

## Original Articles

## Optimizing spectral index to estimate the relative chlorophyll content of the forest under the damage of *Erannis jacobsoni* Djak in Mongolia

Peiling Li<sup>a,b</sup>, Xiaojun Huang<sup>a,b,c,\*</sup>, Shan Yin<sup>a,b</sup>, Yuhai Bao<sup>a,b</sup>, Gang Bao<sup>a,b</sup>, Siqin Tong<sup>a,b</sup>, Ganbat Dashzeveg<sup>d</sup>, Tsagaantsoj Nanzad<sup>d</sup>, Altanchimeg Dorjsuren<sup>e</sup>, Davaadorj Enkhnasan<sup>e</sup>, Mungunkhuyag Ariunaa<sup>d</sup>

<sup>a</sup> College of Geographical Science, Inner Mongolia Normal University, Hohhot 010022, China

<sup>b</sup> Inner Mongolia Key Laboratory of Remote Sensing & Geography Information System, Hohhot 010022, China

<sup>c</sup> Inner Mongolia Key Laboratory of Disaster and Ecological Security on the Mongolia Plateau, Hohhot 010022, China

<sup>d</sup> Institute of Geography and Geology, Mongolian Academy of Sciences, Ulaanbaatar 15141, Mongolia

<sup>e</sup> Institute of Biology, Mongolian Academy of Sciences, Ulaanbaatar 13330, Mongolia



## ARTICLE INFO

## Keywords:

*Erannis jacobsoni* Djak  
Relative chlorophyll content  
Hyperspectral features  
Estimation potential

## ABSTRACT

Jas's Larch Inchworm (*Erannis jacobsoni* Djak) is a Lepidopteran insect pest that seriously threatens larch forest ecosystems in Mongolia. Damage caused by *E. jacobsoni* changes the chlorophyll content of forest trees, leading to significant changes in the color of the larch canopy. Chlorophyll content, an important parameter that reflects the physiological state of plants, is expressed as the relative chlorophyll content (RCC) evaluated using a chlorophyll meter. In this study, we estimate the relative chlorophyll content of forest trees damaged by *E. jacobsoni* by optimizing the spectral index. Four larch forest areas affected by outbreaks of this pest in Ikhtamir, Battsengel, and Tsenkher in the Mongolia Houhangai Province and Binder in the Khentii Province were selected as study areas. Based on the RCC and measured hyperspectral data of forest trees, hyperspectral features such as spectral index (SI) and continuous wavelet coefficients were analyzed. Partial least squares regression (PLSR), support vector machine regression (SVMR), and stepwise multiple linear regression (SMLR) were used to estimate the RCC of the total damage process and different degrees of damage. The optimized spectral index (OSI) showed the highest potential for estimating the total damage process, exhibiting good estimation accuracy and model stability. For example, in the SVMR model, the R2A of OSI-39 SVMR was 0.077, 0.074, 0.014, and 0.115 higher than those of the traditional spectral index (TSI), bior1.5, coif1, and sym3, respectively, while the RMSE was 0.017, 0.021, 0.014, and 0.048 lower than those of TSI, bior1.5, coif1, and sym3, respectively. In the estimation of different degrees of damage, the estimation performance of OSI was significantly improved compared with that of TSI and had the same potential as that of coif1. TSI, OSI, and coif1 showed the best estimation potential in a moderate degree of damage, of which OSI-SVMR had the best effect (R2A = 0.710 44 and RMSE = 0.137).

The optimization and combination method of SI used in this study will help to facilitate future research. Our findings provide insights into the estimation of RCC at the regional scale and for the effective monitoring of forest pest severity.

## 1. Introduction

*Erannis jacobsoni* Djak (Lepidoptera: Geometridae) completes one generation each year. Larvae first appear in early May, begin over-feeding on needles in early June, molt four times, enter the soil, and pupate at the end of July. Adults usually appear in September and infestations continue until mid-October. The females lay eggs in cracks in

the bark and overwinter with their eggs, and are wingless and hence cannot fly; therefore continue to infest the same tree for many years in a row. The damage due to this pest is characterized by slow growth along a 6–7-year cycle which peaks causing a large outbreak, following which the trees continue to lose leaves for 3–4 years, resulting in their death. *E. jacobsoni* Djak. is adapted to dry and continental climates. In recent years, severe infestations of this pest have been reported in coniferous

\* Corresponding author at: College of Geographical Science, Inner Mongolia Normal University, Hohhot 010022, China.

E-mail address: [hxj3s@qq.com](mailto:hxj3s@qq.com) (X. Huang).

<https://doi.org/10.1016/j.ecolind.2023.110714>

Received 13 April 2023; Received in revised form 22 June 2023; Accepted 20 July 2023

Available online 14 August 2023

1470-160X/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

forest areas in northern and northeastern Mongolia (Huang et al., 2018; Xi et al., 2020; Bai et al., 2021) posing a major threat to the security of local forest ecosystems. During an infestation, the appearance (canopy color and structure), as well as internal biochemical components (moisture, pigment, and nitrogen) of the affected trees, change considerably (Lin et al., 2016; Foster et al., 2017; Rahimzadeh-Bajgirani et al., 2018). Chlorophyll, the main component of plant pigments, is directly associated with plant photosynthetic potential, nutrient element content, primary productivity, and change in leaf color (Rumpf et al., 2010; Wang et al., 2015; Sakowska et al., 2018). Therefore, chlorophyll is an important indicator that reflects the physiological status of plants and helps to evaluate forest health. Estimation of chlorophyll content in different states of forest trees is vital for risk assessment of forest diseases and pest outbreaks, early identification of their occurrence, and estimation of the severity of the situation (Wu et al., 2012; Hamzeh et al., 2013; Oeser et al., 2017). The change in chlorophyll content is highly correlated with the spectral features of the leaves, which translates into the optical features that can be utilized as markers for the accurate estimation of chlorophyll concentration (Peñuelas et al., 1993; Ma et al., 2022). The SI constructed based on the spectral bands sensitive to the change of chlorophyll content provides insights into changes in chlorophyll and better estimates the chlorophyll content of plants (Lausch et al., 2013; Zhang et al., 2018). For example, modified chlorophyll absorption in reflectance (MCARI), transformed chlorophyll absorption in reflectance index (TCARI), medium resolution imaging spectrometer–terrestrial chlorophyll index (MTCI), optimizing soil adjusted vegetation index (OSAVI), normalized area over the reflectance curve (NAOC), and other hyperspectral indices can effectively reduce the impact of soil background and leaf area index (Wu et al., 2008). The hyperspectral indices such as red/green pigment indices (RGI), first derivative reflectance of modified simple ratio (BmSR), Gitelson and Merzlyak 1 (GM1), and triangular vegetation index (TVI) constructed using the green light zone and red edge wavelength can effectively alleviate the impact of ground saturation under high chlorophyll concentration (Gitelson and Merzlyak, 1997; Broge and Leblanc, 2001; Liao et al., 2013). Modified simple ratio index (MSRI), photochemical reflectance index (PRI), pigment-specific simple ratio (PSSR), and pigment-specific normalized difference (PSND) can also estimate chlorophyll content under vegetation stress (Blackburn, 1998; Frampton et al., 2013; Li et al., 2015). Spectral indices are used to investigate changes in biochemical components of forest trees under disease and insect stress due to their advantages in universality and ease of operation (Li et al., 2016; Hornero et al., 2020; Iordache et al., 2020). Since forest diseases and insects often have characteristics such as large outbreak areas and fast transmission speed, real-time dynamic monitoring of them is necessary at a large regional scale. Multi-temporal and large-scale unmanned aerial vehicle (UAV) and satellite remote sensing data can well meet this demand (Zhang et al., 2016; Imanyfar et al., 2019; Williams et al., 2021; Shirazinejad et al., 2022). However, the construction of the SI is mostly based on one plant, which cannot be accurately applied to research on different plants or different growth periods of plants (Wang et al., 2016). In recent years, the rapid development of acquisition and processing technologies for various types of hyperspectral data (satellite remote sensing data such as EO-1 Hyperion, Gaofen-5, Sentinel 2, and unmanned aerial vehicle airborne UAV hyperspectral data) has facilitated the study of chlorophyll content and other biochemical components of various forest trees (Fuente et al., 2018; Pulitua et al., 2018; Palaś and Zawadzki, 2020; Astola et al., 2021).

