Peer

Classification of major species in the *sericite–Artemisia* **desert grassland using hyperspectral images and spectral feature identification**

Wenhao Liu**[*](#page-0-0)** , Wanqiang Han**[*](#page-0-0)** , Guili Jin, Ke Gong and Jian Ma

College of Grassland Sciences of Xinjiang Agricultural University, Xinjiang Agriculture University, Urumqi, Xinjiang, China

These authors contributed equally to this work.

ABSTRACT

Background. The species composition of and changes in grassland communities are important indices for inferring the number, quality and community succession of grasslands, and accurate monitoring is the foundation for evaluating, protecting, and utilizing grassland resources. Remote sensing technology provides a reliable and powerful approach for measuring regional terrain information, and the identification of grassland species by remote sensing will improve the quality and effectiveness of grassland monitoring.

Methods. Ground hyperspectral images of a *sericite*–*Artemisia* desert grassland in different seasons were obtained with a Soc710 VP imaging spectrometer. First-order differential processing was used to calculate the characteristic parameters. Analysis of variance was used to extract the main species, namely, *Seriphidium transiliense* (Poljak), *Ceratocarpus arenarius* L., *Petrosimonia sibirica* (Pall), bare land and the spectral characteristic parameters and vegetation indices in different seasons. On this basis, Fisher discriminant analysis was used to divide the samples into a training set and a test set at a ratio of 7:3. The spectral characteristic parameters and vegetation indices were used to identify the three main plants and bare land.

Results. The selection of parameters with significant differences ($P < 0.05$) between the recognition objects effectively distinguished different land features, and the identification parameters also differed due to differences in growth period and species. The overall accuracy of the recognition model established by the vegetation index decreased in the following order: June $(98.87%)$ > September $(91.53%)$ > April (90.37%). The overall accuracy of the recognition model established by the feature parameters decreased in the following order: September $(89.77%)$ > June $(88.48%)$ > April (85.98%).

Conclusions. The recognition models based on vegetation indices in different months are superior to those based on feature parameters, with overall accuracies ranging from 1.76% to 9.40% higher. Based on hyperspectral image data, the use of vegetation indices as identification parameters can enable the identification of the main plants in *sericite– Artemisia* desert grassland, providing a basis for further quantitative classification of the species in community images.

Submitted 23 October 2023 Accepted 10 June 2024 Published 18 July 2024

Corresponding author Guili Jin, jguili@126.com

[Academic editor](https://peerj.com/academic-boards/editors/) [Ahmet Tansel Serim](https://peerj.com/academic-boards/editors/)

[Additional Information and](#page-20-0) [Declarations can be found on](#page-20-0) [page 21](#page-20-0)

DOI **[10.7717/peerj.17663](http://dx.doi.org/10.7717/peerj.17663)**

Copyright 2024 Liu et al.

[Distributed under](https://creativecommons.org/licenses/by-nc/4.0) [Creative Commons CC-BY-NC 4.0](https://creativecommons.org/licenses/by-nc/4.0)

OPEN ACCESS

Subjects Agricultural Science, Plant Science, Natural Resource Management **Keywords** Grassland plants, Spectral characteristics, Identification parameters, Screening, Fisher discriminant

INTRODUCTION

The convenient and rapid identification of the composition of and changes in natural grassland communities is highly important for identifying grassland type, quantity and quality and monitoring grassland degradation and restoration. The development of remote sensing technology compensates for the limitations of human and material resources in traditional grassland resource investigation and monitoring methods (*[Reinermann, Asam](#page-22-0) [& Kuenzer, 2020](#page-22-0)*). Hyperspectral remote sensing, one such technology, is mainly used in spectral feature extraction because of its many bands and narrow channels (*[Yu et al.,](#page-23-0) [2018](#page-23-0)*) and its strong advantages in grassland vegetation classification and identification (*[Qiao et al., 2022](#page-22-1)*). The identification and classification of plant species are the basis of biodiversity monitoring. Clarifying the categories and distributions of plant species has important implications for monitoring ecosystem biodiversity, community reconstruction, and maintaining ecological function (*[Khdery & Yones, 2021](#page-22-2)*). Species composition is an important parameter that describes the structure of vegetation communities and indicates biodiversity and ecosystem health in a region (*[Lyu et al., 2020](#page-22-3)*), and species identification is crucial for environmental monitoring (*[Zhang et al., 2020](#page-23-1)*). Previous studies mostly selected characteristic bands from the period of vigorous vegetation growth to analyse and compare different grassland vegetation or transformed the characteristic bands and used quantitative spectral characteristic parameters and vegetation indices to distinguish different vegetation types, obtaining good research results (*[Melville, Lucieer & Aryal, 2018](#page-22-4)*; *[Yu et al., 2018](#page-23-0)*; *[Frank et al., 2022](#page-21-0)*). Spectral parameters can enhance the ability to separate plants (*[Dmitriev et al., 2022](#page-21-1)*; *[Soubry & Guo, 2021](#page-22-5)*). However, due to the phenological characteristics of vegetation, the selected parameters respond differently depending on the phenological period (*[Zhang et al., 2018](#page-23-2)*), so it is necessary to screen parameters with high response degrees for vegetation classification research. For example, *[Han et al. \(2020\)](#page-21-2)* successfully identified seven species of millet (*Setaria italica*) using visible near infrared (VIS-NIR) spectral information from growing leaves and established a recognition model with a correlation coefficient of up to 0.9994. *[Feng et al. \(2016\)](#page-21-3)* used ground-measured hyperspectral data to identify indicator species of grassland degradation, namely, *Artemisia frigida* Willd, *Stellera chamaejasme* L, and *Potentilla acaulis* L., using a mixed spectral model and found that the recognition accuracy of these three species was 83.3%, 88.9%, and 94.4%, respectively. *[Li et al. \(2016\)](#page-22-6)* identified red, blue and near-infrared bands as sensitive bands and obtained the normalized vegetation index of the blue band $(NDVI_{blue})$ as the best parameter through analysis, and the extraction accuracy of *Stellera chamaejasme* Linn reached 90%. To further strengthen research on vegetation classification and recognition, selecting an appropriate identification period and identification parameters is an important step, especially for grassland communities with no obvious individual plant characteristics (*[Singh et al., 2018](#page-22-7)*).

