treated with H₂O₂ (Sinopharm Chemical Reagent Co., Ltd., Shanghai, China) concentrated H₂SO₄ (Sinopharm Chemical Reagent Co., Ltd., Shanghai, China) until the digestion solution was clear. After cooling, the potassium ion concentration in the digestion solution was measured using a plasma spectrometer iCAP-7000 (Thermo Fisher Scientific, Waltham, MA, USA), and then, the potassium content LKC (%) in each sample was calculated according to Formula (1). Table 2 shows the descriptive statistics of LKC measured in the two experiments.

\[
LKC (%) = \frac{\text{Weight of potassium (g)}}{\text{Dry weight of sample (g)}} \times 100\%.
\] (1)

Table 2. Descriptive statistics of rice LKC (%).

<table>
<thead>
<tr>
<th>Database</th>
<th>Number of Samples</th>
<th>Max</th>
<th>Min</th>
<th>Med</th>
<th>Mean</th>
<th>Sd</th>
<th>Cv (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training data</td>
<td>Exp.1</td>
<td>174</td>
<td>2.48</td>
<td>0.54</td>
<td>1.68</td>
<td>1.64</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>Exp.2</td>
<td>102</td>
<td>3.51</td>
<td>0.97</td>
<td>1.79</td>
<td>1.79</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>All data</td>
<td>276</td>
<td>3.51</td>
<td>0.54</td>
<td>1.73</td>
<td>1.69</td>
<td>0.36</td>
</tr>
<tr>
<td>Testing data</td>
<td>Exp.3</td>
<td>46</td>
<td>2.31</td>
<td>0.61</td>
<td>1.42</td>
<td>1.51</td>
<td>0.43</td>
</tr>
</tbody>
</table>

Max: maximum; Min: minimum; Med: median; Sd: standard deviation; Cv: Coefficient of variation.

2.3. Data Processing
Figure 1 shows the flowchart of the methodology used in this work. The figures involved in the spectral analysis and modeling method are drawn using the following software: MATLAB R2021b (Mathworks Inc., Natick, MA, USA), Excel (Microsoft Office 2021, Redmond, WA, USA), SPSS 26.0 (IBM Inc., Chicago, IL, USA), and Origin 2022 (OriginLab, Hampton, VA, USA).
2.3.1. Transformation of the Spectral Reflectance

In order to reduce the effect of instrument noise on the reflectance, the 350–399 nm region of the band was removed. Then, the four transformed spectra were selected: original reflectance \( R \), first derivative reflectance \( FDR \), continuum-removed reflectance \( CRR \), and normalized reflectance \( NR \), which are shown in Figure 2. The transformed reflectance was defined according to the following equations:

\[
FDR(\lambda) = \frac{R(\lambda_{i+1}) - R(\lambda_{i-1})}{2\Delta\lambda}, \quad (2)
\]

\[
CRR(\lambda) = \frac{R(\lambda)}{R_c(\lambda)}, \quad (3)
\]

\[
NR(\lambda) = \frac{R(\lambda)}{1 + \frac{1}{2500} \sum_{\lambda=400}^{3800} \Delta R(\lambda)}, \quad (4)
\]

where \( R(\lambda_i) \), \( R(\lambda_{i+1}) \), and \( R(\lambda_{i-1}) \) are the reflectance at wavelengths \( i \), \( i + 1 \), and \( i - 1 \), respectively; \( R_c(\lambda) \) is the reflectance of the continuum \( R_c \) at wavelength \( I \); \( \Delta\lambda \) is the wavelength increment.

Figure 2. Mean values and standard errors of \( R \) (a), \( FDR \) (b), \( CRR \) (c), and \( NR \) (d) from samples. Solid lines show the mean values, and shadow parts indicate the mean ± standard deviation (SD) range.
2.3.2. Vegetation Indices (VI)

A generally agreed-upon result has not yet been reached because there have been few investigations on the sensitive bands of the plant’s potassium spectra and the formation of associated VI. It is still necessary to confirm the adaptation of the proposed indices. Based on various modified spectra, 34 VIs were gathered as part of decades-long potassium research to assess the potassium content of rice leaves. The calculation methods and sources of some published VIs are shown in Table 3.

Table 3. Published VIs related to potassium.

