ISSN 1392-3196 / e-ISSN 2335-8947
Zemdirbyste-Agriculture, vol. 108, No. 2 (2021), p. 181–190
DOI 10.13080/z-a.2021.108.024

Machine learning-based estimation of potato chlorophyll content at different growth stages using UAV hyperspectral data

Changchun LI[1], Chunyan MA[1], Peng CHEN[1,2], Yingqi CUI[1], Jinjin SHI[1], Yilin WANG[1]

[1]Henan Polytechnic University, School of Surveying and Land Information Engineering
Jiaozuo, Henan 454000, China
E-mail: mayan@hpu.edu.cn

[2]GZH-HNJ BDS AGR Co. Ltd.
Zhengzhou 450000, China

Abstract

Accurate estimation of chlorophyll (Chl) content is highly significant in monitoring potato growth and improving yield and quality. Fractional differentiation can refine the local information of the spectrum and is conducive to the removal of background noise. In this study, a new method to examine the effects of fractional differentiation on the estimation of the Chl content of crops was developed. Potato (*Solanum tuberosum* L.) was selected as the research object. A fractional derivative was used for unmanned aerial vehicle (UAV) hyperspectral data processing, and an algorithm for estimating the potato Chl content was studied.

The results concluded that the correlation increased after first declining with increasing differential order; the maximum absolute values of the correlation coefficient at different stages were obtained with 1-order differentiation at the budding stage, 0.6-order differentiation during the tuber formation and tuber growth stages and 1.2-order differentiation at the starch accumulation stage. The comparison and analysis of the estimation models of the potato Chl content at different growth stages showed that the support vector machine (SVM) model had the greatest accuracy in estimating the potato Chl content with an R^2 value of 0.83 at the budding stage, followed by R^2 of 0.80 at the tuber forming stage.

Key words: chlorophyll content, fractional differentiation, UAV hyperspectrum, *Solanum tuberosum*.

Introduction

Potatoes are hardy, drought tolerant and barren soil resistant. They are easily planted in both southern and northern China and can meet the self-sufficiency requirements of food crops in the future. To ensure global food security, China's government launched the Potato Staple Food Project. In 2016, the Ministry of Agriculture published the "Guidance on promoting the development of the potato industry"; therefore, large-scale cultivation of potato as a staple food in China is inevitable. The real-time monitoring of the nutritional status of potato crops promotes informed and efficient planting and fertilisation strategies that promote high yields and maximise productivity.

Chlorophyll (Chl) is an important pigment for the light energy utilisation of crops that directly affects the energy and material conversion and transmission process of crops. Changes in the Chl content (C) directly represent the photosynthetic capacity, growth and nutritional state of crops (Gonzalez-Dugo et al., 2015; Sonobe et al., 2021). Moreover, ChlC and growth conditions of potatoes are highly correlated, and Chl is an important index for measuring potato nutritional status (Tilahun et al., 2020). The rapid and accurate monitoring of ChlC is highly significant in the monitoring of the photosynthetic capacity and growth status of potatoes, and guides the improvement and optimization of potato yield and quality (Liu et al., 2020). The traditional methods of ChlC monitoring are indoor high-performance liquid chromatography, atomic absorption spectrometry and spectrophotometry. These methods can accurately measure ChlC but are destructive, unrecoverable and cumbersome with large workloads and other shortcomings and cannot achieve real-time large-scale regional monitoring.

Hyperspectral remote sensing technology has a high spectral resolution and strong band continuity, which are major achievements in the field of earth observation and are at the frontier of remote sensing science and technology. Because the sunlight absorption and reflection characteristics of Chl in crops form a unique spectral curve, rapid and non-destructive high throughput monitoring of ChlC in crops is possible by analysing the hyperspectral characteristics of crops (Li et al., 2020). Unmanned aerial vehicle (UAV) remote sensing technology is fast and high throughput, with low costs, high ease of operation and high resolution. Furthermore, UAV remote sensing has the additional advantage of

Please use the following format when citing the article:
Li C., Ma C., Chen P., Cui Y., Shi J., Wang Y. 2021. Machine learning-based estimation of potato chlorophyll content at different growth stages using UAV hyperspectral data. Zemdirbyste-Agriculture, 108 (2): 181–190. DOI 10.13080/z-a.2021.108.024

being suitable for regional research. Therefore, this paper researches the estimation of potato ChlC using a UAV platform equipped with a hyperspectral camera to provide a scheme for monitoring potato ChlC.

At present, domestic and foreign scholars have made some achievements in estimating the ChlC of crops based on spectral information. Li et al. (2019) estimated the ChlC of winter wheat using a multivariate regression model. Singh et al. (2017) estimated the ChlC of sorghum by using the reflectance composition ratio model at bands of 595, 1676, 595 and 508 nm. Roosjen et al. (2017) determined ChlC at different growth stages of potato by using UAV remote sensing data. Liang et al. (2016) used the correlation between the spectral reflectivity and ChlC of apple trees to select sensitive bands and then used multivariate linear regression, neural networks and a principal component analysis (PCA) method to estimate ChlC.

Piegari et al. (2021) took *Sporobolus densiflorus* as the research object and studied the estimation of leaf area index (LAI) and Chl content using hyperspectral data (HD) coupled with PROSAIL model. Malin Hoeppner et al. (2020), Zhu et al. (2020) and El-Hendawy et al. (2021) used hyperspectral reflectance to construct vegetation index, analysed the relationship between it and canopy Chl content and established an estimation model for Chl content of crops and trees. An et al. (2020) and Yamashita et al. (2020) used the reflectivity of HD and machine learning algorithm to build the estimation model of leaf green content of tea and rice.

Most of the existing studies on the estimation of ChlC using hyperspectral remote sensing technology are based on the relationship between the original spectral characteristics and the ChlC of crops or use the original spectral information to construct a correlation index, analyse the relationship between the original spectral characteristics and the ChlC of crops, and construct an estimation model of the ChlC of crops (Cordon et al., 2016; Croft et al., 2020; Guo et al., 2020; Morley et al., 2020; Qiao et al., 2020; Zhang et al., 2020; Zhou et al., 2020). Spectral differentiation techniques can partially eliminate the influence of atmospheric effects, vegetation shadows and soil and reflect the essential characteristics of vegetation (Afshari et al., 2020).

In recent years, increasing attention has been paid to the use of spectral differential technology to monitor the growth of crops, and some research results on this subject have been obtained. Xu et al. (2019) and Bahrami and Mobasheri (2020) used remote sensing data to carry out 1-order, 2-order and 3-order differential processing on the original data and combined with time series normalized difference vegetation index (NDVI) to realize the identification of the dominant tree species in the forest. Liu et al. (2019) used hyperspectral 1-order differential technology to monitor maize leaf spot disease and achieved good results. He et al. (2016) and Yang et al. (2020) analysed the correlation between the different order spectral differential indices and nitrogen (N) content of crop and selected the spectral differential indices with strong correlation to build the N content estimation model, which achieved good results. Pereira da Conceição et al. (2020) successfully identified two mycotoxicogenic *Fusarium* species associated with maize based on hyperspectral differentiation technology. Basinger et al. (2020) used HD to analyse the effect of phenology on the differentiation of crop and weed species through differential processing.