The ecosystem of the *Seriphidium transiliense* (*Poljak*) desert grassland is fragile and concentrated in Xinjiang, China. Overgrazing and general degradation have seriously damaged the animal husbandry production of herders and threatened the biodiversity of grasslands and the ecological security of oases (*[Jin et al., 2011](#page-21-4)*). At present, hyperspectral research on *S. transiliense* desert grasslands has focused on the analysis of vegetation spectral characteristics, and estimation models of grassland coverage and biomass have been established by using spectral reflectance. *[He et al. \(2014\)](#page-21-5)* found obvious differences among an invasive species (*Polygonum aviculare* L), an increased species (*Petrosimonia sibirica* (*Pall*)) and a reduced species (*S. transiliense*) in the near-infrared band, and in terms of spectral reflectance, the rank was bare ground > *P. aviculare* > *S. transiliense* > *P. sibirica*. *[Wei et al. \(2017a\)](#page-23-3)* and *[Wei et al. \(2017b\)](#page-23-4)* estimated grassland coverage by using specific sensitive bands and the vegetation index through ground and satellite hyperspectral images, and the accuracy of the best model exceeded 85%. They estimated the aboveground biomass inside and outside the fence, and the differences in the estimation models were due to the different types and seasons of the measured grassland communities. The above studies mainly considered the red and near-infrared bands of vegetation, with limited analysis of other bands, and did not include in-depth research or analysis on the identification period and identification parameters of the main species in the focal community.

Therefore, to further improve the identification accuracy and utilization of the identification parameters for the main plants in a grassland community, this study used an imaging spectrometer to collect spectral images of the *S. transiliense* desert grassland community, established spectral characteristic parameters and vegetation indices, and analysed these spectral characteristic parameters and the vegetation indices of the main plants. The sensitive identification parameters in different periods were selected, and the identification objects were classified by Fisher analysis to provide a theoretical basis for the classification and identification of grassland plants.

MATERIALS & METHODS

Study site

The study area is located in the middle part of the northern slope of the Tianshan Mountains in the *S. transiliense* habitat in Baiyanghe Township, Midong District, Urumqi. The geographical location is between $87^{\circ}52'59''$ and $87^{\circ}55'13''E$ and $44^{\circ}00'16''$ and $44^{\circ}01'20''N$ [\(Fig. 1\)](#page-3-0). The mean annual temperature is 16 °C, the highest temperature is 39 °C in July and August, the lowest temperature is −31 ◦C in January, the altitude is 895–954 m, and the annual precipitation is 170–276 mm (National Earth System Science Data Center). The area has a temperate continental semiarid and semidesert climate.

The grassland types in this area are representative and have good growth conditions. The area is home to an important local spring and autumn pasture. The community characteristics are shown in [Table 1.](#page-4-0) The dominant species is *S. artemisia*, and the subdominant species are the annual plants *Ceratocarpus arenarius* L and *P. sibirica.* There are also scattered distributions of *Carpesium abrotanoides* L., *Ceratocephalus orthoceras* DC,

Draba nemorosa L., *Trigonella arcuata* CA, *Koelpinia linearis* (Pall), *Stipa capillata* L. and so on. The 0–5 cm soil moisture content is 11.32%, the soil bulk density is 1.21 $g/cm³$, the conductivity is 199.33 us/cm, the pH is 7.87, the soil C content is 9.10 g/kg, the soil N content is 1.25 g/kg, the soil P content is 0.73 g/kg, and the soil K content is 9.93 g/kg.

Data acquisition

Sample plot establishment

In the study area, a 500 m long transect was established every 150 m, with a total of six parallel transects. Five 30 m \times 30 m transects with a spacing of 30 m were arranged along

Notes.

The numbers in brackets are the importance values of the species, which were calculated with the formula (relative height + relative coverage + relative density + relative biomass)/4.

each transect. Five transects were arranged on two diagonal lines of the transects with a 5-point sampling method as the data acquisition points. According to the field of view of the instrument, the area of the transect was set as 0.5 m \times 0.6 m, with a total of 150 transects [\(Fig. 2\)](#page-4-1).

Hyperspectral image acquisition

An SOC710VP imaging spectrometer was used to collect spectral data of plant communities in the field (spectral range, 400∼1,000 nm; resolution, 4.68 nm; number of bands, 128; and imaging resolution, $1,392 \times 1,040$ pixels), and the spectral measurement time ranged from 11:00 to 16:00 in April, June and September 2018 (solar altitude angle $> 45^{\circ}$). The measurements were taken on sunny, cloudless, and windless or less windy (wind $<$ 3 m/h) days. During measurement, the lens was vertically downwards, and the vertical height from the plant canopy was approximately 1.0 m. Based on the analysis of the spectral characteristics of the three main plants, the other plants distributed sporadically in the quadrat were extracted, and the community canopy spectrum and bare-ground

hyperspectral images were collected in each quadrat. A total of 450 images were collected. In addition, the coverage and biomass of each species in the quadrat were measured to provide a reference for the extraction of spectral data.

Extraction of spectral data

The spectral data of pure pixels of *S. transiliense, C. arenarius, P. sibirica* and bare ground were extracted from the images of plant communities in each measurement period by using SRAnal 710 software of the imaging spectrometer. A total of 150 quadrat images were obtained in each period, yielding 150 reflectance samples of three types of plants and bare land from which the damaged samples were removed. To ensure the consistency of the reflectance samples, 90 reflectance samples of each type were ultimately classified: 360 reflectance samples in April (spring), 273 reflectance samples in total due to the scarcity of *P. sibirica* in June (summer), and 184 samples in total due to the scarcity of *C. arenarius* in September (autumn).