<table>
<thead>
<tr>
<th>No.</th>
<th>Formula</th>
<th>Plant Type</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>VI.1</td>
<td>$(R_{720} - R_{1450})/(R_{720} + R_{1450})$</td>
<td>Wheat (PKC and PKA)</td>
<td>[20]</td>
</tr>
<tr>
<td>VI.2</td>
<td>$(R_{1645} - R_{1715})/(R_{1645} + R_{1715})$</td>
<td>Wheat (PKC and PKA)</td>
<td>[20]</td>
</tr>
<tr>
<td>VI.3</td>
<td>$R_{1875}$</td>
<td>Litchi (LKC)</td>
<td>[34]</td>
</tr>
<tr>
<td>VI.4</td>
<td>$\log(1/FDR_{1377})$</td>
<td>Litchi (LKC)</td>
<td>[34]</td>
</tr>
<tr>
<td>VI.5</td>
<td>$(R_{1705} - R_{1385})/(R_{1705} + R_{1385})$</td>
<td>Rice (LKC)</td>
<td>[37]</td>
</tr>
<tr>
<td>VI.6</td>
<td>$R_{1705} - R_{1385}$</td>
<td>Rice (LKC)</td>
<td>[37]</td>
</tr>
<tr>
<td>VI.7</td>
<td>$R_{1385}/R_{1705}$</td>
<td>Rice (LKC)</td>
<td>[37]</td>
</tr>
<tr>
<td>VI.8</td>
<td>$(R_{1705} - R_{1385})/(R_{1385} + R_{780})$</td>
<td>Rice (LKC)</td>
<td>[37]</td>
</tr>
<tr>
<td>VI.9</td>
<td>$(R_{1705} - R_{1385})/(R_{1705} + R_{1385} - 2 \times R_{790})$</td>
<td>Rice (LKC)</td>
<td>[37]</td>
</tr>
<tr>
<td>VI.10</td>
<td>$(FDR_{1430} - FDR_{1295})/(FDR_{1430} + FDR_{1295})$</td>
<td>Rice (LKC)</td>
<td>[37]</td>
</tr>
<tr>
<td>VI.11</td>
<td>$R_{1385}/FDR_{1295}$</td>
<td>Rice (LKC)</td>
<td>[37]</td>
</tr>
<tr>
<td>VI.12</td>
<td>$(R_{523} - R_{583})/(R_{523} + R_{583})$</td>
<td>Pasture (LKC)</td>
<td>[56]</td>
</tr>
<tr>
<td>VI.13</td>
<td>$(R_{780} - R_{680})/(R_{780} + R_{680})$</td>
<td>Olive orchards (LKC)</td>
<td>[21]</td>
</tr>
<tr>
<td>VI.14</td>
<td>$FDR_{760}/FDR_{930}$</td>
<td>Sainfoin pasture (PKC)</td>
<td>[57]</td>
</tr>
<tr>
<td>VI.15</td>
<td>$R_{780}/R_{680}$</td>
<td>Sainfoin pasture (PKC)</td>
<td>[57]</td>
</tr>
<tr>
<td>VI.16</td>
<td>$(R_{2275} - R_{1875})/(R_{2275} + R_{1875})$</td>
<td>Wheat (LKC)</td>
<td>[58]</td>
</tr>
<tr>
<td>VI.17</td>
<td>$FDR_{1665}/FDR_{2250}$</td>
<td>Wheat (LKC)</td>
<td>[58]</td>
</tr>
<tr>
<td>VI.18</td>
<td>$(R_{2275} - R_{1875})/(R_{2275} + R_{1875} - 2R_{762})$</td>
<td>Wheat (LKC)</td>
<td>[58]</td>
</tr>
<tr>
<td>VI.19</td>
<td>$(R_{2280} - R_{780})/(R_{1875} - R_{780})$</td>
<td>Wheat (LKC)</td>
<td>[58]</td>
</tr>
<tr>
<td>VI.20</td>
<td>$(R_{935} - R_{770})/(R_{935} + R_{770})$</td>
<td>Wheat (PKC)</td>
<td>[58]</td>
</tr>
<tr>
<td>VI.21</td>
<td>$FDR_{1715}/FDR_{690}$</td>
<td>Wheat (PKC)</td>
<td>[58]</td>
</tr>
<tr>
<td>VI.22</td>
<td>$(R_{935} - R_{770})/(R_{935} + R_{770} - R_{1395})$</td>
<td>Wheat (PKC)</td>
<td>[58]</td>
</tr>
<tr>
<td>VI.23</td>
<td>$(NR_{2275} - NR_{445})/(NR_{1875} - NR_{445})$</td>
<td>Wheat (LKC)</td>
<td>[23]</td>
</tr>
<tr>
<td>VI.24</td>
<td>$(NR_{2275} - NR_{1875})(NR_{2275} + NR_{1875} - 2NR_{445})$</td>
<td>Wheat (LKC)</td>
<td>[23]</td>
</tr>
<tr>
<td>VI.25</td>
<td>$(NR_{1400} - 1/NR_{2250}) \times NR_{2200}$</td>
<td>Wheat (LKC)</td>
<td>[23]</td>
</tr>
<tr>
<td>VI.26</td>
<td>$(NR_{2265} - NR_{1880})(NR_{2265} + NR_{1880})$</td>
<td>Wheat (LKC)</td>
<td>[23]</td>
</tr>
<tr>
<td>VI.27</td>
<td>$(NR_{2300} - NR_{2250})(NR_{1850} - NR_{2300})$</td>
<td>Wheat (LKC)</td>
<td>[23]</td>
</tr>
<tr>
<td>VI.28</td>
<td>$(NR_{1850} - NR_{2300})(NR_{1850} + NR_{2300} - 2NR_{2350})$</td>
<td>Wheat (LKC)</td>
<td>[23]</td>
</tr>
<tr>
<td>VI.29</td>
<td>Maximum absorption depth of absorption peak of CRR</td>
<td>Rice</td>
<td>[59]</td>
</tr>
<tr>
<td>VI.30</td>
<td>Absorption peak area of CRR</td>
<td>Rice</td>
<td>[60]</td>
</tr>
<tr>
<td>VI.31</td>
<td>Symmetry of CRR</td>
<td>Rice</td>
<td>[60]</td>
</tr>
<tr>
<td>VI.32</td>
<td>Area normalized maximum absorption depth of CRR</td>
<td>Rice</td>
<td>[40]</td>
</tr>
<tr>
<td>VI.33</td>
<td>The normalized band depth ratio</td>
<td>Pasture (PKC)</td>
<td>[51]</td>
</tr>
<tr>
<td>VI.34</td>
<td>The normalized band depth index</td>
<td>Pasture (PKC)</td>
<td>[51]</td>
</tr>
</tbody>
</table>