Fractional-order differentiation can refine spectral information, make full use of information easily ignored by integer order differentiation, effectively remove image background noise and deeply explore potential information in the spectrum (Cardone, Conte, 2020; Ali et al., 2021). In recent years, fractional differentiation has been widely studied and applied in many fields. Lin et al. (2019) used HD to perform fractional differential processing to build a model to estimate the metal zinc content in the soil. Through analysing the spectral characteristics of differential of different orders, Zhu et al. (2019) obtained the most appropriate differential orders of Co^{2+} and Cu^{2+} respectively based on the multi-objective particle swarm optimization algorithm. Zhang et al. (2016) and Wang et al. (2017) discussed the possibility of using fractional differential technology to estimate soil heavy metal chromium and salt content in HD. The results show that the accuracy and robustness of the model after fractional pre-processing is better than integer order. The differential has been greatly improved. However, there are still relatively few studies on fractional-order differential technology in the field of crop nutrition monitoring.

To verify the effectiveness of fractional differentiation in the hyperspectral estimation of crop ChlC, a new technical method was developed. In this paper, a fractional-order differential algorithm for spectral data processing and carries out the estimation research of potato ChlC was introduced. To obtain 11 types of fractional differentiation spectral data of the canopy at each growth stage, the spectral data of the layer height of the potato canopy in the 454–950 nm band by differential processing of order 0–2 (0.2-order interval) were treated.

Through Pearson correlation analysis, fractional differential spectra with an excellent correlation between the growth stages and potato ChlC were selected, and the correlation coefficients between these spectra and potato ChlC were calculated. Different differential bands were optimized, and a multiple linear regression (MLR), support vector machine (SVM) and random forest (RF) models of the potato ChlC based on the fractional differential spectra were established. The accuracy of the models was verified, and the optimal estimation model was selected.

Materials and methods

Study area and test design. The National Precision Agriculture Research and Demonstration Base of Xiaotangshan, Changping district, Beijing, China was selected as the research area (Figure 1). This area is located in the northeast region of Xiaotangshan, with boundaries ranging from 40°1031″–40°1118″ N lat. to 116°26′10″–116°27′05″ E long. The research site lies in a northern temperate semi-humid continental monsoon climate with an average elevation of 36 m.

As shown in Figure 2, a planting density (T area) and a nitrogen (N) fertiliser (N area) test areas were established at the experimental site using a completely random design. Three levels of planting density: 280 plants ha^{-1} (T1), 320 plants ha^{-1} (T2, control treatment) and 360 plants ha^{-1} (T3) were used. The planting density experiment comprised a total of 6 treatments (3 densities and 2 cultivars) with 3 repetitions per treatment for a total of 18 experimental communities. In the N fertiliser experiment, 4 levels of N: 0 kg ha^{-1} (N0), 0.72 kg ha^{-1} (N1), 1.45 kg ha^{-1} (N2, control treatment) and 2.12 kg ha^{-1} (N3) were used; 1/2 of each level was applied as basal fertilizer, while the other half was applied as jointing fertilizer. The N experiment contained a total of 8 treatments (4 N levels and 2 cultivars) with 3 repetitions per treatment. The total length of the community was 39.6 m from East to West and 45 m from North to South (excluding protective line). In total, 42 experimental plots were established, each with an area of 5 × 6.6 m^2.

The farming method was forefoot rotation, and the previous crop was maize. The tested potato (*Solanum tuberosum* L.) cultivar was ‘Zhongshu 185’.

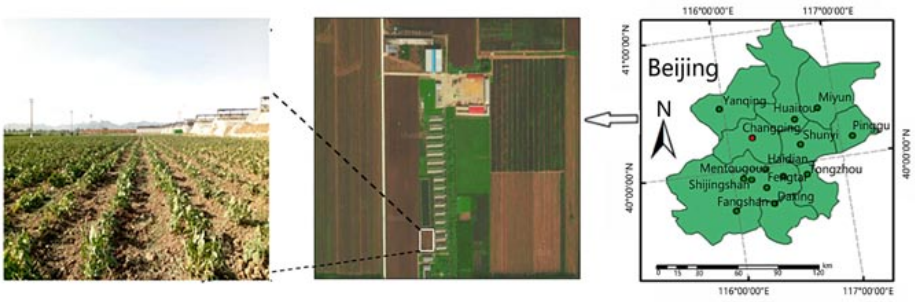


Figure 1. Location of potato test plot

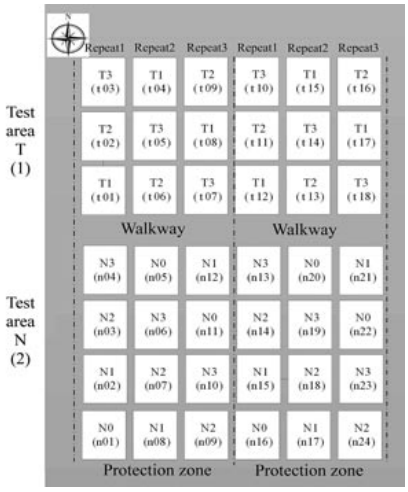


Figure 2. Design of the experiment

Based on the growth habit of this cultivar, as the soil type was selected *Histosol* (WRB, 2014). A spray-type automatic control irrigation system was installed in test plot. Using the automatic control system, the watering amount was set to 4 levels: 1 – very low, 2 – low, 3 – medium and 4 – high watering. After the emergence of the seedlings, only minimal cultivation was applied to keep the soil loose and level 1 of watering was added. At the seedling stage, soil cultivation was continued and the watering amount was set to level 3; the appropriate amounts of nitrogen (N), phosphorus (P) and potassium (K) were added. Before appearance of buds, the amount of watering was set to level 2. At the budding stage, the amount of irrigation water was increased and set to level 4. In order to increase the permeability of the soil, the soil was deeply ploughed and then covered with 3–5 cm thick soil around the roots. At the later growth period (15 days before harvest), foliar fertilizer containing 0.5% urea and 0.3% KH_2PO_4 (potassium dihydrogen phosphate) was sprayed on the plants.

Data acquisition. An unmanned aerial vehicle (UAV) remote sensing data acquisition experiment was carried out synchronously with the field data acquisition and sampling. Radiation correction (field calibration) was completed before the flight. All flight directions were North–South, and the height of the hyperspectral data (HD) acquisition was 40 m. Data were collected during the following five typical potato growth stages: 1) the budding stage (on 15 May, 2018, when the weather was sunny and cloud-free); 2) the tuber formation stage (on 29 May, 2018, when the weather was sunny and cloud-free); 3) the tuber growth stage (on 5 June, 2018, when the weather was sunny and cloud-free); 4) the starch accumulation stage (19 June, 2018, when the weather was sunny and cloud-free); and 5) the mature stage (on 29 June, 2018, when the weather was sunny cloud-free. Due to the poor measurement accuracy of ChlC of the dry leaves at the mature stage, data of the first four growth stages for modelling and model accuracy verification were selected.

Determination of potato chlorophyll content (ChlC). During each growth stage, the ChlC data were obtained synchronously with the UAV hyperspectral remote sensing data. Six leaves were randomly selected from the potato plant samples in each area, and 12 small samples from each leaf were collected using a 0.8 cm diameter puncher. The samples were weighed immediately using a balance of 0.001 g. Samples were then placed in test tubes containing 80 mL of 95% ethanol extract. Then the tubes were placed in the dark and shaken once a day until the leaves turned white. The optical density (OD) of the ethanol solution at visible wavelengths of 440, 655 and 649 nm was measured using a spectrophotometer Ci60 (X-rite, USA), and ChlC was calculated using equation (Dordas, 2011):

$$\text{Chl}\ (\mu g\ cm^{-2}) = (6.10 \times OD_{655} + 20.04OD_{649}) \times V \times 10 / S / 1000, \quad (1),$$

where Chl is the chlorophyll concentration ($\mu g\ cm^{-2}$), OD_{655} and OD_{649} – the absorbance values at 655 and 649 nm, respectively, V – the volume (mL) of the 95% ethanol extract, S – the leaf sample area (dm^2).