Spectral data processing *First-order differential processing*

The first-order differential can eliminate the influence of partial atmospheric and soil backgrounds, highlight the characteristics of the vegetation spectrum, and be beneficial to the extraction of vegetation information such as vegetation indices and the leaf area index.

The first-order differential method was selected to process the spectrum, with the following formula:

$$
FR(\lambda_i) = \frac{R(\lambda_{i+1}) - R(\lambda_{i-1})}{2\Delta\lambda}
$$
\n(1)

where $FR(\lambda_i)$ is the first-order differential spectrum of wavelength λ_i ; $R(\lambda_{i+1})$ is the original spectral reflectance of wavelength *i+1*; *R*(λ*i*−1) is the original spectral reflectance of wavelength *i-1*; and $\Delta\lambda$ is the wavelength difference between wavelength *i* and wavelength *i+1*.

Selection of spectral parameters

To explore the spectral response characteristics of the main species and bare land of the *S. transiliense* desert grassland, 10 common spectral characteristic parameters were selected, including eight location parameters, two area parameters and 11 vegetation indices, including seven wide-band vegetation indices and four narrow-band vegetation indices, as shown in [Tables 2](#page-6-0) and [3.](#page-6-1)

Fisher discriminant analysis is used to establish a linear discriminant function based on the original information of samples of known categories or the feature space information that can represent the information of samples of known categories, project the multidimensional spatial space to one-dimensional space, classify and discriminate the samples of unknown categories through the discriminant criteria, and then verify the classification accuracy of the sample classification information (*[Huang et al., 2019](#page-21-6)*).

$$
S_{\omega} = \sum_{i=1}^{C} \sum_{X \in X_i} (x - m_i)(x - m_i)^T, i = 1, ..., C
$$
 (2)

Table 2 Characteristic parameters.

Table 3 Vegetation index.

$$
S_b = \sum_{i=1}^{C} (x - m_i)(x - m_i)^T
$$

\n
$$
\overline{s_{\omega}} = \sum_{i=1}^{C} \sum_{y \in Y_i} (y - \overline{m_i})^2, i = 1, ..., C
$$

\n
$$
J_F(\omega) = \frac{\omega^T S_b \omega}{\omega^T S_{\omega} \omega}
$$
\n(5)

$$
L(\omega, \lambda) = \omega^T S_b \omega - \lambda (\omega^T S_{\omega} \omega - c)
$$
\n⁽⁶⁾

$$
S_b \omega^* = \lambda S_{\omega} \omega^* \tag{7}
$$

where *mⁱ* is the mean vector of each class of samples in the original high-dimensional space; $\overline{m_i}$ is the mean vector of each class of samples in the one-dimensional space Y after projection; m is the mean vector of all of the samples; *X* is the sample point vector; S_{ω} is the discrete matrix within the sample class; S_b is the discrete matrix between sample classes; $\overline{s_{\omega}}$ is the pooled within-class scatter matrix; *J_F*(ω) is the Fisher criterion function; *L*(ω , λ) is the Lagrange function; λ is the Lagrange multiplier; and ω^* is the maximum value, which is the best projection method.

Screening of sensitive spectral parameters

Variance analysis and a discriminant model were used to screen the eight location parameters, two area parameters and 12 vegetation indices, and the, contributions of these parameters to the identification of the four objects were determined. The *P* value between two identified objects was determined by variance analysis. The smaller the *P* value is, the more significant the difference is. All 10 spectral characteristic parameters and 12 vegetation index parameters were taken as input parameters to obtain the initial discrimination accuracy. The characteristic parameters were removed one by one according to the *P* value. If the discrimination accuracy of the model after removing the parameters was greater than or equal to the initial accuracy, the parameter was removed; otherwise, the parameter was retained in the discrimination model [\(Fig. 3\)](#page-8-0).

Data analysis

The spectral reflectance values of different recognition objects were extracted using SRAnal710 and imported into Excel 2016 to remove abnormal data. First-order differential processing was performed, and feature parameters and vegetation indices were calculated. Principal component analysis was performed in SPSS 20.0 software to screen feature bands, and significance tests were conducted on the feature parameters of each identified object and vegetation index. Fisher discriminant analysis was used to classify samples and verify accuracy. SigmaPlot 14.0 software was used for plotting.

RESULTS

Characteristic parameter analysis

In April, the six spectral parameters of blue edge amplitude M_b , green peak amplitude M_g , red valley amplitude $\rm M_r$, red edge amplitude $\rm M_{re}$, blue edge area $\rm A_b$ and red edge area $\rm A_{re}$ showed significant differences among the four recognition objects ($P < 0.05$); there were significant differences in various spectral parameters between bare land and the three plants $(P < 0.05)$. The spectral parameters in June and September showed significant differences among the four objects, including the red valley position L_r and the red edge area A_{re} , which were significantly different among the identified objects ($P < 0.05$, [Table 4\)](#page-9-0).

Vegetation index analysis

In April, the significant differences in the NDVI, RVI, BNDVI, GNDVI, EVI, PSRI and REEDVI were observed among the four objects were found ($P < 0.05$); there was a significant difference between bare land and the three plants, except for the PRI (*P* < 0.05, [Table 5\)](#page-10-0).

In June, there were significant differences in six vegetation indices (DVI, BNDVI, GNDVI, NDGI, PSRI and VOG) among the four detection objects (*P* < 0.05), but there were no significant differences in the NDVI, RVI or EVI between bare land and *C. arenarius*. The DVI, NDGI and VOG differed from those in April.