2.3.3. Random Forest Feature Selection (RFFS)

An RF is a decision tree-based ensemble learning algorithm proposed by Breiman in 2001 [61]. To create a single decision tree, it bags together randomized samples (extracted using the bootstrap method) from the original data. A split point is chosen for splitting at each node of the decision tree using a random feature subspace. Following a majority vote, these decision trees are joined to produce the final forecast outcome.

To reduce the dimensionality of the data, the RFFS is based on the average error rate computation. In this study, the split point (mtry) value was set to 5; there were 1000 decision trees (ntree), and the training set was rectified by randomly choosing 2/3 of the total data set for replacement [45,59]. The out-of-bag error rate was estimated using the remaining third, often known as out-of-bag (OOB) data. The increase in the OOB means square error (%) could indicate the significance of the chosen variable when the variable was modified.
2.4. Modeling and Evaluation of the LKC

Three machine learning algorithms, namely random forests (RF), support vector machines (SVM), and backpropagation (BP) neural networks, were applied for the regression analysis of rice LKC. The penalty factor “c” and the radial basis function parameter “g” were set as 0.1 and 10 in the SVM model, respectively. The number of hidden nodes, epochs, training rate, and training goals were set as 5, 1000, 0.01, and $1 \times 10^{-5}$ in the BP neural network, respectively. The 276 samples from Exp.1 and Exp.2 were divided into a calibration dataset of 184 (66.7%) samples and a validation dataset of 92 (33.3%) samples, which were used in models that did not require random sampling. The 46 samples from Exp.3 were used to evaluate the performance of prediction models. The predictive accuracy of the model was evaluated by the coefficient of determination ($R^2$) and the root mean square error (RMSE), which have been proven to be effective in the evaluation of most models [62]. The calculation formulas are as follows:

$$R^2 = \frac{\sum_{i=1}^{N}[f(x_i) - \bar{y}]^2}{\sum_{i=1}^{N}[y_i - \bar{y}]^2},$$  

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}[f(x_i) - y_i]^2}{N}},$$  

where $f(x_i)$, $y_i$, and $\bar{y}$ are the predicted values, the observed values, and the average of the observed values of LKC in rice, respectively; $N$ is the number of samples.