Acquisition and processing of UAV hyperspectral data (HD). An electric 8-rotor UAV (DJI Innovation Technology Co. Ltd., China) was used as the loading platform of the HD acquisition system. As shown in Figure 3, the system consists of a flight control system, inertial measurement system (inertial measurement unit, IMU), wireless remote-control system, ground control system, data processing system and sensor (HD multispectral camera and acquisition system).

Figure 3. Unmanned aerial vehicle (UAV) hyperspectral data acquisition system

The hyperspectral sensor adopted the imaging spectrometer Cubert UHD 185 (Cubert GmbH, Germany). The basic technical parameters are listed in Table 1.

UAV data collection was carried out, when the solar radiation intensity was stable, and the sky was clear and cloudless. Before acquiring the hyperspectral images of the UAV, a standard white board was placed near the research area for calibration. The flight altitude was 50 m, and the flight was conducted according to the planned route.

UAV HD processing primarily included radiation correction, image mosaicking and spectral extraction of the average canopy area. First, based on the UHD 185 centre wavelength and half-width of the wavelength, a radiation calibration system was designed in the

Table 1. The parameters of imaging spectrometer Cubert UHD 185

Parameter	Value	Parameter	Value
Spectral resolution nm	4	Spectral range nm	454–950
Pixel resolution nm	0.034	Scanning speed, lines	15–60
Viewing angle °	20	Standard lens focal length mm	25

MATLAB environment to perform radiation correction from the image digital number (DN) value to the surface reflectance. Second, the images of the test area were sieved and spliced using software *Agisoft PhotoScan* (Agisoft LLC, Russia). In the process of image mosaicking, professional software was used to align and transform each image into cue-format data, and then subbands of the image were extracted into a .jpg file. Finally, the subband images of each mosaic were merged to generate regional hyperspectral image data. The average spectrum of 125 band in each study area was extracted using IDL programming language. In the extraction process, the image was resampled using the pixel aggregation method and exported to the corresponding comma-separated values (CSV) format file for later data analysis.

Fractional-order differential is an extension of integer order differential, that is the definition of differential is generalised. When the differential order is a positive integer, the integer-order differentiation becomes a special case of fractional-order differentiation. Three common fractional-order differential forms: the Caputo, Riemann-Liouville and Grünwald-Letnikov definitions, have been developed; however, to process HD, the Grünwald-Letnikov differential definition was adopted (Ngo Van, Ho, 2020):

$$\frac{d^{\alpha} f(\lambda)}{d\lambda^{\alpha}} \approx f(\lambda) + (-a) f(\lambda - 1) + \frac{(-a)(-a+1)}{2} f(\lambda - 2) + \mathrm{L} + \frac{\Gamma(-a+1)}{n!\Gamma(-a+1)} f(\lambda - n) \quad (2),$$

where λ is the corresponding wavelength, Γ – the gamma function, n – the difference between the upper and lower bounds of the differential, a – any order. When α = 0, 1 or 2, it is the original function (original spectrum), 1-order or 2-order differential spectra. When α is a decimal, it is in the form of the Grünwald-Letnikov fractional-order differential. In this study, 11 types of canopy differential spectral data for each growth stage were obtained using the layer height spectral data of potato crowns at the 454–950 nm band, and the Grünwald-Letnikov differential method is defined in equation (2) for 0–2 differential processing with a step length of 0.2.

Pearson correlation analysis. The variation trend of two or more groups of data to determine the degree of closeness of their relationship was assessed. In this approach, the correlation coefficient is often used for discriminant analysis. It was obtained by dividing the covariance of two random variables using the standard deviation. The correlation coefficient was between −1 and 1. A large absolute value indicates a large correlation degree, whereas a value close to 0 indicates no correlation. The calculation method is shown in equation:

$$\rho_{X,Y} = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - E^2(X)}\sqrt{E(Y^2) - E^2(Y)}} \quad (3),$$

where $\rho_{X,Y}$ represents the correlation coefficient, cov(*X,Y*) and σ – covariance and standard deviation, respectively.

Model accuracy evaluation index. To evaluate the model accuracy, the coefficient of determination (R^2), root mean square error (RMSE) and normalized root mean square error (nRMSE) were used. Calculations of the evaluation index are shown in equations:

$$R^2 = \frac{(\sum_{i=1}^{n} y_i - \bar{y})^2}{(\sum_{i=1}^{n} x_i - \bar{y})^2} \quad (4),$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (x_i - y_i)^2}{n}} \quad (5),$$

$$\mathrm{nRMSE} = \sqrt{\frac{\sum_{i=1}^{n} (x_i - y_i)^2}{n}} \Big/ \bar{y} \quad (6),$$

where y_i is the estimated value, $\bar{y}$ – the mean value, *xi* – the measured value, n – the number of samples.

In general, the larger the R^2, the smaller the RMSE. This indicates a good fit of the model. The nRMSE defines the accuracy range in the model validation. A value for nRMSE < 10% indicates that the estimated and measured values are in very good agreement, 10% < nRMSE < 20% – good agreement, 20 ≤ nRMSE < 30% – intermediate agreement and nRMSE ≥ 30% – poor agreement.

Results

Correlation analysis of the canopy original spectrum and chlorophyll content (ChlC). In Figure 4, the spectral curve of the potato canopy presents typical greenery reflection characteristics. The spectral reflectance in the visible band between 0.05 and 0.55 has a typical plant Chl reflection peak at approximately 550 nm. As the wavelength changes, the spectral reflectance shock after 682 nm forms the "red edge" characteristic of typical green vegetation.

The original spectral band had the best correlation with potato ChlC, as determined by Pearson correlation analysis. The result of the correlation was obtained between the original spectrum and the ChlC of the potato canopy at different growth stages (Figure 5).

The data in Figure 5 show the following. (A) At the budding stage, the original potato canopy spectrum had an extremely significant negative correlation with ChlC at the level of 0.01 within the ranges of 502–662 and 662–714 nm, and an extremely significant positive correlation with ChlC within the range of 734–950 nm, reaching 0.01 extremely significant level. As the spectral bands related to ChlC were mainly visible, the bands with the largest 538, 710 and 714 nm correlations were considered within the band interval significantly related to ChlC; the correlation coefficients of these bands were −0.66, −0.77, and −0.76, respectively.

(B) At the potato tuber formation stage, the original canopy spectrum was significantly negatively correlated with ChlC at the 0.01 level within the ranges of 506–654 and 690–726 nm, reaching 0.01 extremely significant level, and no band had an extremely significant positive correlation with ChlC. Therefore, in the visible band ranges of 506–654 and 690–726 nm, there were selected

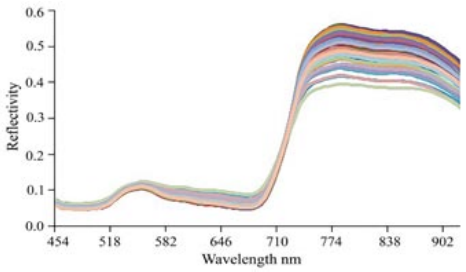


Figure 4. Spectral curve of the potato canopy

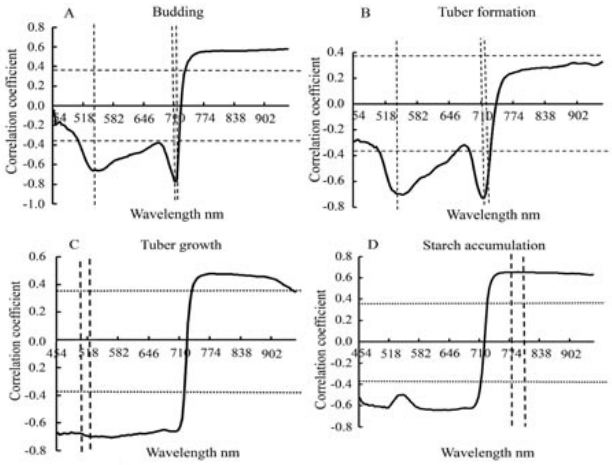


Figure 5. Correlation between the potato canopy original spectrum and chlorophyll content at different growth stages

the 550, 710 and 714 nm bands, which had the highest correlation with ChlC; the corresponding correlation coefficients were 0.70, 0.73 and 0.73, respectively.