Based on the process of change caused by disaster stress associated with *E. jacobsoni* Djak, this study used the TSI, optimized spectral index (OSI), and smooth spectral continuous wavelet coefficients (CWCs) to establish regression models to estimate the RCC. The aims of the current study are to (1) reveal the sensitivity of TSI and CWCs to the RCC of affected trees; (2) analyze OSI band change; (3) compare the estimation potential of TSI, OSI, and CWCs during the total damage process; (4) compare the estimation potential of TSI, OSI, and CWCs at mild,

moderate, and severe damage levels.

## 2. Materials and methods

### 2.1. Data acquisition and preprocessing

#### 2.1.1. Study area

Four typical forest areas affected by *E. jacobsoni* Djak outbreaks in Ikhtamir, Battsengel, and Tsenkher in the Houhangai Province and Binder in the Khentii Province, Mongolia, were selected as the test areas (Fig. 1). The test areas have continental climates and comprise coniferous forest areas with a single tree species. There were no outbreaks of other pests in the forest area during the sampling year, and natural disasters such as drought did not occur.

A total of 110 sample trees were selected in the test areas (Table 1). The health statuses of sample trees were evaluated according to survey data on leaf loss rate. The sample trees were classified as healthy, mildly damaged, moderately damaged, and severely damaged based on leaf losses of 0–5%, 6–30%, 31–70%, and 71–100%, respectively. The number of trees evaluated as healthy and mildly, moderately, and severely damaged were 12, 33, 30, and 35, respectively.

#### 2.1.2. Hyperspectral data

An ASD FieldSpec4 portable ground object spectroradiometer (Analytical Spectral Devices, Boulder, Colorado, USA) was used to collect spectral data. Each sample tree was vertically classified into three levels, namely upper, middle, and lower. At each level, a typical branch representing the physiological state of the tree was selected for spectral measurement. The average value was taken as the spectral reflectance value of the tree. Spectral measurements were performed under clear and cloudless weather conditions from 10:30 to 14:30 Beijing time. During the measurement, the probe used a field of view of 25° and a downward vertical direction from a height of approximately 0.2 m. A reference whiteboard was used for spectral correction before and after measurements. The original spectral reflectance data were smoothed using the Savitzky–Golay method and recorded as smooth spectral reflectance (SSR). The SSR values under different levels of tree damages are shown in Fig. 2a.

Using the 36 mother wavelet bases of the four wavelet systems biorthogonal (bior1.3, bior1.5, bior2.2, bior2.4, bior2.6, bior2.8, bior3.1, bior3.3, bior3.5, bior3.7, bior3.9, bior4.4, bior5.5, and bior6.8), Coiflets (coif1, coif2, coif3, coif4, and coif5), Daubechies (db1, db2, db3, db4, db5, db6, db7, db8, db9, and db10), and symlets (sym2, sym3, sym4, sym5, sym6, sym7, and sym8) to perform continuous wavelet transform on the SSR on the scale of  $2^1$  to  $2^{10}$ , a series of continuous wavelet coefficients were obtained. Forty SIs that performed well in chlorophyll content estimation were selected (Table S1).

#### 2.1.3. Chlorophyll content data

A SPAD-502 portable chlorophyll meter (NANBEI, Henan, China) was used to measure the chlorophyll content of sample trees. SPAD measurement was performed on typical branched used for spectrum measurement, which was arranged during the early morning or late afternoon without clouds to reduce the influence of solar radiation on the SPAD reading. To improve the accuracy of measurement, three twigs of different health levels were selected from each branch, and readings were taken from the tip, middle, and sheath of the needle to calculate the average SPAD reading of each twig; the average value of three twigs was used as the branch SPAD value. Finally, the average SPAD value of three typical branches was taken as the RCC value of the sample tree. A Q–Q diagram was drawn to determine whether RCC data met the modeling requirements (Fig. 2b). The expected value was roughly distributed on a straight line and within the 95% confidence zone. The maximum, minimum, average, and standard deviation values were 50.3, 2.74, 20.76, and 12.58, respectively. This indicated that RCC data met the modeling requirements.

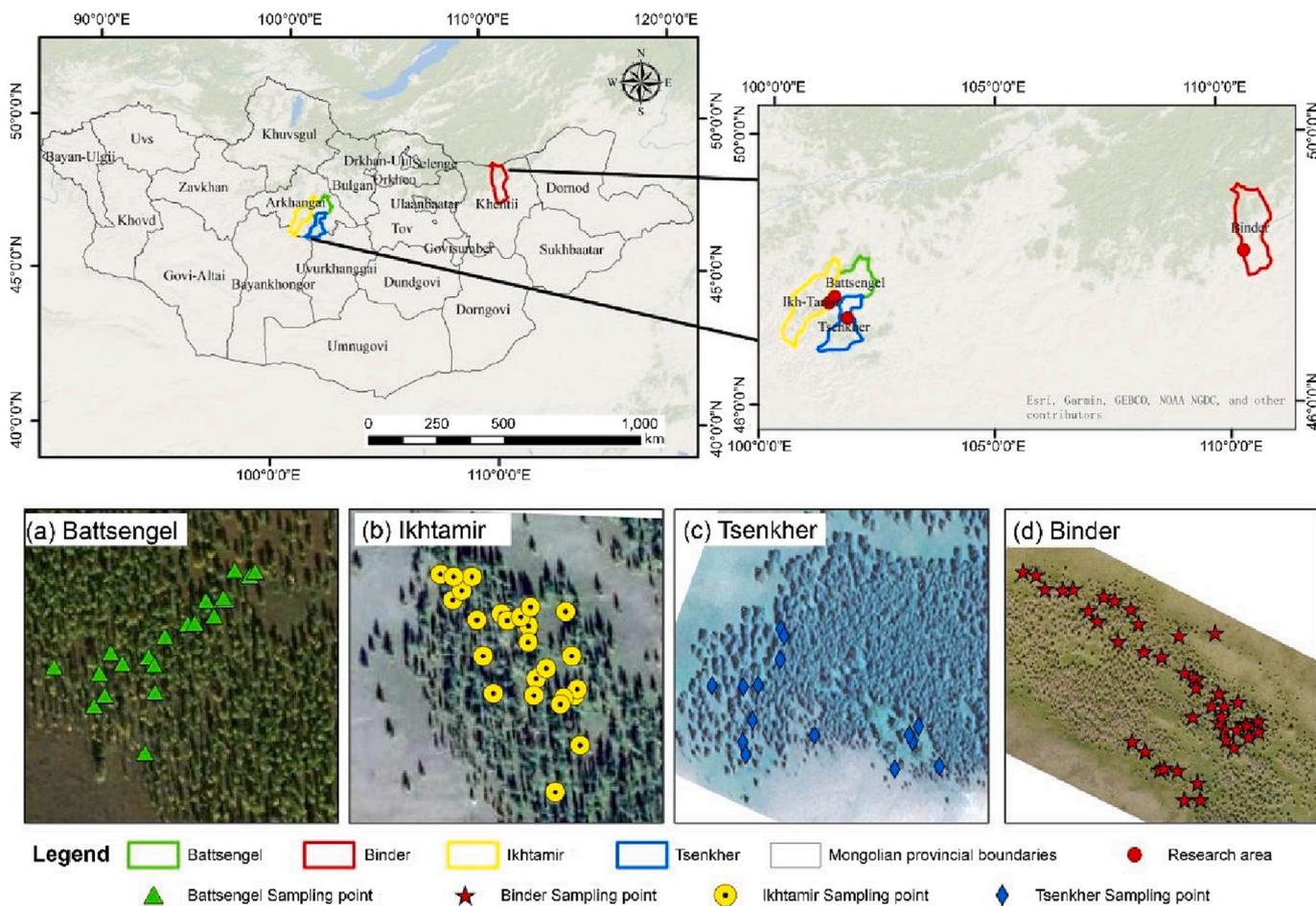


Fig. 1. Study area. (a) Battsengel, (b) Ikhtamir, (c) Tsenkher, and (d) Binder.