3. Results
3.1. Dynamic Change in Spectral Reflectance

Figure 3 shows the dynamic change in the average reflectance of rice leaves under different treatments, and the reflectance of several key bands is used to reflect the response of the spectra to the condition change. The leaves of K1 in Figure 3a,b can reflect more light at the two reflection platforms of 800–1000 nm and 1650–1800 nm. The reflectance of N1–N3 increases progressively, with K3 being the lowest. The reflectance of K3 is higher and the difference in reflectance between N1 and N3 is less noticeable at the two peaks of 550 and 2200 nm. The consecutive increases in the LKC concentration of K1–K3 in Figure 3d,e show that increasing potassium fertilizer can facilitate potassium absorption in rice leaves. The LKC concentration of N1–N3 exhibits a trend of gradually increasing (2021) and initially increasing and then dropping (2022), suggesting that sufficient nitrogen fertilizer can aid in the absorption of potassium in leaves but excessive nitrogen fertilizer may impede potassium absorption. This might be brought on by ammonium ions inhibiting potassium uptake in plant roots under conditions of high nitrogen concentration [63]. The LKC of rice steadily declines with rice growth, which is consistent with the changing trend of the leaf reflectance at 2200 nm, as shown in Figure 3c,f.

3.2. Relationship between the LKC and the Transformed Spectra

The relationship between R, FDR, CRR, NR, and rice LKC is shown in Figure 4. The values on both sides of the dashed line are significant at the significance level of 0.01. When the wavelength is greater than 690 nm, the reflectance of R is negatively correlated with LKC. It is significantly improved after 1263 nm, and two troughs are formed at 1410 and 1881 nm; the maximum correlation coefficient reaches $-0.593$ and $-0.583$. The correlation coefficients at 52 and 2214 nm exceeded 0.6, measuring 0.61, 0.66, $-0.663$, $-0.667$, $-0.61$, and $-0.696$, respectively. The change curve of the CRR rises sharply in the range of 2142–2211 nm to the maximum, which is 0.603, and forms two troughs at 1404 and 1878 nm, and the correlation coefficients are $-0.538$ and $-0.555$. The two trough positions of NR are close to CRR, and the maximum negative correlation appears at 1404 nm with $-0.578$ and 1881 nm with $-0.563$. The only spectrum with a stronger correlation in the visible
range is NR, which has a correlation value of 0.412 at 731 nm when compared to the other three transformed spectra’s −0.226, −0.288, and −0.319.

3.3. Screening of Variables
3.3.1. Single Band Variable (BV) Screening

As a single band variable (BV) set for rice LKC prediction, Table 4 contains all bands with linear correlation coefficients greater than 0.5 between the four transformed spectra and rice LKC. Regarding all four types of transformed spectra, most of the characteristic bands are found in the SWIR, where the R, CRR, and NR at 1382–1480 nm, the R and FDR at 1996–2003 nm, and the FDR and CRR at 2205–2216 nm are all sensitive to rice LKC. Each of the four transformed spectra in the 1884–1888 nm region is responsive to the change in rice LKC. The link between FDR and rice LKC is the strongest in terms of the correlation coefficient. The number of bands with a correlation coefficient greater than 0.5 is 309, which is the highest, and NR is 129, which is the lowest. The correlation
between the four transformed spectra and rice LKC is FDR > R > CRR > NR. The NR has the lowest variable compression rate at 5.86%, which means that the proportion of redundant bands removed from the complete reflectance is the highest from the perspective of the variable compression effect. In addition, the accuracy of single band variables is evaluated by $R^2$ and RMSE. The ranking of the transformed spectra characteristic bands is NR > R > FDR > CRR.

Table 4. All single bands with linear correlation coefficients greater than 0.5 between the four transformed spectra and rice LKC.

<table>
<thead>
<tr>
<th>Reflectance</th>
<th>Characteristic Bands</th>
<th>Number of Characteristic Bands</th>
<th>Variable Compression Rate (%)</th>
<th>$R^2$</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>FDR</td>
<td>2182–2222, 2238–2249, 2412, 2414, 2418, 2419, 2421, 2423–2430, 2432–2443, 2445–2448, 2450–2452, 2454, 2458–2461, 2463, 2470, 2471 nm</td>
<td>309</td>
<td>14.04</td>
<td>0.62314</td>
<td>0.17481</td>
</tr>
<tr>
<td>CRR</td>
<td>1382–1497, 1865–1888, 2205–2216 nm</td>
<td>152</td>
<td>6.91</td>
<td>0.58752</td>
<td>0.18562</td>
</tr>
<tr>
<td>NR</td>
<td>1381–1480, 1865–1893 nm</td>
<td>129</td>
<td>5.86</td>
<td>0.67974</td>
<td>0.16979</td>
</tr>
</tbody>
</table>

3.3.2. Vegetation Index Variable (IV) Screening

Figure 5 and Table 5 illustrate the results of the principal component analysis of vegetation indicators. It was found that there were six principal components with eigenvalues greater than one, namely PC1, PC2, PC3, PC4, PC5, and PC6. According to the weight coefficients, the first 20 vegetation indices corresponding to these PCs were listed and their accuracy was evaluated. Based on the result of $R^2$ and RMSE, the ranking of the six PCs is PC1 > PC5 > PC4 > PC2 > PC6 > PC3. Therefore, the 20 component vegetation indices of PC1 were used as the results of the IV variable screening.