(C) At the potato tuber growth stage, the original canopy spectrum was negatively correlated with ChlC at the 0.01 level within the range of 454–718 nm and reached the 0.01 level within the range of 738–934 nm, reaching 0.01 extremely significant level. Therefore, the 558 and 566 nm visible bands, which had the largest correlation with ChlC, were selected; the corresponding correlation coefficients were −0.71 and −0.72, respectively.

(D) At the starch accumulation stage, the original potato canopy spectrum had an extremely significant negative correlation with ChlC at the level of 0.01 within the range of 454–710 nm, and an extremely significant positive correlation with ChlC at the level within the range of 726–950 nm, reaching 0.01 extremely significant level. Therefore, in the range of 726–950 nm visible bands, the bands with the highest correlation were 786 nm and 790 nm; the corresponding correlation coefficients were 0.65 and 0.66, respectively.

Correlation analysis of the fractional differential spectrum and ChlC. The potato canopy HD in the 454–950 nm band in the differential method of equation (2) was used. The 0–2-order differential processing was carried out (the order interval was 0.2), resulting in 11 types of canopy differential spectral data in four typical growth periods. The correlations between the different differential spectra and ChlC are shown in Figure 6.

According to the correlations between the potato ChlC and fractional differential canopy spectrum at different growth stages, the relationship between the canopy spectrum of each order differential and the absolute value of the maximum of ChlC at different growth stages was analysed, and order curves showing the absolute value of the maximum of the correlation coefficients of ChlC and various differential orders at different growth stages were drawn (Figure 7).

Figures 6 and 7 show that as the differential order increased, the differential order that yielded the maximum absolute correlation coefficient value in different growth periods differed according to the analysis of the absolute correlation coefficient values for potato ChlC and the differential orders at different growth stages. At the budding stage, the absolute value of the correlation coefficient was highest for 1-order differentiation (662 and 686 nm). At the potato tuber formation and tuber growth stages, the absolute value of the correlation coefficient was highest for 0.6-order differentiation (538, 526, 506 and 558 nm). At the starch accumulation stage, the absolute value of the correlation coefficient was highest for 1.2-order differentiation (822 nm). Although the "red edge" and the "green" bands are closely related to the crop ChlC, only the "red edge" and the "green" bands information are used to estimate the ChlC. The effective validity of the spectral data cannot be fully exploited leading to the saturation phenomenon in the process of ChlC estimation. Using fractional differentiation refines the effective information of spectral data to improve the spectral data sensitivity and enhance the correlation between the information of the "red edge" and "green" bands and ChlC.

To avoid as much collinearity among different differential bands as possible, the absolute correlation coefficient values of ChlC were sorted from large to small, and the top 10 scores in each typical growth stage at a significance level of 0.01 were screened out. Then, a matrix diagram of the correlation coefficients of ChlC and the fractional differential spectrum in potatoes at different growth stages was plotted (Figure 8).

Figure 8 shows that at the budding stage, 0-order differentiation occurs in the 710 and 714 nm bands, 0.2-order differentiation – in the 706 nm band, 0.4-order differentiation – in the 702 and 698 nm bands, 0.6-order differentiation – in the 690 and 694 nm bands, 0.8-order differentiation – in the 514 nm band and 1-order differentiation – in the 662 and 686 nm bands. The absolute value of the correlation coefficient is between 0.74 and 0.85, reaching 0.01 significance level. At the tuber growth stage, 0-order differentiation occurs in the 538 and 574 nm bands, 0.2-order differentiation – in the 562, 566 and 570 nm bands, 0.6-order differentiation – in the 506 and 558 nm bands, 1.2-order differentiation – in the 454 and 658 nm bands and 2.0-order differentiation – in the 714 nm band. The absolute value of the correlation coefficient is between 0.61 and 0.71, reaching 0.01 significance level. At the starch accumulation stage, 0-order differentiation occurs in the 790 nm band, 0.4-order differentiation – in the 750 nm band, 0.6-order differentiation – in the 582 and 786 nm bands, the 0.8-order differentiation – in the 558-nm band, 1.0-order differentiation – in the 734 nm band, 1.2-order differentiation – in the 822 nm band, 1.4-order differentiation – in the 662 nm band, 1.6-order differentiation – in the 794 nm band and 1.8-order differentiation – in the 718 nm band. The absolute value of the correlation coefficient is above 0.65, reaching 0.01 significance level.

Optimal model selection of potato ChlC based on the fractional differentiation of the canopy spectrum. The ChlC was used as the dependent variable from 0- to 2-order fractional differentiation to determine the optimal band, which was the independent variable. The model was established using a fractional-order differential spectrum and ChlC, multiple linear regression (MLR), support vector machine (SVM) and random forest (RF) models. The accuracy of the models was verified, and the optimal estimation model was screened (Table 2).

The data of Table 2 show the following. (1) At the budding stage, compared with the MLR and RF models, for the SVM, the R^2 of the modelling accuracy increased by 5% and 6%, while the RMSE decreased by 0.18 and 0.06 μg cm^{-2}, respectively. The R^2 of verification accuracy increased by 9% and 6%, and the RMSE decreased by 0.41 and 0.25 μg cm^{-2}, respectively. These results show that the accuracy of potato ChlC estimation model was higher for SVM than for MLR and RF models. For the validation model, the R^2 of the MLR was higher than that of the SVM and RF models. The RMSE and nRMSE of the MLR were close to those of the SVM and RF models. This shows that the stability of the MLR was higher than that of the SVM and RF models.

(2) At the tuber formation stage, in the verification model, the R^2 of the RF was 3% and 2% higher than the R^2 of the SVM and MLR models, respectively, and the RMSE value of the RF was similar to that of the SVM and MLR models. The accuracy of the ChlC estimation