In September, there were significant differences among the seven vegetation indices: the NDVI, DVI, EVI, NDGI, PSRI, RENDVI and VOG (*P* < 0.05). The NDVI, EVI and PSRI were the same in June and April; and the DVI, NDGI, PSRI and VOG were the same in June and April. Only the EVI differed from that in June and April.

Notes. The different lowercase letters in the table indicate significant differences between different recognition objects within the same characteristic parameter and season (*^P* < 0.05).

Liu et al. (2024),

Notes.

The different lowercase letters in the table indicate significant differences between different recognition objects within the same vegetation index and season (*^P* < 0.05).

Table 6 Mean *P* **value from variance analysis of characteristic parameters among the identified objects.**

Table 7 Mean *P* **value from variance analysis of the vegetation indices among the identified objects.**

Discriminant parameter analysis and screening *Discriminant parameter analysis*

In April, there were no significant differences in the characteristic parameters among the identified objects, which were L_b , L_{re} , L_g and A_{re} . For June, the parameters were L_b , L_{re} , $\rm L_g$, $\rm M_g$, $\rm M_r$, $\rm M_{re}$ and $\rm A_b$, and for September, they were $\rm L_b$, $\rm L_{re}$, $\rm M_b$, $\rm L_g$, $\rm M_g$, $\rm M_r$, $\rm M_{re}$ and $\rm A_b$ [\(Table 6\)](#page-11-0).

In April, there were no significant differences in the vegetation indices among the identified objects, which were the DVI, NDGI, RI, PRI, PSRI, and VOG. For June, the measured indices were the PRI and RENDVI, and for September, they were the RVI, BNDVI, GNDVI and PRI [\(Table 7\)](#page-11-1).

Discriminant parameter screening

Since the average *P* value can indirectly reflect the size of the difference, the parameters that did not have very significant differences were eliminated one by one from small to large according to the *P* value ($P < 0.01$), and the discrimination accuracy of the four recognition objects and the total discrimination accuracy of the test set were calculated. The group with more eliminated parameters and an accuracy no less than that of the noneliminated parameters was taken as the optimal set of parameters. The eliminated parameters and retained parameters of sericite, *S. transiliense*, *C. arenarius*, *P. sibirica* and bare land in April, June and September are shown in [Tables 8](#page-12-0) and [9.](#page-12-1)

Recognition object discrimination

Characteristic parameter discrimination

In April, 1,200 samples of the three plants and bare land were divided into a training set and a test set at a ratio of 7:3. Taking the selected spectral characteristic parameters as input variables, a discrimination model was established, and the recognition effect is shown in [Fig. 4.](#page-13-0) Discriminant analysis of *S. transiliense* (A), *C. arenarius* (B), *P. sibirica*

Table 8 Characteristic parameter screening of the identified objects in different seasons.

Table 9 Vegetation index screening of the identified objects in different seasons.

(C) and bare land (D) was performed. When classifying *S. transiliense* and the other three objects, *S. transiliense* was incorrectly identified 17 times, with an accuracy of 81.11%. The other objects were incorrectly identified as *S. transiliense* 102 times, with an accuracy of 62.22%, yielding a total accuracy of 71.67%. When classifying *C. arenarius* with the other three objects, *C. arenarius* was incorrectly identified six times, with an accuracy of 95.56%, and other objects were incorrectly identified as *C. arenarius* 51 times, with an accuracy of 79.26%, yielding a total accuracy of 87.41%. When classifying *P. sibirica* and the other three objects, the other objects were incorrectly identified as *P. sibirica* 52 times, with an accuracy of 82.96%, yielding a total accuracy of 85.93%. When classifying the bare land and the three plants, the bare land and all plants were classified into the correct category with an accuracy of 100%.

Full-size [DOI: 10.7717/peerj.17663/fig-4](https://doi.org/10.7717/peerj.17663/fig-4)

In June, 910 samples of the three plants and bare land were divided into a training set and a test set at a ratio of 7:3. The test set samples of *S. transiliense, C. arenarius, P. sibirica* and bare land were classified. When classifying *S. transiliense* and the other three objects, *S. transiliense* was incorrectly identified four times, with an accuracy of 95.56%. The other objects were incorrectly identified as *S. transiliense* three times, with an accuracy of 98.36%, yielding a total accuracy of 96.96%. When classifying *C. arenarius* with the other three objects, *C. arenarius* was incorrectly identified 15 times, with an accuracy of 83.33%, and the other objects were incorrectly identified 32 times as *C. arenarius*, with an accuracy of 82.51%, yielding a total accuracy of 82.92%. When classifying *P. sibirica* and the other three objects, *P. sibirica* was incorrectly identified once, with an accuracy of 66.66%, and the other objects were incorrectly identified three times as *P. sibirica*, with an accuracy of 98.89%, yielding a total accuracy of 82.78%. When classifying the bare land and the three plants, the bare land and all plants were placed in the correct category, and the other objects

Full-size [DOI: 10.7717/peerj.17663/fig-5](https://doi.org/10.7717/peerj.17663/fig-5)

were incorrectly identified as bare land 32 times, with an accuracy of 82.51%, yielding a total accuracy of 91.26% [\(Fig. 5\)](#page-14-0).

A total of 614 samples of three plants and bare land were divided into a training set and a test set at a ratio of 7:3. When classifying *S. transiliense* and the other three objects, *S. transiliense* was incorrectly identified zero times, with an accuracy of 100%, and the other objects were incorrectly identified as *S. transiliense* zero times, with an accuracy of 100%, yielding a total accuracy of 100%. When classifying *C. arenarius* with other objects, *C. arenarius* was incorrectly identified once, with an accuracy of 75%, and the other objects were incorrectly identified as *C. arenarius* 56 times, with an accuracy of 68.89%, yielding a total accuracy of 71.95%. When classifying the bare land and the other two plants, all the bare land was placed in the correct category, and the other objects were incorrectly identified as bare land four times, with an accuracy of 95.74%, yielding a total accuracy of 97.37% [\(Fig. 6\)](#page-15-0).