Figure 5. PCs with their eigenvalues.

3.3.3. Importance Assessment of Variables by RFFS

Out of the four transformed spectral characteristic bands, the BV with the highest correlation coefficients is chosen to form a total of variables set with the IV. The RFFS method is used to evaluate the importance of variables to achieve the purpose of second feature screening. Finally, the top 40 variables ranked by importance are extracted (Figure 6) as mixed variables (MV). The importance of the VI is usually higher than that of a single band.
Among all the mentioned VIs, the ratio of the VIs based on the four transformed spectra, namely R, FDR, CRR, and NR, is 6:2:3:2, and the two highest ranked indices are \((\frac{R_{2275} - R_{1875}}{R_{2275} + R_{1875}})\) and \((\frac{R_{2275} - R_{1875}}{R_{2275} + R_{1875} - 2R_{1875}})\) on the R, followed by \((\frac{NR_{2275} - NR_{1875}}{NR_{2275} + NR_{1875} - 2NR_{1875}})\) and \((\frac{NR_{2275} - NR_{1875}}{NR_{1875} - NR_{1875}})\) based on the NR. The relative error rates of the four OOBs exceed 0.6. The absorption peak area of the CRR and the normalized band depth ratio are ranked 10th and 16th, respectively. Single bands make up more than the VIs, accounting for 67.5% of all the MV. The ratio of the single band number based on the four transformed spectra is 8:4:9:5. The 1407 nm of the CRR, the 1409 nm of NR, and the 1414 nm of the R reflectance rank highest among all single bands of the four transformed spectra, indicating that the 1407–1414 nm reflectance is the most sensitive to the response of rice LKC changes.

### Table 5. PCs with their component VIs.

<table>
<thead>
<tr>
<th>PCs</th>
<th>VIs</th>
<th>R²</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>PC1</td>
<td>VI.1, VI.6, VI.4, VI.5, VI.18, VI.17, VI.25, VI.15, VI.22, VI.19, VI.23, VI.3, VI.10, VI.16, VI.34, VI.33, VI.24, VI.30, VI.32, VI.11</td>
<td>0.7237</td>
<td>0.18265</td>
</tr>
<tr>
<td>PC2</td>
<td>VI.27, VI.26, VI.28, VI.24, VI.21, VI.19, VI.2, VI.1, VI.34, VI.33, VI.3, VI.23, VI.22, VI.10, VI.13, VI.19, VI.15, VI.20, VI.25, VI.16</td>
<td>0.71854</td>
<td>0.18967</td>
</tr>
<tr>
<td>PC3</td>
<td>VI.27, VI.29, VI.12, VI.32, VI.30, VI.14, VI.31, VI.1, VI.33, VI.34, VI.28, VI.19, VI.21, VI.25, VI.11, VI.3, VI.15, VI.19, VI.22, VI.20</td>
<td>0.70674</td>
<td>0.19237</td>
</tr>
<tr>
<td>PC4</td>
<td>VI.27, VI.11, VI.2, VI.16, VI.19, VI.21, VI.14, VI.1, VI.33, VI.31, VI.12, VI.28, VI.30, VI.13, VI.10, VI.23, VI.20, VI.19</td>
<td>0.72114</td>
<td>0.18562</td>
</tr>
<tr>
<td>PC5</td>
<td>VI.27, VI.19, VI.21, VI.31, VI.20, VI.11, VI.10, VI.19, VI.28, VI.24, VI.26, VI.16, VI.17, VI.2, VI.29, VI.23, VI.11, VI.18, VI.14, VI.13</td>
<td>0.72113</td>
<td>0.18365</td>
</tr>
<tr>
<td>PC6</td>
<td>VI.27, VI.18, VI.20, VI.11, VI.14, VI.19, VI.21, VI.12, VI.16, VI.10, VI.9, VI.28, VI.29, VI.33, VI.12, VI.13, VI.34, VI.32, VI.30, VI.26</td>
<td>0.70712</td>
<td>0.19122</td>
</tr>
</tbody>
</table>