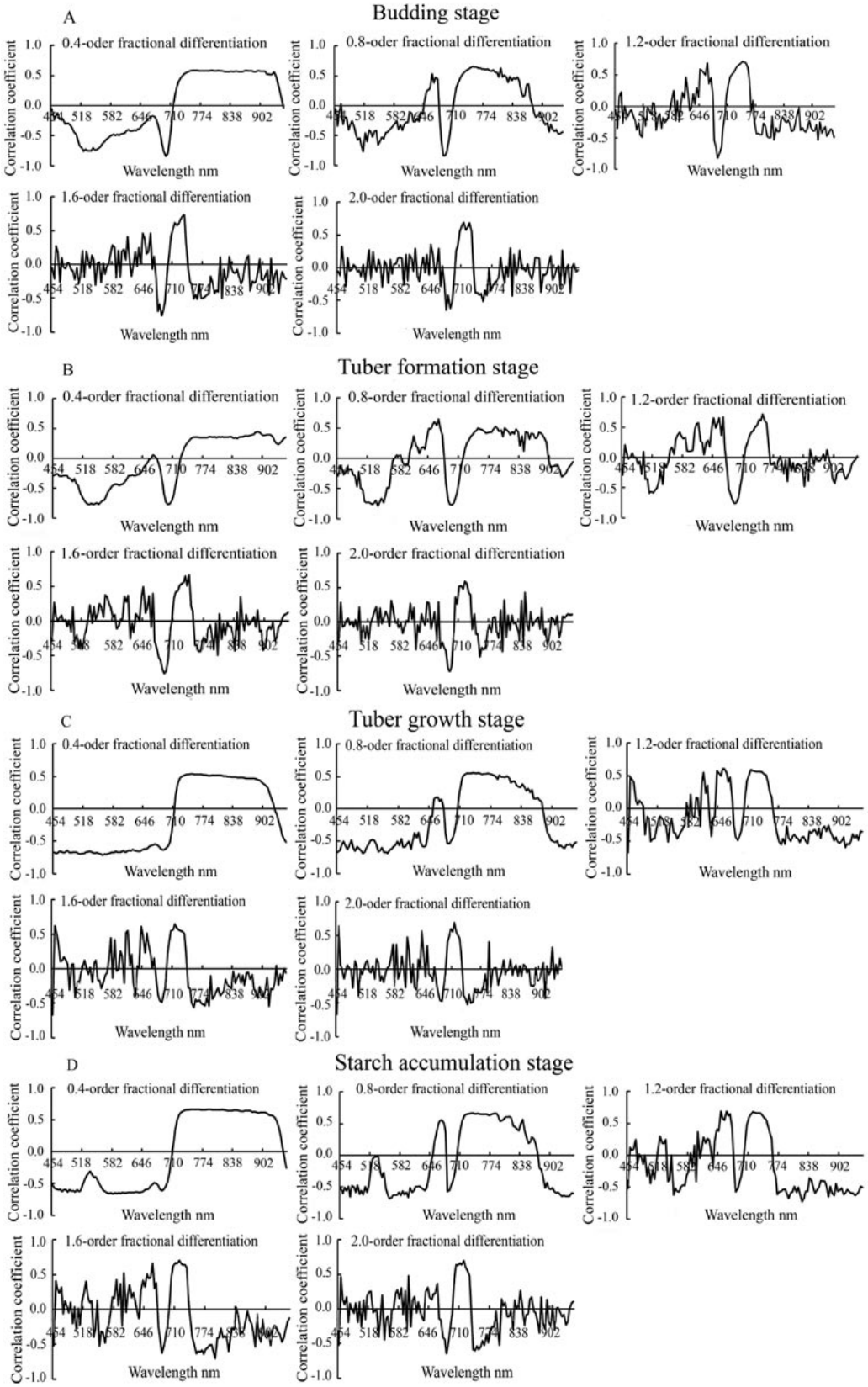


Figure 6. Correlation between the chlorophyll content and fractional differential spectrum of potato

model based on the SVM was higher than that based on the MLR and RF models. The stability of the RF was higher than that of the SVM and MLR models.

(3) At the tuber growth stage, in the verification model, the accuracy of the ChlC estimation model based on the MLR was higher than that based on the SVM and RF models. The stability of the MLR was higher than that of the SVM and RF models.

(4) At the starch accumulation stage, the accuracy of the ChlC estimation model based on the MLR was higher than that based on the RF and SVM models. The stability of the MLR was higher than that of the RF and SVM models.

In summary, the SVM was selected to construct the model to estimate potato ChlC at the budding and tuber formation stages, and the MLR was selected to

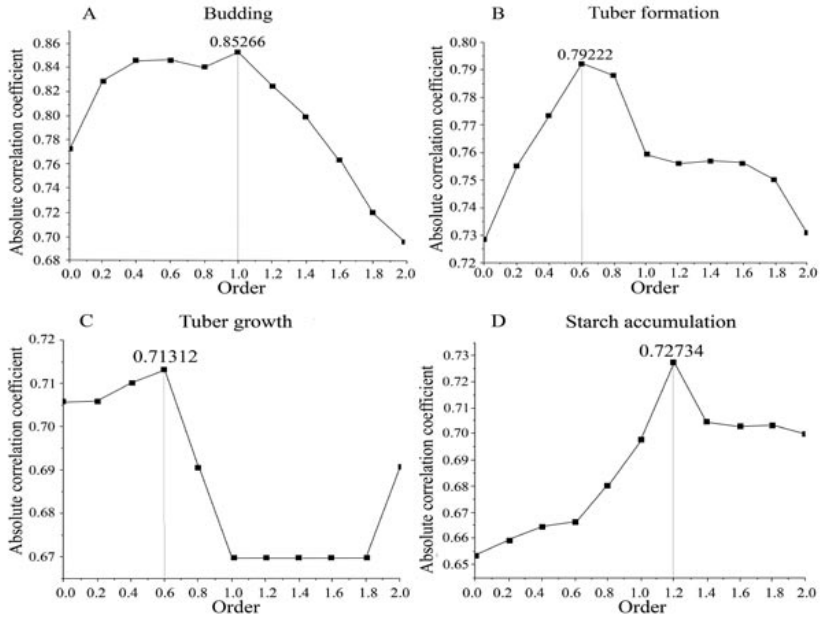


Figure 7. Absolute correlation coefficient and order curve of the chlorophyll content at different growth stages of potato