Table 1  
Overview of the study area.

Larch forest area	Sampling date	Number of sample trees	Average altitude (m)	Area (m <sup>2</sup> )	Age of sampled trees		
					Young	Middle	Old
Ikhtamir	June 18, 2016	27	1,766.75	19,923	7	11	9
Battsengel	June 18, 2016	21	1,808.95	14,851	8	10	3
Tsenkher	June 19, 2016	16	1,728.89	10,001	5	10	1
Binder	June 12, 2019	44	1,160.75	308,709	17	23	4

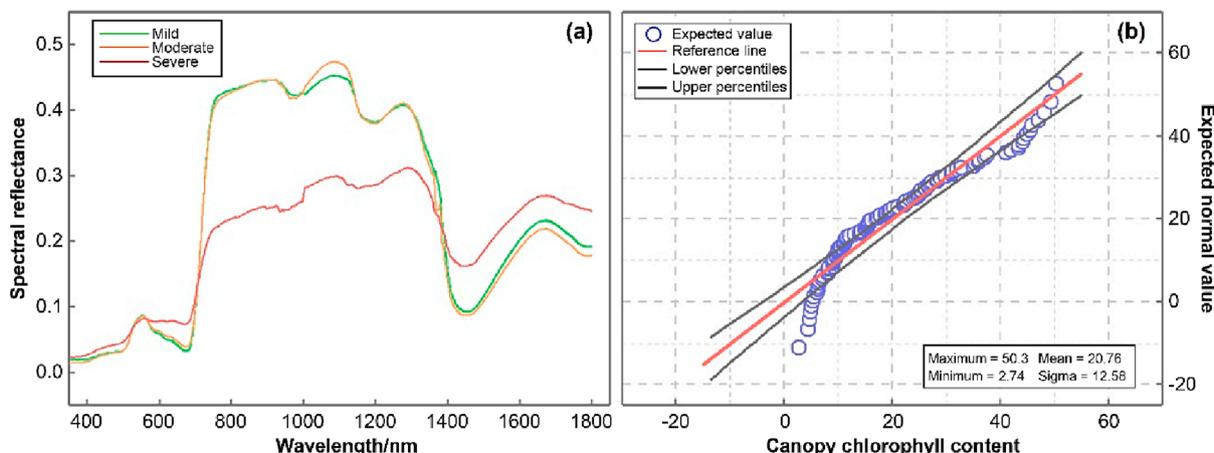


Fig. 2. (a) Specific simple ratio (SSR) under pest infestation severities, (b) Quantile–quantile (Q–Q) plot of relative chlorophyll content (RCC) data.

2.2. Methods

Hyperspectral feature extraction and estimation model establishment is shown in Fig. 3. First, the Pearson correlation analysis method was used for sensitivity analysis. The sensitivity of hyperspectral characteristics to changes in RCC was analyzed using the determination coefficient ( $R^2$ ) of the correlation coefficient,  $R$  (Liao et al., 2013), and the spectral index significantly correlated with RCC under the condition of  $P = 0.01$  was selected. Thereafter, the find peaks function–successive projection algorithm (Findpeaks-SPA) pattern (Huang et al., 2019) was used to construct TSI and CWCs that were ultimately used for modeling. The accuracy and stability of the model was verified using the correction determination coefficient ( $R_A^2$ ) and root mean square error (RMSE) of the model. To a large extent, model accuracy and stability may characterize model performance.  $R_A^2$  and the RMSE may be calculated using the following equations:

$$R_A^2 = 1 - \left[ 1 - \left( 1 - \frac{\sum_i (\hat{y}_i - \bar{y}_i)^2}{\sum_i (y_i - \bar{y}_i)^2} \right) \right] \frac{(n-1)}{(n-k)} \tag{8}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \tag{9}$$

where  $\hat{y}_i$  is the predicted value of canopy chlorophyll content,  $y_i$  is the measured value of canopy chlorophyll content,  $\bar{y}_i$  is the average value of canopy chlorophyll content,  $n$  is the number of samples, and  $k$  is the number of variables. The closer  $R_A^2$  is to 1, the better the model accuracy, and the closer RMSE is to 0, the better the stability.

The index used to build TSI was optimized to obtain OSI. The spectral index set (SI<sub>j</sub>) was calculated by considering the arrangement and combination of all wavelengths in this band range. The correlation between SI<sub>j</sub> and RCC was analyzed. The SI with the highest  $R^2$  was selected to adjust the original SI with its corresponding wavelength. The adjusted process was as follows:

$$\begin{cases} X_1 = [A_{m1} A_{m2} \dots A_{mm}] \\ X_2 = [B_{n1} B_{n2} \dots B_{nn}] \\ \vdots \\ X_i = [C_{k1} C_{k2} \dots C_{kk}] \end{cases} \tag{1}$$

$$[Y_i, Y_{i-1}, \dots, Y_1] = ndgrid(X_i, X_{i-1}, \dots, X_1) \tag{2}$$

$$Z_{ji} = [Y_1(:), Y_2(:), \dots, Y_i(:)] \tag{3}$$

$$SI_j = [D_1, D_2, \dots, D_j] \tag{4}$$

where  $X_1, X_2,$  and  $X_i$  are the row matrices comprising wavelengths of the 1st, 2nd, and  $i^{th}$  bands of values,  $A_{mm}, B_{nn},$  and  $C_{kk}$  are the maximum wavelength values,  $ndgrid()$  is a multi-dimensional rectangular grid function, and  $Y_1, Y_2, \dots, Y_i$  is an  $i$ -dimensional rectangular grid array generated through the  $ndgrid()$  function by copying the matrix of rows such as  $X_1, X_2, \dots, X_i$  (each grid value represents the wavelength value).  $Z_{ji}$  is a  $j$  row and  $i$  column matrix stacked and constructed by all wavelengths in each band of the SI.  $SI_j$  is a column matrix formed by  $j$  spectral indices ( $D_1, D_2, \dots, D_j$ ) constructed with different wavelength combinations; it comprises all possible wavelength combinations for one SI.

3. Results and analysis

3.1. Sensitivity analysis of hyperspectral features

3.1.1. SI

The sensitivity of the spectral index to RCC is listed in Table 2. The correlation between mSR705 and PSSRb and RCC was highest at  $> 0.6$ . In addition, MSRI, SR, MCARI1, MTVI1, PSSRa, and PRI were  $> 0.58$ , which also showed a high correlation. BRI1 showed the worst correlation at 0.056.