![Figure 6. Variable importance is ranked by random forest regression models, as measured using the percentage increase in OOB error (%).](image)

### 3.4. Prediction of Potassium Content

#### 3.4.1. Potassium Content Prediction Based on BV

As shown in Figure 7, the FDR model provides the best predictive effect among all BV-based potassium content prediction outcomes. The validation R² of the FDR–RF, the FDR–SVM, and the FDR–BP are 0.62301, 0.58347, and 0.59979, and the RMSE are 0.24174, 0.18898, and 0.25767, respectively (Figure 7d–f). Followed by the NR–BP and CRR–BP (Figure 7i), the validation R² are 0.59046 and 0.57752, and the RMSEs are 0.24017 and 0.21914. The prediction accuracy of the R–RF is the lowest among all models, and the validation R² is only 0.44813 (Figure 7a), which is the only one with an R² lower than 0.5, indicating that the rice LKC predicted by the model is quite different from the actual value and the model is not suitable for the regression of rice LKC analysis. From the perspective of model selection, the optimal models corresponding to the four transformed spectra are...
the R–SVM, the FDR–RF, the CRR–BP, and the NR–BP. The model combined with the BP neural network and the BV of the four transformed spectra shows a better predictive ability of rice LKC (Figure 7b,d,i,l).

**Figure 7.** Scatterplots, coefficient of determination ($R^2$), root mean square error (RMSE), fitting curve, and the number of samples (N) of predicted LKC using the single band variables of R (a–c), FDR (d–f), CRR (g–i), NR (j–l), and the RF, the SVM, and the BP neural network, respectively.

### 3.4.2. Potassium Content Prediction Based on IV

Figure 8a–c illustrate the prediction effects of the three regression models IV–RF, IV–SVM, and IV–BP on rice LKC. The $R^2$ of the validation are 0.7387, 0.71424, and 0.72311, and the RMSEs are 0.15045, 0.17362, 0.15182, respectively.
the lowest RMSE. The specific performance is that the prediction ability of extreme values is weak and the dispersion of data distribution at both ends is high. Compared with the IV–SVM, the IV–BP has a close RMSE, but the $R^2$ of the latter is significantly higher than the former, so the predictive ability of the three models for rice LKC is RF > BP > SVM.

Figure 7. Scatterplots, coefficient of determination ($R^2$), root mean square error (RMSE), fitting curve, and number of samples (N) of predicted LKC using mixed variables and the RF (a), the SVM (b), the BP (c) neural network, respectively.

3.4.3. Prediction of Potassium Content Based on MV

In the regression model based on MV, the prediction accuracy of the MV–RF, the MV–SVM, and the MV–BP for rice LKC improves as compared with the model based on the BV and the IV (Figure 9a–c). The validation $R^2$ reaches 0.77817, 0.7377, and 0.75566, and the RMSE is 0.14913, 0.15617, and 0.15644, respectively. Figure 9 shows that all models are more accurate in predicting LKC in the range of 1.0–2.2%, while the prediction deviation of LKC in the range of less than 0.1% and greater than 2.2% is relatively large. The MV–RF has the highest $R^2$ and the lowest RMSE, which is the best among all rice LKC regression models (Figure 9a).

Figure 8. Scatterplots, coefficient of determination ($R^2$), root mean square error (RMSE), fitting curve, and number of samples (N) of predicted LKC using vegetation index variables and RF (a), SVM (b), the BP (c) neural network, respectively.

3.5. Evaluation and Testing of the MV–RF Model

The 46 samples from Exp.3 were used to evaluate and test the accuracy of the MV–RF model, and the linear regression and prediction errors are shown in Figure 10. The $R^2$ and the RMSE of the prediction model were 0.77593 and 0.17313 (Figure 10a), which was approximately equal to $R^2$ and had a 16.1% increase in RMSE when compared with that in Figure 9a, respectively. It indicated that the MV–RF model could effectively improve the prediction accuracy of rice LKC.
water status of leaves (water content, pressure potential, osmotic potential) responded was feasible to establish a spectral diagnosis mode for potassium deficiency through the characteristic bands. These differences in characteristic bands can provide an important way to distinguish potassium deprivation might diminish the activity of water channel proteins, and hydraulic conductance and water supply in plants were suppressed [18,66].