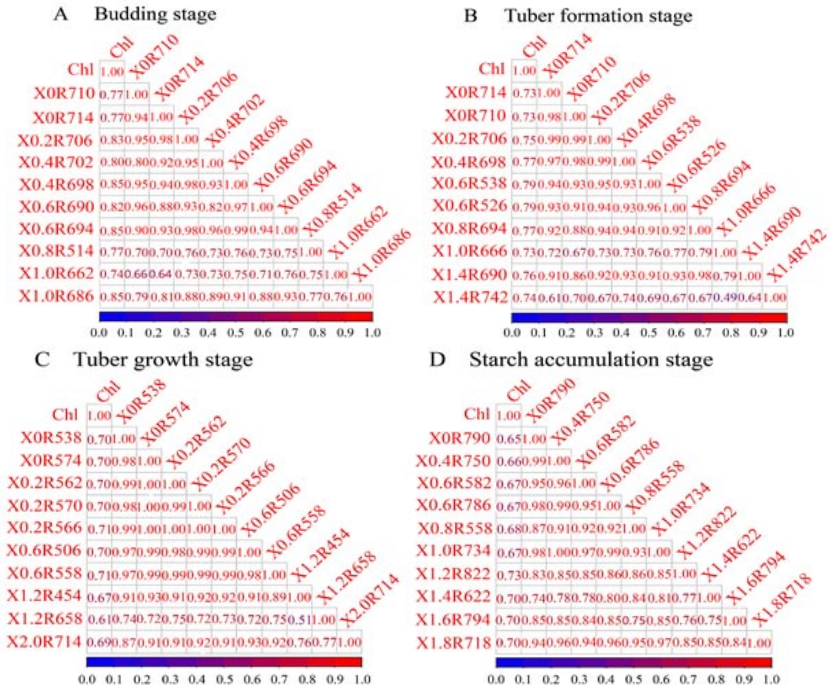
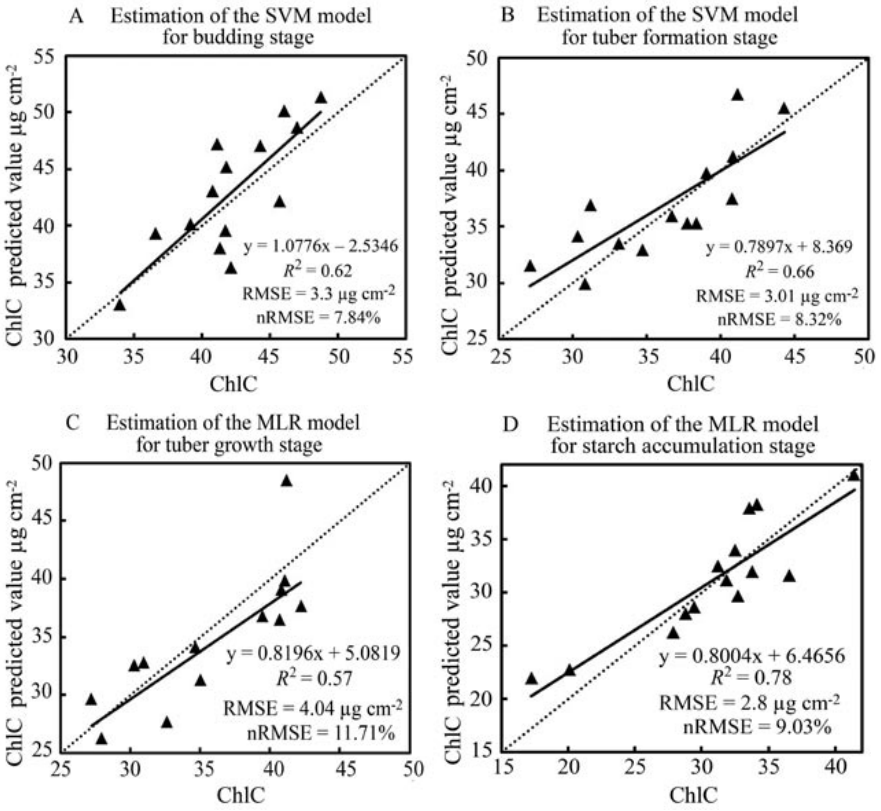


Figure 8. Correlation coefficient matrix of the chlorophyll content and fractional differential spectrum of potato

Table 2. Accuracy of estimation model of chlorophyll content at different growth stages of potato

Growth stage	Modelling method	Modelling accuracy			Verification accuracy		
		R^2	RMSE	nRMSE	R^2	RMSE	nRMSE
(1) Budding	MLR	0.78**	2.76	6.38%	0.71**	2.96	6.89%
	SVM	0.83**	2.58	5.92%	0.62**	3.37	7.84%
	RF	0.77**	2.97	6.84%	0.65**	3.21	7.48%
(2) Tuber formation	MLR	0.78**	2.79	7.40%	0.67**	2.77	7.51%
	SVM	0.80**	2.86	7.52%	0.66**	3.07	8.32%
	RF	0.63**	3.94	10.34%	0.69**	2.70	7.31%
(3) Tuber growth	MLR	0.64**	3.56	9.77%	0.57**	4.04	11.71%
	SVM	0.55**	3.84	10.39%	0.47**	4.86	13.73%
	RF	0.47**	4.25	11.37%	0.39	4.94	14.31%
(4) Starch accumulation	MLR	0.67**	3.49	11.56%	0.78**	2.81	9.03%
	SVM	0.63**	3.52	11.47%	0.60**	4.10	14.01%
	RF	0.50**	4.40	14.31%	0.46**	4.20	14.39%

MLR – multiple linear regression, SVM – support vector machine, RF – random forest; ** – significant at $p < 0.01$



SVM – support vector machine, MLR – multiple linear regression

Figure 9. Chlorophyll content (ChlC) predicted and measured values at different growth stages of potato

construct the model to estimate ChlC at the tuber growth and starch accumulation stages (Figure 9).

Discussion

The advantages of hyperspectral data (HD) include high spectral resolution and abundant spectral information of ground objects. The spectral characteristics of ground objects can be analysed from multiple angles, and the detailed information of crops can be expressed more comprehensively. Song et al. (2016) showed that HD having a high spectral resolution can be used to capture the changes in crop biochemical components: various pigments, N, Chl, lignin, cellulose and water. This results in a tiny position and depth difference in the spectral reflection spectrum curve, which enables an accurate estimation of crop biochemical parameters.

The results of the present study showed that at the budding, tuber formation, tuber growth and starch accumulation stages, the mean and maximum values of R^2 of the optimal model for ChlC estimation were 0.74 and 0.83, respectively, indicating a good estimation effect. The results are consistent.

He et al. (2018) treated the HD with integer-order differentiation, constructed the differential ratio index D_i / D_j, normalized index $(D_i – D_j / D_i + D_j)$ and constructed an estimation model for the ChlC of crops. The results showed that the R^2 of the model was 0.77 and 0.68, respectively. The estimation accuracy of ChlC was lower than that of the fractional differential estimation model in the present study. This is mainly because the original spectrum and integer-order differential spectrum ignore the gradual change information of the spectral curve. Fractional differential can refine the local information of the spectrum, which aids in the reduction of background noise, the extraction of detailed information, improvement in the sensitivity of the spectral data and enhancement of the correlation between the information of "red edge" and "green" bands and ChlC. The conclusions are consistent with those of Faghih and Mokhtary (2021).

Results of our experiment demonstrate that the best growth period for potato ChlC estimation is the budding stage. This is mainly due to the rapid growth of stems and leaves during this growth period, the rapid increase in ChlC as well as photosynthesis efficiency and the obvious difference in spectral response characteristics (the correlation with ChlC was the strongest). After plants enter the flowering stage, the ChlC gradually decreases, and the difference in spectral response characteristics gradually weakens, which is consistent with the research conclusions of Liu et al. (2020).

Conclusion

Fractional differentiation can effectively remove the background noise of the image and fully excavate the potential information in the hyperspectral spectrum. Therefore, the paper applied fractional-order differentiation technology to unmanned aerial vehicle (UAV) hyperspectral data (HD) processing. Using the spectral data processed by fractional differentiation, based on the models of multiple linear regression (MLR), support vector machine (SVM) and random forest (RF), the potato chlorophyll content (ChlC) at different growth stages was estimated.

The results showed that at the budding, tuber formation, tuber growth and starch accumulation stages, the R^2 of the optimal ChlC estimation model reached 0.83, 0.80, 0.64 and 0.67, respectively, indicating that the hyperspectral fractional differential technique applied to crop ChlC estimation is feasible and the estimation effect is good. When constructing the potato ChlC estimation model, the statistical model was used, which was simple, quick and easy to operate. However, this method relies

on remote sensing data and ground measured data, which simplifies the radiation transmission process and does not consider the radiation transmission mechanism.

The radiation transmission model considers the influence of crop physiological parameters on the spectral reflectance, but the model is too complex, having numerous parameters that affect one another. Therefore, in follow-up research, the combination of the radiation transmission mechanism and statistical model should be considered in order to further improve the robustness of the model. Moreover, the small size of the sample has a certain impact on the robustness of the model. In future studies, the sample size, number of crop cultivars, sample age range and sample collection area can all be increased, thereby improving the universality and robustness of the model.

Acknowledgments

This study was supported by the Natural Science Foundation of China (41871333), the Important Project of Science and Technology of the Henan Province (212102110238; 182102310844) and Key scientific research project of Henan college and university (20B420002).

We thank Haikuan Feng for the image data and field sampling collection.

We are grateful to the anonymous reviewers for their valuable comments and recommendations.