Figure 6 (A–C) Classification based on characteristic parameters in September. Full-size [DOI: 10.7717/peerj.17663/fig-6](https://doi.org/10.7717/peerj.17663/fig-6)

Vegetation index model and accuracy evaluation

In April, 1,200 samples of three plants and bare land were divided into a training set and a test set at a ratio of 7:3. The samples in the test set were classified as *S. transiliense, C. arenarius, P. sibirica* and bare ground. When classifying *S. transiliense* and the other three objects, *S. transiliense* was incorrectly identified 10 times, with an accuracy of 88.89%. The other objects were incorrectly identified as *S. transiliense* 57 times, with an accuracy of 78.89%, yielding a total accuracy of 83.89%. When classifying *C. arenarius* with the other three objects, *C. arenarius* was incorrectly identified seven times, with an accuracy of 92.22%, and the other objects were incorrectly identified as *C. arenarius* 46 times, with an accuracy of 82.96%, yielding a total accuracy of 86.30%. When classifying *P. sibirica* and the other three objects, *P. sibirica* was incorrectly identified three times, with an accuracy of 96.67%. The other objects were incorrectly identified as *P. sibirica* 38 times, with an accuracy of 85.93%, yielding a total accuracy of 91.30%. When classifying the bare land and the three plants, the bare land and all plants were placed in the correct category, with an accuracy of 100% [\(Fig. 7\)](#page-16-0).

In June, 910 samples of three plants and bare land were divided into a training set and a test set at a ratio of 7:3. When classifying *S. transiliense* and the other three objects, *S. transiliense* was incorrectly identified zero times, with an accuracy of 100%. When classifying *C. arenarius* with the other three objects, *C. arenarius* was incorrectly identified as *S. transiliense* four times, with an accuracy of 95.56%, and the other objects were incorrectly identified as *C. arenarius* eight times, with an accuracy of 95.63%, yielding a total accuracy of 95.60%. When classifying *P. sibirica* and the other three objects, *P. sibirica* was incorrectly identified zero times, with an accuracy of 100%. Other objects were incorrectly identified as *P. sibirica* zero times, with an accuracy of 100%, yielding a total accuracy of 100%. When classifying the bare land and the three plants, the bare land and all plants were placed in the correct category, and the other objects were incorrectly identified as bare land 15 times, with an accuracy of 91.80%, yielding a total accuracy of 95.90% [\(Fig.](#page-17-0) [8\)](#page-17-0).

In September, 614 samples of three plants and bare land were divided into a training set and a test set at a ratio of 7:3. When classifying *S. transiliense* and the other three objects, *S. transiliense* was incorrectly identified zero times, with an accuracy of 100%, and the

Full-size [DOI: 10.7717/peerj.17663/fig-7](https://doi.org/10.7717/peerj.17663/fig-7)

other objects were incorrectly identified three times as *S. transiliense*, with an accuracy of 96.81%. The total accuracy was 98.41%. When classifying *C. arenarius* and other objects, *C. arenarius* was incorrectly identified once, with an accuracy of 75%, and the other objects were incorrectly identified as *S. transiliense* 25 times, with an accuracy of 86.11%, yielding a total accuracy of 80.56%. When classifying bare land and the other two plants, bare land was incorrectly identified five times, with an accuracy of 94.44%, and the other objects were incorrectly identified three times as bare land, with an accuracy of 96.81%, yielding a total accuracy of 95.63% [\(Fig. 9\)](#page-17-1).

DISCUSSION

Differences in identification parameters among seasons

The recognition parameters showed different responses in different recognition seasons and for different recognition objects, revealing a certain complexity. The parameters of the

Figure 9 (A–C) Classification based on the vegetation indices in September. Full-size [DOI: 10.7717/peerj.17663/fig-9](https://doi.org/10.7717/peerj.17663/fig-9)

same plant and the remaining recognition objects are significantly different in different seasons, and accurate classification can be achieved based on differences in plant phenology (*[Weisberg et al., 2021](#page-23-5)*). In this study, both the spectral characteristic parameters and the vegetation indices reflected this result; for example, there were six spectral characteristic parameters in April (the period of seedling establishment) for *S. transiliense*, which significantly differed from the number for other recognition objects, while there were two in June (the peak period) and September (the dry and yellow periods). The reason is that the spectra of the same plant differ depending on the phenological period (*[Ouyang et](#page-22-8) [al., 2013](#page-22-8)*; *[Soubry & Guo, 2021](#page-22-5)*; *[Wang et al., 2021](#page-23-6)*). Although the spectral characteristics of different plants are roughly the same, they are still different in terms of reflectivity or some individual bands (*[Ge et al., 2006](#page-22-9)*; *[Liang, Yang & Xing, 2021](#page-22-10)*; *[Zhou, Li & Xu, 2022](#page-23-7)*). Both the characteristic parameters and vegetation indices are based on the spectra, explaining this result. However, in June (peak period) and September (dry and yellow period), there were red valley positions L_r and red edge areas. These two spectral characteristic parameters showed significant differences among the identified objects, which may be because the red light band is the most sensitive to vegetation and because the *S. transiliense* desert grassland is affected by drought year round. Vegetation reflectance in the dry and yellow seasons is similar to the spectral reflectance and characteristics of litter on the ground (*[He et al.,](#page-21-5) [2014](#page-21-5)*).