3.1.2. CWCs

The sensitivities of CWCs to RCC are shown in Fig. 4. CWCs have the highest sensitivity in the green peak and red edge of the visible light band, and sensitive bands are also distributed in the near-infrared band that is sensitive to changes in cell structure and the shortwave infrared band that is sensitive to changes in water content. For example, in  $coif1$  (scale 5), the sensitivity was high, showing ranges of 453–462 nm, 487–512 nm, 22–567 nm, 574–597 nm, 620–627 nm, 684–712 nm, 725–773 nm, 1239–1249 nm, 1259–1276 nm, 1574–1592 nm, and 1726–1738 nm ( $R^2 > 0.4$ ). The highest sensitivity was in the ranges 593–596 and 601–609 nm ( $R^2 > 0.6$ ).

3.2. OSI band change

Using the Fp-SPA algorithm, TCARI and mSR705 were finally selected to build the TSI. The optimized spectral index wavelength showed signs of moving toward the short wavelength direction (Table 3). The wavelength changed from 700 nm to 690 nm, 670 nm to 660 nm, and 550 nm to 540 nm, while its sensitivity increased by 0.104; mSR705: wavelength changed from 750 nm to 740 nm, 445 nm to 435 nm, and 705 nm to 695 nm, and its sensitivity increased by 0.041.

Table 2  
Sensitivity of SI to RCC.

SI	R <sup>2</sup>	SI	R <sup>2</sup>	SI	R <sup>2</sup>
RVI	0.505	RDVI	0.566	GRVI	0.469
NDVI	0.490	REIP	0.134	GNDVI	0.416
MTCI	0.417	PSSRa	0.584	ARVI	0.527
MCARI1	0.585	PSSRb	0.609	IPVI	0.491
TCARI	0.519	PSNDa	0.474	ARI	0.282
PRI	0.583	PSNDb	0.528	SR	0.592
SIPI	0.249	LIC1	0.474	mSR705	0.602
TVI	0.467	LIC2	0.167	RGI	0.486
VOG1	0.519	CTR2	0.466	ARI2	0.355
VOG2	0.475	BRI1	0.056	GM1	0.512
VOG3	0.485	BRI2	0.204	GM2	0.572
NPCI	0.418	ARI1	0.496	MTVI1	0.585
MSRI	0.592	mND705	0.563	DD	0.541
BmSR	0.479	-	-	-	-

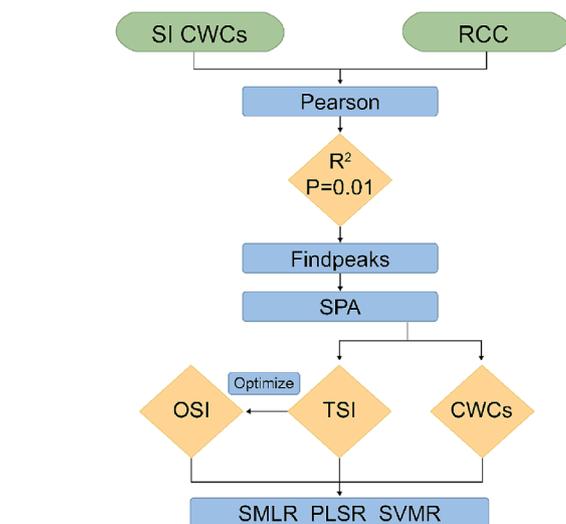


Fig. 3. Flow chart of hyperspectral feature extraction and estimation model establishment.

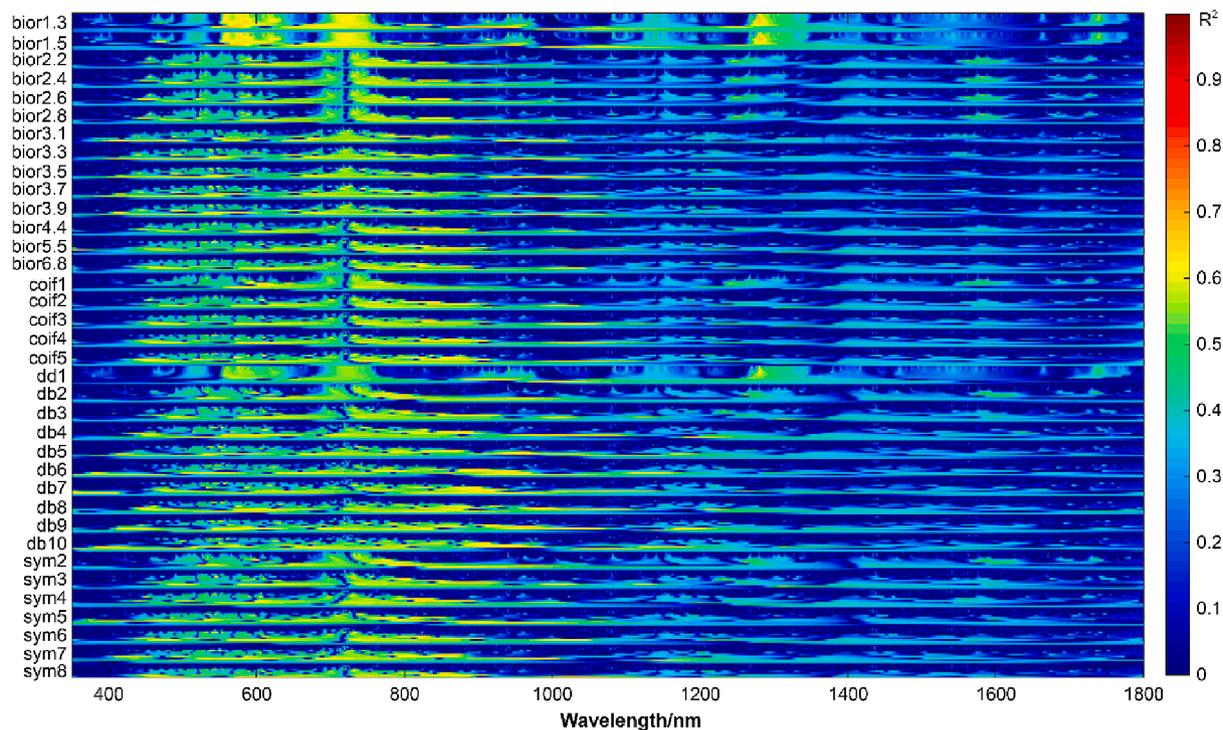


Fig. 4. Sensitivities of the four continuous wavelet coefficients (CWCs) to relative chlorophyll content (RCC). Data with 36 mother wavelet bases of 4 wavelet coefficients on the scale of  $2^1-2^{10}$ . For example, coif1 includes  $2^1$  (scale<sup>1</sup>),  $2^2$  (scale<sup>2</sup>) ..... $2^{10}$  (scale<sup>10</sup>).

Table 3  
Wavelengths of TSI and OSI.

Before optimized				After optimized			
SI		Wavelength (nm)	R <sup>2</sup>	SI		Wavelength (nm)	R <sup>2</sup>
TSI	TCARI	700, 670, 550	0.519	OSI	TCARI	690, 660, 540	0.624
—	mSR705	750, 445, 705	0.602	—	mSR705	740, 435, 695	0.643

### 3.3. Estimation of RCC

#### 3.3.1. Estimation of RCC in the total damage process

The estimation model of RCC for the total process of damage based on TSI, OSI, bior1.5, coif1, db4, and sym3 as input variables are listed in Table 4, which shows only the results of the modeling of the four optimal mother wavelet bases; all results are also shown in Table S2. OSI showed the best estimation accuracy in the three models. In the SVMR model, OSI was higher by 0.077, 0.074, 0.014, 0.094, and 0.115 than those of TSI, bior1.5, coif1, db4, and sym3, respectively. Modeling of OSI was also relatively stable and second only to db4, but the estimation accuracy of db4 was very poor. In CWCs, the estimation performance of coif1 was the best. The highest value (0.764) was shown by the Coif1-SVMR model. In summation, the model built based on OSI showed the best estimation effect as well as good accuracy and stability simultaneously,

Table 4  
Estimation model of RCC for the total damage process.