Received 07 06 2020
Accepted 01 02 2021

References

Afshari H., Jarad F., Abdeljawad T. 2020. On a new fixed point theorem with an application on a coupled system of fractional differential equations. Advance in Difference Equations, 2020 (1): 461. https://doi.org/10.1186/s13662-020-02926-0

Ali A., Shah K., Abdeljawad T., Mahariq I., Rashdan M. 2021. Mathematical analysis of nonlinear integral boundary value problem of proportional delay implicit fractional differential equations with impulsive conditions. Boundary Value Problems, 7 (2021). https://doi.org/10.1186/s13661-021-01484-y

An G., Xing M., He B., Liao C., Huang X., Shang J., Kang H. 2020. Using machine learning for estimating rice chlorophyll content from *in situ* hyperspectral data. Remote Sensing, 12 (18): 3104. https://doi.org/10.3390/rs12183104

Bahrami M., Mobasheri M. R. 2020. Plant species determination by coding leaf reflectance spectrum and its derivatives. European Journal of Remote Sensing, 53 (1): 258–273. https://doi.org/10.1080/22797254.2020.1816501

Basinger N. T., Jennings K. M., Hestir E. L., Monks D. W., Jordan D. L., Everman W. J. 2020. Phenology affects differentiation of crop and weed species using hyperspectral remote sensing. Weed Technology, 34 (6): 897–908. https://doi.org/10.1017/wet.2020.92

Cardone A., Conte D. 2020. Stability analysis of spline collocation methods for fractional differential equations. Mathematics and Computers in Simulation, 178: 501–514. https://doi.org/10.1016/j.matcom.2020.07.004

Cordon G., Lagorio M. G., Paruelo J. M. 2016. Chlorophyll fluorescence, photochemical reflective index and normalized difference vegetative index during plant senescence. Journal of Plant Physiology, 199: 100–110. https://doi.org/10.1016/j.jplph.2016.05.010

Croft H., Arabian J., Chen J. M., Shang J., Liu J. 2020. Mapping within-field leaf chlorophyll content in agricultural crops for nitrogen management using Landsat-8 imagery. Precision Agriculture, 21 (4): 856–880. https://doi.org/10.1007/s11119-019-09698-y

Dordas C. A. 2011. Nitrogen nutrition index and its relationship to N use efficiency in linseed. European Journal of Agronomy, 34 (2): 124–132. https://doi.org/10.1016/j.eja.2010.11.005

El-Hendawy S., Elsayed S., Al-Suhaibani N., Alotaibi M., Tahir M. U., Mubushar M., Attia A., Hassan W. M. 2021. Use of hyperspectral reflectance sensing for assessing growth and chlorophyll content of spring wheat grown under simulated saline field conditions. Plants (Basel), 10 (1): 101. https://doi.org/10.3390/plants100101101

Faghih A., Mokhtary P. 2021. A new fractional collocation method for a system of multi-order fractional differential equations with variable coefficients. Journal of Computational and Applied Mathematics, 383: 113139. https://doi.org/10.1016/j.cam.2020.113139

Gonzalez-Dugo V., Hernandez P., Solis I., Zarco-Tejada P. J. 2015. Using high-resolution hyperspectral and thermal airborne imagery to assess physiological condition in the context of wheat phenotyping. Remote Sensing, 7 (10): 13586–13605. https://doi.org/10.3390/rs71013586

Guo Y., Yin G., Sun H., Wang H., Chen S., Senthilnath J., Wang J., Fu Y. 2020. Scaling effects on chlorophyll content estimations with RGB camera mounted on a UAV platform using machine-learning methods. Sensors (Basel), 20 (18): 5130. https://doi.org/10.3390/s20185130

He T., Li J.-D., Liu G.-P., Wang G.-J., Li D. 2016. Estimation models of maze total nitrogen content based on hyperspectral remote sensing. Journal of Shenyang Agricultural University, 47 (3): 257–265 (in Chinese). https://doi.org/10.3969/j.issn.1000-1700.2016.03.001

He R., Li H., Qiao X., Jiang J. 2018. Using wavelet analysis of hyperspectral remote-sensing data to estimate canopy chlorophyll content of winter wheat under stripe rust stress. International Journal of Remote Sensing, 39 (12): 4059–4076. https://doi.org/10.1080/01431161.2018.1454620

Li W. G., Sun Z. Q., Lu S., Omasa K. 2019. Estimation of the leaf chlorophyll content using multiangular spectral reflectance factor. Plant, Cell and Environment, 42 (11): 3152–3165. https://doi.org/10.1111/pce.13605

Li C. C., Chen P., Ma C. Y., Feng H. K., Wei F. Y., Wang Y. L., Shi J. J., Cui Y. Q. 2020. Estimation of potato chlorophyll content using composite hyperspectral index parameters collected by an unmanned aerial vehicle. International Journal of Remote Sensing, 41 (1): 8176–8197. https://doi.org/10.1080/01431161.2020.1757779

Liang L., Qin Z. H., Zhao S. H., Di L. P., Zhang C., Deng M. X., Lin H., Zhang L. P., Wang Z. X. 2016. Estimating crop chlorophyll content with hyperspectral vegetation indices and the hybrid inversion method. International Journal of Remote Sensing, 37 (13): 2923–2949. https://doi.org/10.1080/01431161.2016.1186850

Lin X., Su Y. C., Shang J., Sha J., Li X., Sun Y., Ji J., Jin B. 2019. Geographically weighted regression effects on soil zinc content hyperspectral modeling by applying the fractional-order differential. Remote Sensing, 11 (6): 636. https://doi.org/10.3390/rs11060636

Liu J., Wang L., Yang F., Yang L. 2019. Spring corn leaf blight monitoring based on hyperspectral derivative index. Chinese Agricultural Science Bulletin, 35 (6): 143–150 (in Chinese). https://doi.org/10.11924/j.issn.1000-6850.casb18090024

Liu N., Qiao L., Xing Z. Z., Li M. Z., Sun H., Zhang J. Y., Zhang Y. 2020. Detection of chlorophyll content in growth potato based on spectral variable analysis. Spectroscopy Letters, 53 (6): 476–488. https://doi.org/10.1080/00387010.2020.1772827

Malin Hoeppner J., Skidmore A. K., Darvishzadeh R., Heurich M., Chang H.-C., Gara T. W. 2020. Mapping canopy chlorophyll content in a temperate forest using airborne hyperspectral data. Remote Sensing, 12 (21): 3573. https://doi.org/10.3390/rs12213573

Morley P. J., Jump A. S., West M. D., Donoghue D. N. M. 2020. Spectral response of chlorophyll content during leaf senescence in European beech trees. Environmental Research Communications, 2 (7): 071002. https://doi.org/10.1088/2515-7620/aba7a0

Ngo Van H. N., Ho V. 2020. A study of fractional differential equation with a positive constant coefficient via hilfer fractional derivative. Mathematical Problems in Engineering, 2020: 2749138. https://doi.org/10.1155/2020/2749138

Pereira da Conceição R. R., Ferreira Simeone M. L., Vieira Queiroz V. A., de Medeiros E. P., Borges de Araujo J., Macedo Coutinho W., da Silva D. D., de Araujo Miguel R., de Paula Lana U. G., de Resende Stoianoff M. A. 2020. Application of near-infrared hyperspectral (NIR) images combined with multivariate image analysis in the differentiation of two mycotoxicogenic *Fusarium* species associated with maize. Food Chemistry, 344: 128615. https://doi.org/10.1016/j.foodchem.2020.128615