In this study, the random forest feature selection method was used to screen the variables in the BV and the IV to build the MV. The single bands from the BV were primarily dispersed in the ranges of 1402–1428 nm and 1871–1907 nm from the distribution of significant bands. The location of the potassium sensitive band is extremely close to the peaks of the water absorption characteristic (1450 and 1950 nm); this is consistent with previous conclusions from ecological research about variations in potassium levels impacting the laws of plant water absorption. It has been proposed that potassium ions in plant cells might modify leaf osmotic pressure to control stomatal conductance and, consequently, leaf water content [13,19,37]. Besford discovered that tomato leaves retained more water in the foliage at high potassium levels when compared to low potassium levels, leading to greater leaf fresh weight [64]: Mengel proposed that potassium significantly affected the leaves (particularly young leaf) of cucumber, sunflower, and wheat. The water status of leaves (water content, pressure potential, osmotic potential) responded more sensitively to potassium supply than dry matter production [65]. When plants were subjected to potassium stress, water uptake by crop root cells would also decline since potassium deprivation might diminish the activity of water channel proteins, and hydraulic conductance and water supply in plants were suppressed [18,66].

4. Discussion

4.1. Differences in Nitrogen and Potassium Sensitive Bands

The key to establishing the relationship between potassium and rice leaf spectra is to obtain characteristic bands that are different from other elements. Lu [37] considered that it was feasible to establish a spectral diagnosis mode for potassium deficiency through the combination of SWIR, which was different from nitrogen deficiency and other nutrients. In this study, comparing rice LKC under different nitrogen and potassium levels, Figure 3a,c show that the spectral characteristics of potassium-deficient rice leaves are significantly different from those under nitrogen deficiency in the visible light and SWIR bands. The reflectance of potassium-deficient rice leaves is higher than that of normal leaves near 700 nm and it decreases near 2200 nm in the SWIR bands. The SWIR of nitrogen-deficient rice leaves is lower than that of normal leaves and the difference is not obvious around 2200 nm. These differences in characteristic bands can provide an important way to distinguish potassium and nitrogen deficiency in rice by leaf spectral analysis.

4.2. Spectral Characteristic Bands of Potassium

In this study, the random forest feature selection method was used to screen the variables in the BV and the IV to build the MV. The single bands from the BV were marked in Figure 11 among all the variables included in the MV. The single bands are primarily dispersed in the ranges of 1402–1428 nm and 1871–1907 nm from the distribution of significant bands. The location of the potassium sensitive band is extremely close to the peaks of the water absorption characteristic (1450 and 1950 nm); this is consistent with previous conclusions from ecological research about variations in potassium levels impacting the laws of plant water absorption. It has been proposed that potassium ions in plant cells might modify leaf osmotic pressure to control stomatal conductance and, consequently, leaf water content [13,19,37]. Besford discovered that tomato leaves retained more water in the foliage at high potassium levels when compared to low potassium levels, leading to greater leaf fresh weight [64]: Mengel proposed that potassium significantly affected the leaves (particularly young leaf) of cucumber, sunflower, and wheat. The water status of leaves (water content, pressure potential, osmotic potential) responded more sensitively to potassium supply than dry matter production [65]. When plants were subjected to potassium stress, water uptake by crop root cells would also decline since potassium deprivation might diminish the activity of water channel proteins, and hydraulic conductance and water supply in plants were suppressed [18,66].
4.3. VIs of Potassium Related

Figure 11 demonstrated that the combination of bands from the IV were mainly distributed in the ranges of 1799–1881 nm and 2276–2350 nm. Moreover, 20 of 34 indices were reserved based on the PCA because they had a great eigenvalue. These VIs had a strong correlation with the visible light band and contained few SWIR bands. Figure 11 and Table 5 indicate that vegetation indices based on the visible region and the combination of visible and near-infrared bands in previous studies were not effective for potassium monitoring in this study. This result was different from the conclusions of Gómez-Casero [21], Kawamura [56], and Albayrak [57]. They hold the view that these VIs, including the visible light band, were also significantly correlated with the LKC of plants, and a linear model was developed to link the VIs with the LKC. The reason for this phenomenon was the absence of the SWIR from the initial measurement range or the result of differences between the canopy and leaf levels. In conclusion, these indices did not play a role in this study, and further research was needed to develop the generality of VI. Therefore, potassium-sensitive bands in this study can be used as a basis to provide some reference for the establishment of potassium-related VIs in the future.