Hyperspectral features	PLSR		SVMR		SMLR	
	R <sub>A</sub> <sup>2</sup>	RMSE	R <sub>A</sub> <sup>2</sup>	RMSE	R <sub>A</sub> <sup>2</sup>	RMSE
TSI	0.699	0.144	0.710	0.140	0.684	0.149
OSI	0.776	0.124	0.781	0.123	0.776	0.124
bior1.5	0.702	0.141	0.686	0.144	0.702	0.141
coif1	0.763	0.140	0.764	0.137	0.763	0.140
db4	0.682	0.121	0.670	0.122	0.682	0.121
sym3	0.662	0.173	0.663	0.171	0.588	0.193

followed by CWCs and TSI.

The three-model validation sets of OSI, TSI, and coif1 were used to perform a 1:1 straight-line fitting analysis to test the performances of the models (Fig. 5). The results show that the data points of all test sets were closely distributed near the fitting line. In PLSR, SVMR, and SMLR, OSI had the highest degree of fit, followed by TSI. For OSI, OSI-SVMR with the best estimation accuracy had the highest fitting effect, and the data points of the test set were more evenly distributed.

#### 3.3.2. Estimation of RCC at different damage levels

To reveal the potential of the SI for estimating RCC at different levels of damage, TSI, OSI<sub>2</sub>, and coif1 were used to establish estimation models under mild, moderate, and severe damage levels. Results are shown in Table 5. At mild damage level, RA(2) values of OSI were higher by 0.139, 0.158, and 0.068 compared with those of TSI. At moderate damage level, OSI was higher by 0.135, 0.120, and 0.138. The severe damage level, OSI were higher by 0.110, 0.145, and 0.070. The estimated performance of OSI and coif1 differed slightly at each level of damage. In the PLSR and SMLR models for mild damage and the PLSR model for moderate and severe damage, the RA(2) of OSI was higher than coif1. In other models, it was marginally lower than coif1. TSI, OSI, and coif1 provided the best estimation of RCC at the moderate damage level. In summary, the estimability of OSI was better than that of TSI and comparable to that of coif1. TSI, OSI, and coif1 showed the highest accuracy in estimating RCC at the moderate level. Model validation sets of TSI, OSI, and coif1 were used to perform a 1:1 straight-line fitting analysis (Fig. 6) to test the performance of the model. The fit of OSI was

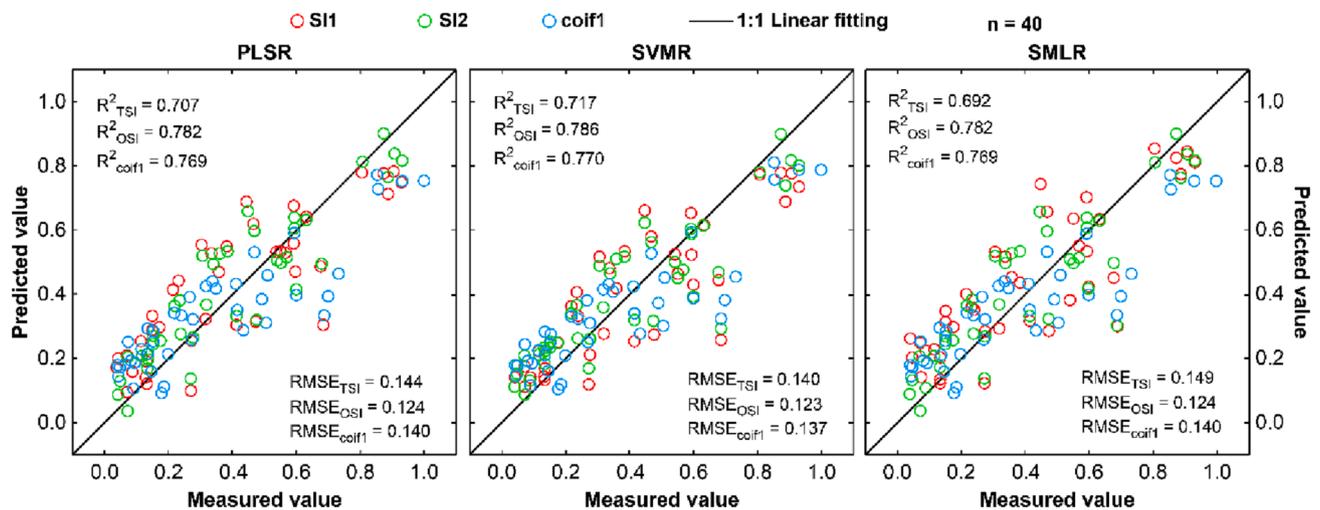


Fig. 5. 1:1 straight-line fitting based on the estimation model of the validation set for the total damage process.

Table 5

Estimation of RCC for different damage levels.

Damage level	Hyperspectral features	PLSR		SVMR		SMLR	
		$R_A^2$	RMSE	$R_A^2$	RMSE	$R_A^2$	RMSE
Mild	TSI	0.414	0.282	0.454	0.212	0.357	0.259
	OSI	0.553	0.148	0.612	0.154	0.425	0.175
	coif1	0.549	0.147	0.631	0.165	0.416	0.127
Moderate	TSI	0.552	0.128	0.591	0.195	0.411	0.189
	OSI	0.688	0.133	0.710	0.130	0.549	0.147
	coif1	0.677	0.126	0.732	0.137	0.552	0.103
Severe	TSI	0.407	0.222	0.425	0.177	0.321	0.192
	OSI	0.517	0.142	0.570	0.132	0.391	0.176
	coif1	0.526	0.161	0.544	0.160	0.402	0.139

higher than that of TSI in all cases. The fit of OSI in the PLSR and SMLR models for mild and moderate damage, as well as in the SVMR model for severe damage, was marginally higher than that of coif1. Compared with mild and severe damage, TSI, OSI, and coif1 showed the best fit for moderate damage. The above characteristics were consistent with the accuracy of the corresponding chlorophyll content estimation model.

#### 4. Discussion

The results of the sensitivity of TSI and CWC to RCC (Table 2 and Fig. 3) showed that bands with high sensitivity ( $R^2 > 0.6$ ) of mSR705, PSSRb, and coif1 (scale 5), such as 593–596 and 601–609 nm, were all in the visible light band, where various pigments are the main regulatory factors of the spectral response of plants, where the role of chlorophyll is particularly important. The plant cell structures of coniferous trees infested by pests are directly damaged due to overfeeding by larvae, resulting in a reduction in the number of green coniferous leaves and inhibited chlorophyll synthesis. Under different degrees of disaster, the position and slope of the red edge and green peak are sensitive to the change of RCC per unit area of plants and can effectively respond to plant stress (Li et al., 2015; He et al., 2018). The optimized mSR705 adopts the wavelengths of 705 and 750 nm, far from the main absorption band of chlorophyll, thereby avoiding the interference of supersaturation and maintaining the highest sensitivity to change in RCC (Schlemmer et al., 2013). The TCARI index is calculated using reflectance at the minimum and maximum absorption bands of vegetation chlorophyll, which reduces canopy structure and soil background (Haboudane et al., 2002; Liao et al., 2013). The hyperspectral features composed of optimized mSR705 and TCARI consider the effects of canopy structure, soil background, and saturation conditions and are

more sensitive and responsive to RCC changes. Considering the total process and different degrees of damage (Tables 4 and 5), the accuracies of the three OSI models are not only better than that of TSI but are also close to or even exceed those of CWCs. Therefore, this SI optimization and combination model can effectively extract the index sensitive to changes in RCC and compose SI features with a better estimation potential, which can be further applied to estimate other physiological indicators of trees. TSI, OSI, and coif1 showed the highest estimation accuracy at a moderate degree of damage. During a pest outbreak, the internal biochemical parameters and external structure of trees would change with the severity of disaster (Cheng et al., 2014).