Identification parameter screening

This study included 10 spectral characteristic parameters and 12 vegetation indices. Through verification *via* analysis of variance and the discrimination model, it was found that the characteristic parameters and vegetation indices removed and retained in different seasons were different. When these parameters are eliminated one by one, the precision is improved or the same compared with the original precision, which shows that the eliminated parameters do not play a corresponding role in the discrimination of the four recognition objects; in contrast, there will be errors in the identification results due to the similarity of the information included. For example, in April, the three parameters $L_{\rm b}$, $L_{\rm r}$, and $L_{\rm re}$ show only small, nonsignificant differences among plants and are thus eliminated, while L_g does not show significant differences among the four recognition objects but is retained. This may be because L_g is strongly affected by plant chlorophyll content (*[Datt, 2010](#page-21-7)*). *[Dmitriev et al. \(2022\)](#page-21-1)* also reported that the vegetation indices used to identify *Populus* plants vary among periods, which is related to the chlorophyll content and photosynthetic intensity. Therefore, whether there are significant differences in parameters among plants cannot be used as the only criterion for elimination, and different characteristic parameters have different contributions and sensitivities in the determination of recognition objects. For example, through a continuous projection algorithm, *[Cao et](#page-21-8) [al. \(2023\)](#page-21-8)* reported that among 27 spectral parameters, colour parameters played a more important role in the classification of 18 oolong tea cultivars and can be accurately divided into early, mesophytic, and late species, with a recognition accuracy of more than 90%.

Identification of different parameters in the same season

In the same season, the characteristic parameters and vegetation indices were used to classify and identify the three types of plants and bare land. The discrimination accuracy of bare land is always high because the spectral reflectance and characteristics of bare land are quite different from those of vegetation (*[Qiao et al., 2022](#page-22-1)*); therefore, bare land had little effect on the ground hyperspectral identification of the three types of plants. However, there were some differences in the accuracy obtained by using different types of recognition parameters; for example, in April (returning green stage), the accuracy of the characteristic parameters was 85.51%, and those of the vegetation indices were 89.50%, 92.50% and 98.01% in June (peak period) and 89.85% and 86.16% in September (dry and yellow period), respectively. In this study, different characteristic parameters were selected to establish the discrimination model, and the vegetation index model had the best recognition effect and the highest accuracy. When using a single model for identification in different periods, the overall recognition accuracy of the characteristic parameters was 85.98% in April, 88.48% in June and 89.77% in September; the overall accuracy of the vegetation index model was 90.37% in April, 97.88% in June and 91.53% in September. Characteristic parameters and vegetation indices are also widely used as identification parameters, but the species, quantity and identification accuracy of vegetation indices selected when identifying different plants are different (*[Li et al., 2016](#page-22-6)*). *[Du et al. \(2021\)](#page-21-9)* classified four dominant plant communities in wetlands by calculating vegetation indices through screening feature bands. The highest recognition accuracy for the dominant species, *Phragmites australis,* was 89.17%, while the recognition performance for *Scirpus triqueter* was relatively poor at 72.80%. This study also showed the best recognition effect for the dominant species *Seriphidium transiliense,* with an overall accuracy of 89.54% for the characteristic parameters and a vegetation index of 94.10%.

Identification of the same parameter in different seasons

The accuracy of classification and identification using the same identification parameters also differed between periods. For example, the characteristic parameters used in April (turning green period) yielded an accuracy of 85.51% in April, 92.50% in June (peak period), 89.85% in September (dry yellow period), 89.50% in April (turning green period), 98.01% in June (peak period), and 86.16% in September (dry yellow period). Scholars have also studied the identification and classification of different vegetation or plants at different scales in different periods and determined the optimal period. *[Liu et al. \(2002\)](#page-22-11)* compared the classification results of different band combinations and finally generated a vegetation map, concluding that the images obtained in late spring and summer were better than those obtained in other seasons. *Lu* $\&$ *He (2017)* used spectral indices to classify Canadian grassland plants and found that the overall accuracy in June was the highest at 86%, which is consistent with the result of this study that the optimal period for the identification of *S. transiliense* desert grassland was June (the peak period). Furthermore, *[Tarantino et al. \(2019\)](#page-23-8)* studied the classification results of multiseasonal WorldView-2 (WV-2) satellite images, which were used to map the invasive species *Ailanthus altissima*, and concluded that the combination of remote sensing images of communities and corresponding growing seasons was the best strategy, which was helpful for vegetation interpretation and classification (*[Wei et al., 2017a](#page-23-3)*; *[Wei et al., 2017b](#page-23-4)*), similar to the results of this study. Therefore, in this study, the identification parameters were used to identify

the optimal combination of vegetation indices in June (peak period) for three main plants in the *S. transiliense* desert grassland.

CONCLUSIONS

The identification model established based on the vegetation indices in June had the best discrimination, with the highest accuracy for the test samples of *S. transiliense*, *C. arenarius*, *P. sibirica* and bare land, reaching 100.00%, 95.60%, 100.00% and 95.90%, respectively. Considering the identification period and parameters, the vegetation index model is the best way to identify land features in the *sericite*–*Artemisia* desert grassland in June.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This research was funded by the National Natural Science Foundation of China (31960360) and the Open Project of Key Laboratory in Xinjiang Uygur Autonomous Region (2019D04012). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors: National Natural Science Foundation of China: 31960360. Open Project of Key Laboratory in Xinjiang Uygur Autonomous Region: 2019D04012.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- [Wenhao Liu](#page-0-1) conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- [Wanqiang Han](#page-0-2) conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- [Guili Jin](#page-0-3) conceived and designed the experiments, performed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- [Ke Gong](#page-0-4) performed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- [Jian Ma](#page-0-5) performed the experiments, prepared figures and/or tables, and approved the final draft.