4.4. Comparison of Transformed Spectra

The comprehensive analysis of the single band variable, the index variable, and the LKC regression model of several transformed spectra showed that the spectral transformation process could promote the improvement of LKC prediction ability. The results of the linear correlation analysis showed that the FDR had the best correlation with rice LKC, which directly led to the regression model based on the single band screened by the FDR performing the best among all four transformed spectral prediction models. This is consistent with the results of the effect of the FDR on plant potassium content, which was conducted by Chen [34], Peng [35], and Azadnia [50]. The advantage of the NR was reflected in the VIs based on the NR. In the process of variable screening, three of the six VIs based on the NR were retained. On the one hand, it benefited from the good removal effect of the NR on redundant information; on the other hand, it was closely related to the establishment of the VIs and the selection of characteristic bands. Most of these bands were within the sensitive range in Figure 11. The CRR significantly increased the importance of single band variables. The prediction results of several models based on the CRR for rice LKC were relatively balanced, and it was the most stable among the four transformed spectra.

4.5. Variable Importance Evaluation with RFFS

In this study, the linear correlation analysis method was used to conduct the primary screening of the characteristic single bands and the VIs of the four transformed spectra, respectively. Then, RFFS was applied to carry out the second screening of the variable optimization. There were two main reasons: First, the traditional linear model had a single choice of variables, while the RFFS used all variables as an indiscriminate sampling
set and considered the role of variables rather than the source, which was an important guarantee for maximizing the regression effect of the model. Second, high-dimensional data contained many useless features. The selection of the RFFS subspace cannot guarantee the effectiveness of the selected features and the quality of the constructed feature subset. When many useless feature subsets were used to build trees, the performance of a single decision tree was greatly reduced, and then, the overall performance of the random forests was affected. The studies of Jiang [45], Viljanen [47], and Karunaratne [48] all applied different types of variables to the random forests to obtain better prediction results than single variables. The results of this study showed that the MV dataset obtained by applying RFFS second screening as compared with the single variable set did improve the prediction accuracy of rice LKC, which is verified by previous research.

4.6. Evaluations of Modeling Variables and Methods

In this study, three machine learning methods of RF, SVM, and BP neural networks were combined with three different types of variables to conduct regression analysis. Among them, the MV–RF obtained the highest prediction accuracy and was the best prediction model for rice LKC under this study’s conditions. In general, the RF has a better performance on the IV and the MV than the BV. In some fitting results, it was found that the selected regression models had a weaker predictive ability for the extremum, with data distribution at both ends of the x-axis displaying higher dispersion. During the experimental measurement process, it was inevitable to obtain some high or low data, which were still within the reasonable range during various stages of plant growth. These experimental data were real and effective. For the extremum, non-linear modeling could be used for a better fit. In the process of regression analysis and data fitting, the number of the extremum accounted for a very small proportion of the entire dataset, which was less than 3%. Therefore, the linear fitting method that was applicable to most normal data was applied in this study.

5. Conclusions

This study verified the feasibility of the RFFS method in the prediction of potassium in rice leaves. Using this method, 80 single bands with the highest linear correlation coefficient and 20 VIs constituting the highest eigenvalue PCs were used for a second screening and then formed the MV. The optimal MV had a higher predictive ability than the BV and the IV. The three methods of the RF, the SVM, and the BP neural network were selected for regression analysis, and it was found that the processing ability of the RF for the IV and the MV was better than that of the BV. In addition, this study compared the R with the FDR, the CRR, and the NR. The FDR has the largest number of characteristic bands. The CRR has the strongest correlation. The VIs based on the NR have the strongest predictive ability. The comprehensive analysis of important bands in all models reveals that it is of great importance that the single band lies in the ranges of 1402–1428 nm and 1871–1907 nm and that the VI is composed of bands in the ranges of 1799–1881 nm and 2276–2350 nm, which provides a reference for the establishment of a universal VI related to potassium.

Author Contributions: Conceptualization, Y.Y.; methodology, Y.Y.; software, Y.Y.; validation, X.L.; formal analysis, X.L.; investigation, X.L.; resources, H.Y. and Y.S.; data curation, Y.Y. and X.L.; writing—original draft preparation, Y.Y.; writing—review and editing, H.Y., L.Z. and Y.S.; visualization, Y.Y.; supervision, L.Z.; project administration, Y.S.; funding acquisition, Y.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (32001418), the program of Jiangsu Province and Education Ministry Cosponsored Synergistic Innovation Center of Modern Agriculture Equipment (XTCX1006), the Jilin Province Science and Technology Development Program (20200402015NC).

Data Availability Statement: Data will be made available on request.
Acknowledgments: We are grateful for support from the projects listed above for their generous support of this work. We also thank Minghao Qu, Jingjing Guo, and Junhe Zhang for their help in rice experiments and data collection.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References


**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.