In the case of a mild degree of damage, the RCC changes slightly, while in the case of moderate and severe degrees of damage, which is more conducive to capturing spectral information sensitive to changes in RCC, the RCC changes significantly. However, in case of a severe degree of damage, the RCC remains at a very low level, and the rate of leaf loss reaches the highest (Huang et al., 2019). By contrast, the interference of supersaturation and soil background is small under conditions involving moderate damage, which is conducive to the estimation of RCC. At a moderate degree of damage, the accuracy of the PLSR model of coif1 was lower and higher than OSI in SVMR and SMLR models, respectively. CWCs have excellent information extraction ability and can capture changes in forest biochemical indicators; therefore, they have been used in the estimation of RCC, nitrogen, water content, and other biochemical components (Song et al., 2011; Cheng et al., 2014). Bior1.5, which performed well in the RCC estimation model for the total damage process, also showed a high estimation accuracy in studies of forest water content conducted by Fang and Ju (2015) and Cheng et al. (2011). This study did not achieve good estimation accuracy under mild and severe damage conditions. In future research, structural parameters, such as the

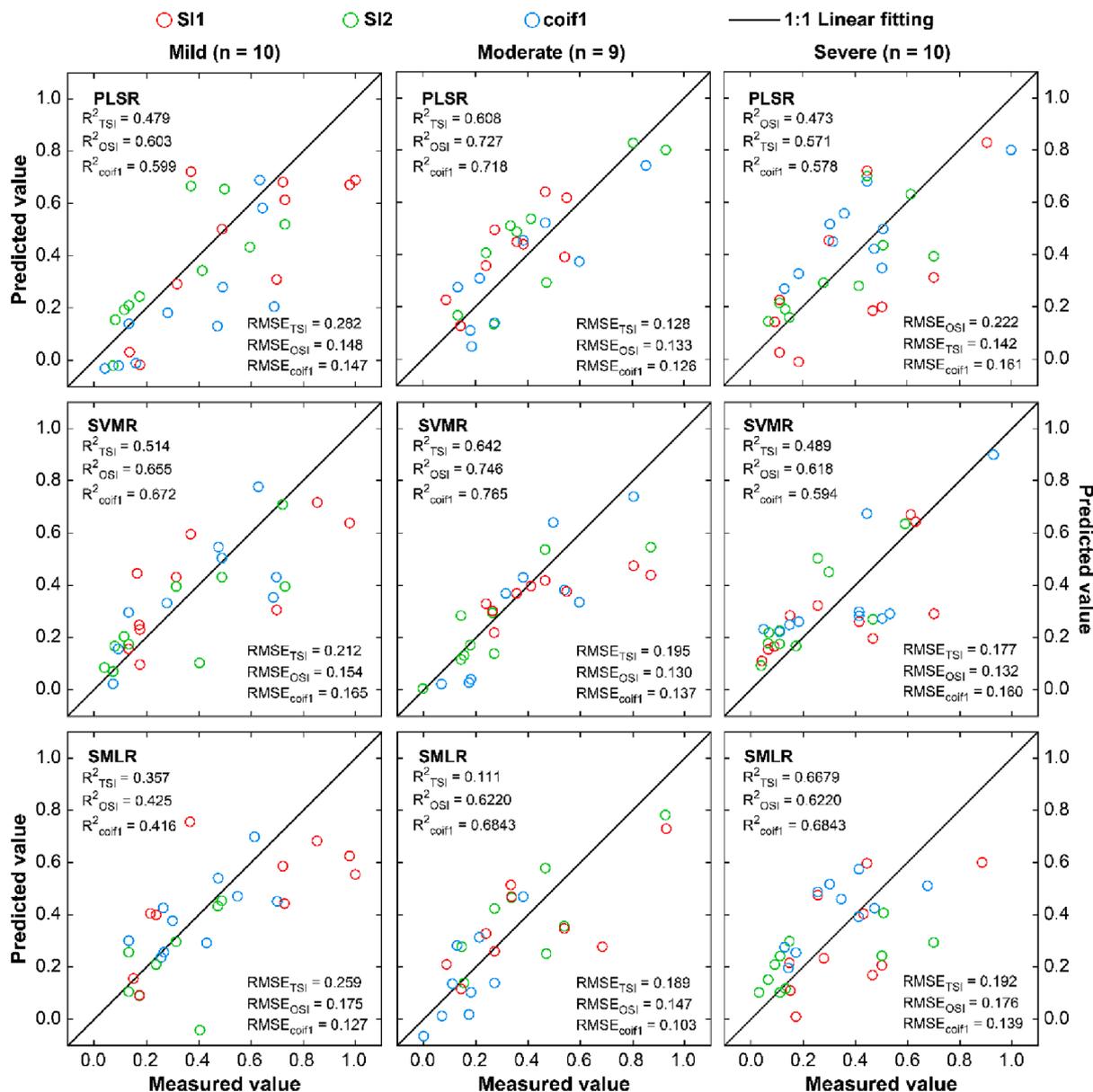


Fig. 6. 1:1 straight-line fitting based on the estimation model of the validation set for different damage levels.

cell diameter of needles can be obtained by combining hyperspectral and LiDAR data (Lin et al., 2016; Rahman et al., 2022), and used to improve the estimation accuracy of RCC under mild and severe damage conditions.

### 5. Conclusions

In this study, we used TSI, OSI, and CWCs to estimate the RCC corresponding to total damage as well as to different degrees of damage. The bands corresponding to the RCC response were distributed more in the green light region and the red edge band.

Regarding the total process, the estimation capability of OSI was significantly improved compared with that of TSI, and better than that of CWCs. Regarding different degrees of damage, the estimation performance of OSI was close to or even better than that of coil1. TSI, OSI, and coil1 showed the highest estimation accuracy at moderate degrees of damage. Changing one or more wavelength positions of SI, allowed the index to be better adjusted to any specific absorption region, which improved the estimation performance of TSI, suggesting a new method that involves using SI to estimate RCC and other forest biochemical

components. Notably, forest diseases and insect pests have the characteristics of large outbreak areas and fast spread, and the application of drones and aerospace remote sensing image data is more widespread (Hunt and Rondon, 2017; Tane et al., 2018; Li et al., 2018). Using wavelet coefficients to carry out relevant research will generate intensive calculations, and the data processing and information extraction process is also relatively cumbersome. In contrast, the optimized spectral index can reach a similar estimation accuracy to that of wavelet coefficients, but it is more convenient and faster to use. This study optimized the spectral indices based on data obtained from handheld hyperspectral instruments. The optimized indices can be applied to UAV hyperspectral data, greatly facilitating large-scale research. However, some band information for optimized indices cannot be corresponded one-to-one in hyperspectral satellites, resulting in the inability of the optimized indices to be directly used in some satellite data. Therefore, we intend to further explore the potential of Fp-SPA algorithm in combining spectral indices to improve the accuracy of estimating chlorophyll content using satellite data by optimizing the combination of spectral indices. This study will help to facilitate future research on RCC remote sensing monitoring at a large regional scale based on

multivariate hyperspectral data and plays an important part in the early identification and severity estimation of the locust pest, *E. Jacobson* Djak. Our findings can help local forestry departments to efficiently perform pest control work.

## Funding

This research was funded by the Natural Science Foundation of Inner Mongolia Autonomous Region (grant number: 2022MS04005), Inner Mongolia Autonomous Region Science and Technology Project (grant number: 2021GG0183), Inner Mongolia University Youth Science and Technology Talent Support Program Project (grant number: NJYT22030), and Inner Mongolia Normal University Introduces High-level Talents Scientific Research Start-up Funding Project (grant number: 2020YJRC051).