Data Availability

The following information was supplied regarding data availability: The original data is available in the [Supplementary File.](http://dx.doi.org/10.7717/peerj.17663#supplemental-information)

Supplemental Information

Supplemental information for this article can be found online at [http://dx.doi.org/10.7717/](http://dx.doi.org/10.7717/peerj.17663#supplemental-information) [peerj.17663#supplemental-information.](http://dx.doi.org/10.7717/peerj.17663#supplemental-information)

REFERENCES

- **Cao Q, Zhao CJ, Bai BN, Cai J, Chen LY, Wang F, Xu B, Meng XY, Yang GJ. 2023.** Oolong tea cultivars categorization and germination period classification based on multispectral information. *Frontiers in Plant Science* **14**:1251418 [DOI 10.3389/fpls.2023.1251418.](http://dx.doi.org/10.3389/fpls.2023.1251418)
- **Datt B. 2010.** Visible/near infrared reflectance and chlorophyll content in Eucalyptus leaves. *International Journal of Remote Sensing* **20(14)**:2741–2759 [DOI 10.1080/014311699211778.](http://dx.doi.org/10.1080/014311699211778)
- **Dmitriev PA, Kozlovsky BL, Dmitrieva AA, Rajput VD, Minkina TM, Varduni TV. 2022.** Identification of species of the genus *Populus* L. based on the data of hyperspectral crown survey for climate change monitoring. *Environmental Challenges* **9**:100619 [DOI 10.1016/j.envc.2022.100619.](http://dx.doi.org/10.1016/j.envc.2022.100619)
- **Du BJ, Mao D, Wang ZM, Qiu ZQ, Yan HQ, Feng KD, Zhang BZ. 2021.** Mapping wetland plant communities using unmanned aerial vehicle hyperspectral imagery by comparing object/pixel-based classifications combining multiple machine-learning algorithms. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **14**:8249–8258 [DOI 10.1109/JSTARS.2021.3100923.](http://dx.doi.org/10.1109/JSTARS.2021.3100923)
- **Feng SS, Tian B, Hu YC, Wang X, Xue RH. 2016.** The hyper–spectra characteristics of degradation indicator species in thegrassland of Bashang region. *Journal of Arid Land Resources and Environment* **30(03)**:133–139.
- **Frank T, Smith A, Houston B, Lindsay E, Guo XL. 2022.** Differentiation of six grassland/forage types in three Canadian ecoregions based on spectral characteristics. *Remote Sensing* **14(9)**:2121 [DOI 10.3390/rs14092121.](http://dx.doi.org/10.3390/rs14092121)
- **Ge S, Everitt J, Carruthers R, Gong P, Anderson G. 2006.** Hyperspectral characteristics of canopy components and structure for phenological assessment of an invasive weed. *Environmental Monitoring & Assessment* **120(1/3)**:109–126 [DOI 10.1007/s10661-005-9052-1.](http://dx.doi.org/10.1007/s10661-005-9052-1)
- **Han X, Yang W, Song H, Zhang Z, Zuo Y, Duan Z, Zhang X. 2020.** Identification of foxtail millet varieties using leaf surface spectral information. *Sensors and Materials* **32(4)**:1557–1566 [DOI 10.18494/SAM.2020.2718.](http://dx.doi.org/10.18494/SAM.2020.2718)
- **He L, An SZ, Jin GL, Fan YM, Zhang T. 2014.** Analysis on high spectral characteristics of degraded *Seriphidium transiliense* Desert Grassland. *Acta Agrestia Sinica* **22(2)**:271–276 [DOI 10.11733/j.issn.1007-0435.2014.02.010.](http://dx.doi.org/10.11733/j.issn.1007-0435.2014.02.010)
- **Huang LS, Wu ZC, Huang WJ, Ma HQ, Zhao JL. 2019.** Identification of fusarium head blight in winter wheat ears based on fisher's linear discriminant analysis and a support vector machine. *Applied Sciences* **9(18)**:3894 [DOI 10.3390/app9183894.](http://dx.doi.org/10.3390/app9183894)
- **Jin GL, Zhu JZ, Liu HL, Tang SM, Wang CH. 2011.** Study on physiology/ecology adaptation of main plant in degraded *Seriphidium transiliense* Desert Rangeland. *Acta Agrestia Sinica* **19(1)**:26–30 [DOI 10.1093/mp/ssq070.](http://dx.doi.org/10.1093/mp/ssq070)
- **Li JZ, Liu YM, Mo CH, Wang L, Pang GW, Cao MM. 2016.** IKONOS image-based extraction of the distribution area of *Stellera chamaejasme* L. in Qilian County of Qinghai province, China. *Remote Sensing* **8(2)**:148 [DOI 10.3390/rs8020148.](http://dx.doi.org/10.3390/rs8020148)
- **Liang F, Yang W, Xing D. 2021.** Fruit tree identification based on multi-source remote sensing image data—taking pomegranate tree as an example. *IOP Conference Series Earth and Environmental Science* **697(1)** 012006 [DOI 10.1088/1755-1315/697/1/012006.](http://dx.doi.org/10.1088/1755-1315/697/1/012006)
- **Liu QJ, Takamura T, Takeuchi N, Shao G. 2002.** Mapping of boreal vegetation of a temperate mountain in China by multitemporal Landsat TM imagery. *International Journal of Remote Sensing* **23(17)**:3385–3405 [DOI 10.1080/01431160110076171.](http://dx.doi.org/10.1080/01431160110076171)
- **Lu B, He Y. 2017.** Species classification using Unmanned Aerial Vehicle (UAV)-acquired high spatial resolution imagery in a heterogeneous grassland. *ISPRS Journal of Photogrammetry and Remote Sensing* **128(6)**:73–85 [DOI 10.1016/j.isprsjprs.2017.03.011.](http://dx.doi.org/10.1016/j.isprsjprs.2017.03.011)