Piegari E., Gossn J. I., Grings F., Barraza Bernadas V., Juarez A. B., Mateos-Naranjo E., Gonzales Trilla G. 2021. Estimation of leaf area index and leaf chlorophyll content in *Sporobolus densiflorus* using hyperspectral measurements and PROSAIL model simulations. International Journal of Remote Sensing, 42 (4): 1181–1200. https://doi.org/10.1080/01431161.2020.1826058

Qiao L., Gao D.., Zhang J.., Li M.., Sun H., Ma J. 2020. Dynamic influence elimination and chlorophyll content diagnosis of maize using UAV spectral imagery. Remote Sensing, 12 (16): 2650. https://doi.org/10.3390/rs12162650

Roosjen P. P. J., Brede B., Suomalainen J. M., Bartholomeus H. M., Kooistra L., Clevers J. G. P. W. 2017. Improved estimation of leaf area index and leaf chlorophyll content of a potato crop using multi-angle spectral data – potential of unmanned aerial vehicle imagery. International Journal of Applied Earth Observation and Geoinformation, 66: 14–26. https://doi.org/10.1016/j.jag.2017.10.012

Singh S. K., Houx J. H., Maw M. J. W., Fritschi F. B. 2017. Assessment of growth, leaf N concentration and chlorophyll content of sweet sorghum using canopy reflectance. Field Crops Research, 209: 47–57. https://doi.org/10.1016/j.fcr.2017.04.009

Song X., Xu D. Y., He L., Feng W., Wang Y., Wang Z., Coburn C. A., Guo T. 2016. Using multi-angle hyperspectral data to monitor canopy leaf nitrogen content of wheat. Precision Agriculture, 17 (6): 721–736. https://doi.org/10.1007/s11119-016-9445-x

Sonobe R., Yamashita H., Mihara H., Morita A., Ikka T. 2021. Hyperspectral reflectance sensing for quantifying leaf chlorophyll content in wasabi leaves using spectral pre-processing techniques and machine learning algorithms. International Journal of Remote Sensing, 42 (4): 1311–1329. https://doi.org/10.1080/01431161.2020.1826065

Tilahun S., An H. S., Solomon T., Baek M. W., Choi H. R., Lee H. C., Jeong C. S. 2020. Indices for the assessment of glycoalkaloids in potato tubers based on surface color and chlorophyll content. Horticulturae, 6 (4): 107–118. https://doi.org/10.3390/horticulturae6040107

Wang J., Tashpolat T., Zhang D. 2017. Spectral detection of chromium content in desert soil based on fractional differential. Transactions of the Chinese Society for Agricultural Machinery, 48 (5): 152–158 (in Chinese). https://doi.org/10.6041/j.issn.1000-1298.2017.05.018

WRB. 2014. World reference base for soil resources. World Soil Resources Reports No. 106. FAO, 189 p.

Xu K.-J., Tian Q.-J., Xu N.-X., Yue J.-B., Tang S.-F. 2019. Classifying forest dominant trees species based on high dimensional time-series NDVI data and differential transform methods. Spectroscopy and Spectral Analysis, 39 (12): 3794–3800 (in Chinese).

Yamashita H., Sonobe R., HironoY., Morita A., Ikka T. 2020. Dissection of hyperspectral reflectance to estimate nitrogen and chlorophyll contents in tea leaves based on machine learning algorithms. Scientific Reports, 10: 17360. https://doi.org/10.1038/s41598-020-73745-2

Yang J., Du L., Gong W., Shi S., Sun J. 2020. Estimating leaf nitrogen concentration based on the combination with fluorescence spectrum and first-derivative. Royal Society Open Science, 7 (2): 1–8. https://doi.org/10.1098/rsos.191941

Zhang D., Tiyip T., Ding J., Zhang F., Nurmemet I., Kelimu A., Wang J. 2016. Quantitative estimating salt content of saline soil using laboratory hyperspectral data treated by fractional derivative. Journal of Spectroscopy, 2016: 1081674. https://doi.org/10.1155/2016/1081674

Zhang J., Sun H., Gao D., Qiao L., Liu N., Li M., Zhang Y. 2020. Detection of canopy chlorophyll content of corn based on continuous wavelet transform analysis. Remote Sensing, 12 (17): 2741. https://doi.org/10.3390/rs12172741

Zhou X., Zhang J., Chen D., Huang Y., Kong W., Yuang L., Ye H., Huang W. 2020. Assessment of leaf chlorophyll content models for winter wheat using Landsat-8 multispectral remote sensing data. Remote Sensing, 12 (16): 2574. https://doi.org/10.3390/rs12162574

Zhu H. Q., Chen J. M., Yang C. H., Li Y. G., Gong J. 2019. Spectral pretreatment method for detection of trace Cu^{2+} and Co^{2+} in Zinc solution. Acta Optica Sinica, 39 (1): 468-476 (in Chinese). https://doi.org/10.3788/AOS201939.0130001

Zhu W. X., Sun Z. G., Yang T., Li J., Peng J., Zhu K., Li S., Gong H., Lyu Y., Li B., Liao X. 2020. Estimating leaf chlorophyll content of crops via optimal unmanned aerial vehicle hyperspectral data at multi-scales. Computer and Electronics in Agriculture, 178: 105786. https://doi.org/10.1016/j.compag.2020.105786

Išmaniosiomis sistemomis pagrįstas chlorofilo kiekio įvertinimas bulvėse skirtingais augimo tarpsniais naudojant bepiločiu orlaiviu gautus hiperspektrinius duomenis

C. Li[1], C. Ma[1], P. Chen[1,2], Y. Cui[1], J. Shi[1], Y. Wang[1]

[1]Henan politechnikos universiteto Topografijos ir žemės inžinerijos informacijos mokykla, Kinija
[2]GZH-HNJ BDS AGR Co. Ltd., Kinija

Santrauka

Siekiant pagerinti bulvių derlingumą ir kokybę, labai svarbu yra tiksliai įvertinti chlorofilo (Chl) kiekį. Frakcinė diferenciacija gali patikslinti vietinę spektro informaciją ir yra naudinga šalinant foninį triukšmą. Siekiant ištirti frakcinės diferenciacijos poveikį Chl kiekio nustatymui augaluose, buvo sukurtas naujas metodas. Tyrimo objektas – valgomoji bulvė (*Solanum tuberosum* L.). Bepiločiu orlaiviu surinkti hiperspektriniai duomenys apdoroti naudojant frakcinį išskaidymą, ir buvo tirtas bulvių Chl kiekio įvertinimo algoritmas.

Tyrimo rezultatai parodė, kad koreliacija padidėjo po pirmojo išskaidymo didėjant diferenciniam laipsniui; maksimalios absoliučios koreliacijos koeficiento reikšmės skirtingais etapais buvo gautos su 1 laipsnio diferenciacija pumpurų formavimosi tarpsniu, 0,6 laipsnių diferenciacija gumbų formavimo ir gumbų augimo tarpsniais ir 1,2 laipsnio diferenciacija krakmolo kaupimosi metu. Lyginant ir analizuojant bulvių Chl kiekio skirtingais augimo tarpsniais nustatymo modelius padaryta išvada, kad vertinant bulvių Chl kiekį tiksliausias buvo atramos vektoriaus klasifikatoriaus (SVM) modelis, kurio R^2 vertė pumpurų formavimosi tarpsniu buvo 0,83, gumbų formavimo tarpsniu – 0,80.

Reikšminiai žodžiai: chlorofilo kiekis, frakcinė diferenciacija, hiperspektriniai duomenys, *Solanum tuberosum*.