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Data will be made available on request.

## Acknowledgments

We thank the researchers at the Institute of Biology and the Institute of Ecology of the Mongolian Academy of Science for their contribution toward data collection in the field and their cooperation for the smooth progress of this study.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ecolind.2023.110714>.

## References

- Astola, H., Seitsonen, L., Halme, E., Molinier, M., Lönnqvist, A., 2021. Deep neural networks with transfer learning for forest variable estimation using sentinel-2 imagery in boreal forest. *Remote Sens.* 13, 2392–2432.
- Bai, L.G., Huang, X.J., Dashzebegd, G., Ariunaad, M., Nanzadd, T., Dorjsuren, A., Bao, G., Tong, S.Q., Bao, Y.H., Yin, S., Davadorj, E., 2021. Estimation of the population density of *Erannis jacobsoni* (Lepidoptera: Geometridae) based on hyperspectral features. *Acta Entomol. Sin.* 64, 711–721. <https://doi.org/10.16380/j.kcxb.2021.06.007>.
- Blackburn, G.A., 1998. Spectral indices for estimating photosynthetic pigment concentrations: A test using senescent tree leaves. *Int. J. Remote Sens.* 19, 657–675. <https://doi.org/10.1080/014311698215919>.
- Broge, N.H., Leblanc, E., 2001. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sens. Environ.* 76, 156–172. [https://doi.org/10.1016/S0034-4257\(00\)00197-8](https://doi.org/10.1016/S0034-4257(00)00197-8).
- Cheng, T., Rivard, B., Sánchez-Azofeifa, A., 2011. Spectroscopic determination of leaf water content using continuous wavelet analysis. *Remote Sens. Environ.* 115, 659–670. <https://doi.org/10.1016/j.rse.2010.11.001>.
- Cheng, T., Riaño, D., Ustin, S.L., 2014. Detecting diurnal and seasonal variation in canopy water content of nut tree orchards from airborne imaging spectroscopy data using continuous wavelet analysis. *Remote Sens. Environ.* 143, 39–53. <https://doi.org/10.1016/j.rse.2013.11.018>.
- de la Fuente, B., Saura, S., Beck, P.S.A., Fortin, M.-J., 2018. Predicting the spread of an invasive tree pest: the pine wood nematode in Southern Europe. *J. Appl. Ecol.* 55 (5), 2374–2385.
- Fang, M., Ju, W., 2015. An inversion model for remote sensing of leaf water content based on the leaf optical property. *Spectrosc. Spect. Anal.* 35, 167–171. [https://doi.org/10.3964/j.issn.1000-0593\(2015\)01-0167-05](https://doi.org/10.3964/j.issn.1000-0593(2015)01-0167-05).
- Farhadur Rahman, M.d., Onoda, Y., Kitajima, K., 2022. Forest canopy height variation in relation to topography and forest types in central Japan with LiDAR. *Forest Ecol. Manag.* 503, 119792.
- Foster, A.C., Walter, J.A., Shugart, H.H., Sibold, J., Negron, J., 2017. Spectral evidence of early-stage spruce beetle infestation in Engelmann spruce. *Forest Ecol. Manag.* 384, 347–357. <https://doi.org/10.1016/j.foreco.2016.11.004>.
- Frampton, W.J., Dash, J., Watmough, G., Milton, E.J., 2013. Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. *ISPRS J. Photogrammetry Remote Sens.* 82, 83–92. <https://doi.org/10.1016/j.isprsjprs.2013.04.007>.
- Gitelson, A.A., Merzlyak, M.N., 1997. Remote estimation of chlorophyll content in higher plant leaves. *Int. J. Remote Sens.* 18 (12), 2691–2697.
- Haboudane, D., Miller, J.R., Tremblay, N., Zarco-Tejada, P.J., Dextraze, L., 2002. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sens. Environ.* 81, 416–426. [https://doi.org/10.1016/S0034-4257\(02\)00018-4](https://doi.org/10.1016/S0034-4257(02)00018-4).
- Hamzeh, S., Naseri A.A., AlaviPanah S. K., Mojaradi B., Bartholomeus H.M., Clevers J.G. P.W., Behzad M., 2013. Estimating salinity stress in sugarcane fields with spaceborne hyperspectral vegetation indices. *Int. J. Appl. Earth Obs. Geoinf.* 21, 282–290. <https://doi.org/10.1016/j.jag.2012.07.002>.
- 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. *Int. J. Remote Sens.* 39, 4059–4076. <https://doi.org/10.1080/01431161.2018.1454620>.
- Hornero, A., Hernández-Clemente, R., North, P.R.J., Beck, P.S.A., Boscia, D., Navas-Cortes, J.A., Zarco-Tejada, P.J., 2020. Monitoring the incidence of *Xylella fastidiosa* infection in olive orchards using ground-based evaluations, airborne imaging spectroscopy and Sentinel-2 time series through 3-D radiative transfer modelling. *Remote Sens. Environ.* 236, 1–15. <https://doi.org/10.1016/j.rse.2019.111480>.
- Huang, X., Xie, Y., Bao, Y., 2018. Suitable distribution area of Jas' larch inchworm in Mongolia Plateau. *J. Northwest. A&F Univ. (Nat. Sci. Ed.)* 46, 98–106.
- Huang, X.J., Xie, Y.W., Bao, Y.H., Bao, G., Qing, S., Bao, Y.L., 2019. Estimation of leaf loss rate in larch infested with *Erannis jacobsoni* Djak based on differential spectral continuous wavelet coefficient. *Spectrosc. Spect. Anal.* 39, 2732.
- Hunt, E.R., Rondon, S.I., 2017. Detection of potato beetle damage using remote sensing from small unmanned aircraft systems. *J. Appl. Rem. Sens.* 11, 1. <https://doi.org/10.1117/1.JRS.11.026013>.
- Imanyfar, S., Hasanlou, M., Zadeh, V.M., 2019. Mapping oak decline through long-term analysis of time series of satellite images in the forests of Malekshahi. *Iran. Int. J. Remote Sens.* 40, 8705–8726. <https://doi.org/10.1080/01431161.2019.1620375>.
- Iordache, M.D., Mantas, V., Baltazar, E., Pauly, K., Lewycky, N., 2020. A machine learning approach to detecting pine wilt disease using airborne spectral imagery. *Remote Sens.* 12, 1–23. <https://doi.org/10.3390/rs12142280>.
- Lausch, A., Heurich, M., Gordalla, D., Dobner, H.J., Gwilym-Margianto, S., Salbach, C., 2013. Forecasting potential bark beetle outbreaks based on spruce forest vitality using hyperspectral remote-sensing techniques at different scales. *Forest Ecol. Manag.* 308, 76–89. <https://doi.org/10.1016/j.foreco.2013.07.043>.
- Li, D., Li, F., Hu, Y.C., Bodo, M., Schmidhalter, U., 2016. Study on the estimation of nitrogen content in wheat and maize canopy based on band optimization of spectHal parameters. *Guang Pu Xue Yu Guang Pu Fen Xi Spectrosc Spect Anal.* 36, 1150–1157.
- Li, X., Liu, X., Liu, M., Wang, C., Xia, X., 2015. A hyperspectral index sensitive to subtle changes in the canopy chlorophyll content under arsenic stress. *Int. J. Appl. Earth Obs. Geoinf.* 36, 41–53. <https://doi.org/10.1016/j.jag.2014.10.017>.
- Li, C., Zhu, X., Wei, Y., Cao, S., Guo, X., Yu, X., Chang, C., 2018. Estimating apple tree canopy chlorophyll content based on Sentinel-2A remote sensing imaging. *Sci. Rep.* 8, 3756. <https://doi.org/10.1038/s41598-018-21963-0>.
- Liao, Q., Wang, J., Yang, G., Zhang, D., Li, H., Fu, Y., Li, Z., 2013. Comparison of spectral indices and wavelet transform for estimating chlorophyll content of maize from hyperspectral reflectance. *J. Appl. Remote Sens.* 7 (1), 073575.
- Lin, Q.N., Huang, H.G., Chen, L., Yu, L.F., Huang, K., 2016. Simulation of needle reflectance spectrum and sensitivity analysis of biochemical parameters of *Pinus yunnanensis* in different healthy status. *Guang Pu Xue Yu Guang Pu Fen Xi* 36, 2538–2545.
- Ma, L., Huang, X.J., Dashzebegd, G., Ariunaad, M., Nanzadd, T., Dorjsuren, A., Bao, G., Tong, S.Q., Bao, Y.H., Davadorj, E., 2022. Monitoring forest insect pests by different remote sensing sensors: Research progress and prospect. *Chinese Agr. Sci. Bull.* 38, 91–99. <https://doi.org/10.11924/j.issn.1000-6850.casb2021-0872>.
- Oeser, J., Pflugmacher, D., Senf, C., Heurich, M., Hostert, P., 2017. Using intra-annual Landsat time series for attributing forest disturbance agents in central Europe. *Forests* 8, 251–275. <https://doi.org/10.3390/f8070251>.
- Pałaś, K.W., Zawadzki, J., 2020. Sentinel-2 imagery processing for tree logging observations on the białowieża forest world heritage site. *Forests* 11, 857–873. <https://doi.org/10.3390/f11080857>.
- Peñuelas, J., Gamon, J.A., Griffin, K.L., Field, C.B., 1993. Assessing community type, plant biomass, pigment composition, and photosynthetic efficiency of aquatic vegetation from spectral reflectance. *Remote Sens. Environ.* 46, 110–118. [https://doi.org/10.1016/0034-4257\(93\)90088-F](https://doi.org/10.1016/0034-4257(93)90088-F).
- Pulitua, S., Saarelab, S., Gobakkena, T., Ståhlb, G., Næsseta, E., 2018. Combining UAV and Sentinel-2 auxiliary data for forest growing stock volume estimation through hierarchical model-based inference. *Remote Sens. Environ.* 204, 485–497. <https://doi.org/10.1016/j.rse.2017.10.007>.
- Rahimzadeh-Bajgiran, P., Weiskittel, A., Kneeshaw, D., MacLean, D., 2018. Detection of annual spruce budworm defoliation and severity classification using landsat imagery. *Forests* 9, 357. <https://doi.org/10.3390/f9060357>.
- Rumpf, T., Mahlein, A.K., Steiner, U., Oerke, E.C., Dehne, H.W., Plümer, L., 2010. Early detection and classification of plant diseases with support vector machines based on hyperspectral reflectance. *Comput. Electron. Agric.* 74, 91–99. <https://doi.org/10.1016/j.compag.2010.06.009>.
- Sakowska, K., Alberti, G., Genesio, L., Peressotti, A., Delle Vedove, G., Gianelle, D., Colombo, R., Rodeghiero, M., Panigada, C., Juszczak, R., Celesti, M., Rossini, M., Haworth, M., Campbell, B.W., Mevy, J.-P., Vescovo, L., Cendrero-Mateo, M.P.,