- **Lyu X, Li XB, Dang DL, Dou HS, Xuan XJ, Liu SY, Li MY, Gong JR. 2020.** A new method for grassland degradation monitoring by vegetation species composition using hyperspectral remote sensing. *Ecological Indicators* **114**:106310 [DOI 10.1016/j.ecolind.2020.106310.](http://dx.doi.org/10.1016/j.ecolind.2020.106310)
- **Khdery G, Yones M. 2021.** Innovative spectral library for identification common wild plants using hyperspectral technology in Northwestern Coast, Egypt. *The Egyptian Journal of Remote Sensing and Space Sciences* **24(3)**:485–491 [DOI 10.1016/j.ejrs.2020.08.002.](http://dx.doi.org/10.1016/j.ejrs.2020.08.002)
- **Melville B, Lucieer A, Aryal J. 2018.** Assessing the impact of spectral resolution on classification of lowland native grassland communities based on field spectroscopy in Tasmania, Australia. *Remote Sensing* **10(2)**:308:2–19 [DOI 10.3390/rs10020308.](http://dx.doi.org/10.3390/rs10020308)
- **Ouyang ZT, Gao Y, Xie X, Guo HQ, Zhang TT, Zhao B. 2013.** Spectral discrimination of the invasive plant *Spartina alterniflora* at multiple phenological stages in a saltmarsh wetland. *PLOS ONE* **8(6)**:e67315 [DOI 10.1371/journal.pone.0067315.](http://dx.doi.org/10.1371/journal.pone.0067315)
- **Qiao X, Liu XH, Wang FK, Sun ZY, Yang L, Pu XJ, Huang YQ, Liu SY, Qian WQ. 2022.** A method of invasive alien plant identification based on hyperspectral images. *Agronomy* **12(11)**:2825 [DOI 10.3390/agronomy12112825.](http://dx.doi.org/10.3390/agronomy12112825)
- **Reinermann S, Asam S, Kuenzer C. 2020.** Remote sensing of grassland production and management—a review. *Remote Sensing* **12(12)**:1949 [DOI 10.3390/rs12121949.](http://dx.doi.org/10.3390/rs12121949)
- **Singh KK, Chen YH, Smart L, Gray J, Meentemeyer RK. 2018.** Intra-annual phenology for detecting understory plant invasion in urban forests. *ISPRS Journal of Photogrammetry and Remote Sensing* **142(AUG.)**:151–161 [DOI 10.1016/j.isprsjprs.2018.05.023.](http://dx.doi.org/10.1016/j.isprsjprs.2018.05.023)
- **Soubry I, Guo XL. 2021.** Identification of the optimal season and spectral regions for shrub cover estimation in grasslands. *Sensors* **21(9)**:3098 [DOI 10.3390/s21093098.](http://dx.doi.org/10.3390/s21093098)
- **Tarantino C, Casella F, Adamo M, Lucas R, Beierkuhnlein C, Blonda P. 2019.** Ailanthus altissima mapping from multi-temporal very high resolution satellite images. *ISPRS Journal of Photogrammetry and Remote Sensing* **147**:90–103 [DOI 10.1016/j.isprsjprs.2018.11.013.](http://dx.doi.org/10.1016/j.isprsjprs.2018.11.013)
- **Wang X, Wang L, Tian J, Shi C. 2021.** Object-based spectral-phenological features for mapping invasive *Spartina alterniflora*. *International Journal of Applied Earth Observation and Geoinformation* **101(2)**:102349 [DOI 10.1016/j.jag.2021.102349.](http://dx.doi.org/10.1016/j.jag.2021.102349)
- **Wei XH, Jin GL, Fan YM, An SZ, Zhu XW, Wang HN. 2017a.** Community coverage estimation of degraded *Seriphidium transiliense* desert Grassland based on hyperspectral remote sensing. *Chinese Journal of Grassland* **39(6)**:33–39 + 46 [DOI 10.16742/j.zgcdxb.2017-06-06.](http://dx.doi.org/10.16742/j.zgcdxb.2017-06-06)
- **Wei XH, Jin GL, Fan YM, An SZ, Zhu XW, Wang HN. 2017b.** Estimation of biomass based on HJ-HSI data in *Seriphidium transiliense* desert grassland. *Pratacultural Science* **34(12)**:2504–2511 [DOI 10.11829/i.issn.1001-0629.2017-0059.](http://dx.doi.org/10.11829/i.issn.1001-0629.2017-0059)
- **Weisberg PJ, Dilts TE, Greenberg JA, Johnson KN, Pai H, Sladek C, Kratt C, Tyler SW, Ready A. 2021.** Phenology-based classification of invasive annual grasses to the species level. *Remote Sensing of Environment* **263**:112568 [DOI 10.1016/j.rse.2021.112568.](http://dx.doi.org/10.1016/j.rse.2021.112568)
- **Yu H, Bo K, Wang GX, Sun H. 2018.** Hyperspectral database prediction of ecological characteristics for grass species of alpine grasslands. *The Rangeland Journal* **40(1)**:19–29 [DOI 10.1071/RJ17084.](http://dx.doi.org/10.1071/RJ17084)
- **Zhang C, Atkinson PM, George C, Wen Z, Diazgranados M, Gerard F. 2020.** Identifying and mapping individual plants in a highly diverse high-elevation ecosystem using UAV imagery and deep learning. *ISPRS Journal of Photogrammetry and Remote Sensing* **169**:280–291 [DOI 10.1016/j.isprsjprs.2020.09.025.](http://dx.doi.org/10.1016/j.isprsjprs.2020.09.025)
- **Zhang Y, Qin QM, Ren HZ, Sun YH, Li MZ, Zhang TY, Ren SL. 2018.** Optimal hyperspectral characteristics determination for winter wheat yield prediction. *Remote Sensing* **10(12)**:2015 [DOI 10.3390/rs10122015.](http://dx.doi.org/10.3390/rs10122015)
- **Zhou B, Li H, Xu F. 2022.** Analysis and discrimination of hyperspectral characteristics of typical vegetation leaves in a rare earth reclamation mining area. *Ecological Engineering* **174**:106465 [DOI 10.1016/j.ecoleng.2021.106465.](http://dx.doi.org/10.1016/j.ecoleng.2021.106465)