- Rascher, U., Miglietta, F., 2018. Leaf and canopy photosynthesis of a chlorophyll deficient soybean mutant. *Plant Cell Environ.* 41, 1427–1437. <https://doi.org/10.1111/pce.13180>.
- Schlemmer, M., Gitelson, A., Schepers, J., Ferguson, R., Peng, Y., Shanahan, J., Rundquist, D., 2013. Remote estimation of nitrogen and chlorophyll contents in maize at leaf and canopy levels. *Int. J. Appl. Earth Obs. Geoinf.* 25, 47–54.
- Shirazinejad, G., Zoj, M.V., Latifi, H., 2022. Applying multirate Sentinel-2 data for forest-type classification in complex broadleaf forest stands. *Forestry* 95, 363–379. <https://doi.org/10.20944/preprints202101.0235.v1>.
- Song, C., White, B.L., Heumann, B.W., 2011. Hyperspectral remote sensing of salinity stress on red (*Rhizophora mangle*) and white (*Laguncularia racemosa*) mangroves on Galapagos Islands. *Remote Sens. Lett.* 2, 221–230. <https://doi.org/10.1080/01431161.2010.514305>.
- Tane, Z., Roberts, D., Koltunov, A., Sweeney, S., Ramirez, C., 2018. A framework for detecting conifer mortality across an ecoregion using high spatial resolution spaceborne imaging spectroscopy. *Remote Sens. Environ.* 209, 195–210. <https://doi.org/10.1016/j.rse.2018.02.073>.
- Wang, H.F., Huo, Z.G., Zhou, G.S., Laio, Q.-H., Feng, H.-K., Wu, L., 2016. Estimating leaf SPAD values of freeze-damaged winter wheat using continuous wavelet analysis. *Plant Physiol. Biochem.* 98, 39–45. <https://doi.org/10.1016/j.plaphy.2015.10.032>.
- Wang, J., Wang, T., Skidmore, A.K., Shi, T., Wu, G., 2015. Evaluating different methods for grass nutrient estimation from canopy hyperspectral reflectance. *Remote Sens.* 7, 5901–5917. <https://doi.org/10.3390/rs70505901>.
- Williams, P.T., Wynne, R.H., Thomas, V.A., DeFries, R., 2021. Mapping smallholder forest plantations in Andhra Pradesh, India using multitemporal harmonized landsat sentinel-2 S10 data. *Land Degrad. Dev.* 32, 4212–4226. <https://doi.org/10.1002/ldr.4027>.
- Wu, N., Liu, J., Yan, R., Zhou, G., 2012. Using spectral feature parameters to estimate the chlorophyll content of Chinese fir under disease stress. *Plant Prot.* 38 (72–76), 88. <https://doi.org/10.3390/rs70505901>.
- Wu, C., Niu, Z., Tang, Q., Huang, W., 2008. Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation. *Agric. For. Meteorol.* 148, 1230–1241. <https://doi.org/10.1016/j.agrformet.2008.03.005>.
- Xi, G.L., Huang, X.J., Bao, Y.H., Bao, G., Tong, S.Q., Dashzebegd, G., Nanzadd, T., Dorjsuren, A., Davaadorj, E., Ariunaad, M., 2020. Hyperspectral discrimination of different canopy colors in *Erannis jacobsoni* Djak-infested larch. *Spectrosc. Spect. Anal.* 40, 2925–2931.
- Zhang, S.N., Wang, Q., Jin, J., Xu, L., Guan, H.Y., 2016. Application of hyperspectral indices for estimating leaf chlorophyll content of assimilating shoots of *Tamarix ramosissima*. *Arid Zone Res.* 33, 1088–1097.
- Zhang, W., Wang, X.M., Pan, Q.M., Xie, J.Z., Zhang, J.S., Meng, P., 2018. Hyperspectral response characteristics and chlorophyll content estimation of *Phyllostachys violascens* leaves under drought stress. *Acta Ecol. Sin.* 38, 6677–6684. <https://doi.org/10.5846/stxb20180330>.

### Further reading

- Lin, Q., Huang, H., Wang, J., Huang, K., Liu, Y., 2019. Detection of pine shoot beetle (PSB) stress on pine forests at individual tree level using UAV-based hyperspectral imagery and lidar. *Remote Sens.* 11, 2540. <https://doi.org/10.3390/rs11212540>.