Disturbed Boundaries Extraction in Coal-grain Overlap Areas With High Groundwater Levels Using UAV-based Visible and Multispectral Imagery

Yunqi Guo
China University of Mining and Technology  https://orcid.org/0000-0001-8332-4016

Yanling Zhao (✉ zhaoyl7677@163.com)
China University of Mining Technology: China University of Mining and Technology

Haoyue Yan
China University of Mining Technology: China University of Mining and Technology

Research Article

Keywords: High groundwater level, UAV, Multispectral imagery, Elevation threshold segmentation, OTSU segmentation, Vegetation index, Out-of-production boundary, Reduced-production boundary

Posted Date: November 11th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-924902/v1

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Version of Record: A version of this preprint was published at Environmental Science and Pollution Research on April 4th, 2022. See the published version at https://doi.org/10.1007/s11356-022-19966-6.
Disturbed boundaries extraction in coal-grain overlap areas with high groundwater levels using UAV-based visible and multispectral imagery

Yunqi Guo · Yanling Zhao · Haoyue Yan

Abstract
Coal-grain overlap areas (CGOA) with high groundwater levels are vulnerable to subsidence and water logging during a series of mining activities, which have adverse impacts on crop yields. Such damage requires full reports of disturbed boundaries for the agricultural reimbursement and ongoing reclamation. Since direct measurements are difficult in such a case because of vast, unreachable areas, so it is necessary to be able to identify out-of-production boundary (OB) and reduced-production boundary (RB) in the corresponding region. In this study, OB was extracted by setting thresholds through characteristics of the cultivated land elevation based on UAV-generated digital surface model (DSM) and digital orthophoto map (DOM). Meanwhile, aboveground biomass (AGB), the soil plant analysis development (SPAD) value of chlorophyll contents, and leaf area index (LAI), were used to select the appropriate vegetation indexes (VI) to perform a reduced-production map (RM) based on power regression (PR), exponential regression (ER), multiple linear regression (MR) and random forest (RF) algorithms. Finally, an improved OTSU segmentation algorithm was applied to extract mild RB and severe RB. The results show elevation threshold segmentation method and the improved OTSU segmentation method can accurately recognize and extract disturbed boundaries, which are consistent with the tonal difference after crop damage in the image. This study provides reference methods and theoretical supports for disturbed boundaries determination in CGOA with high groundwater levels for further agricultural compensation and reclamation processes.

Keywords High groundwater level · UAV · Multispectral imagery · Elevation threshold segmentation · OTSU segmentation · Vegetation index · Out-of-production boundary · Reduced-production boundary

Corresponding author: Yanling Zhao (zhaoyl7677@163.com)
1 Institute of Land Reclamation and Ecological Restoration, China University of Mining and Technology (Beijing), Beijing 100083, China
Introduction

China is one of the largest coal producers and processors in the world. As an indispensable pillar industry of China, coal-mining has occupied a major stake in the field of energy and economy. Nevertheless, the large-scale coal-mining evitability brings significant landscape damages and detrimental environmental pollution, which has become a looming threat to the wellbeing of all mankind (Bell et al. 2000; Bian et al. 2009).

In coal mines with high groundwater levels, the groundwater tends to rise upwards easily after subsidence, thus forming a unique water-land compound ecological environment (Ren et al. 2020). Such outcome is most obvious for which mining areas overlap cultivated lands. The Coal-grain overlap areas (CGOA) refers to a region with underground mining and further processing, meanwhile agricultural production above-ground is still ongoing (Hu et al. 2014). In this circumstance, crops have been soaked in stagnant water for a long time, resulting in dramatic yield losses. According to regulations of agricultural reimbursement and reclamation, the scope of disturbed boundaries shall affect the final decision of crops losses. However, due to the lack of policies for reporting the loss of cultivated lands in CGOA with high groundwater levels, the damaged degree is difficult to quantitatively describe and there is still no consensus on criteria of determining disturbed boundaries and corresponding compensation levels. Therefore, the precise extraction of disturbed boundaries in subsidence areas is a prerequisite as well as an important basis for realizing reimbursement and reclamation.

Currently, China does not have a unified standard for defining boundaries of damaged cultivated lands. It is ambiguous to apply the most frequently used 10-mm subsidence contour as disturbed boundaries, because most of cultivated lands are only slightly affected. In such case, this blurred definition will lead to unnecessary increments in project budgets, thereby giving extra burdens to national economic expenditure and wasting valuable resources. In order to solve this uncertainty, many scholars have carried out studies based on geological information to stimulate surface subsidence (Liu et al. 2021; Yuan et al. 2021; Komnitsas et al. 2010). These methods greatly improve the accuracy of boundary extraction results without a doubt. However, those studies, which mostly rely on tremendous field operations, have been vigorously challenged. Researchers need to carry out toilsome surveys outdoors, which consume a lot of manpower, material resources, and time. On the other hand, it cannot handle the problem of temporal and spatial heterogeneity in the study area. In this case, it appears that the use of remote sensing techniques may be an excellent alternative.

With the development of remote sensing technologies, many predecessors have utilized different kinds of satellites to extract farmland boundaries. Lu et al. (2021) created a method that considered the characteristics of spectrum, shape and texture to calculate the optimal segmentation scale of cultivated lands using the Quickbird satellite. Xue et al. (2021) extracted the crop boundary in GF-2 remote sensing images based on an improved watershed segmentation algorithm. Watkins et al. (2019) modified the OBIA method to realize automatically identify and delineate agricultural fields from Sentinel-2 imagery. However, satellites have shortcomings, such as high cost, long operation period, and severe restriction of meteorological conditions (Zheng et al. 2018). In recent years, the unmanned aerial vehicle (UAV) as an emerging technology has a series of advantages such as relatively flexible, low cost, short data acquisition cycle, and
low risk (Li and Hu, 2019). At the same time, the UAV technology in mining areas can directly serve the
dynamic monitoring of land use and subsidence, geological environment and disaster evaluation, mineral
resources survey, acquisition of three-dimensional images of the surface, landscape planning and ecological
reconstruction activities (Xiao et al. 2017). Therefore, the use of UAV technology to obtain the boundary in
the coal mining subsidence area not only enhances the timeliness of extraction, but also reduces the workload
for field workers.

Due to its high efficiency and strong stability, vegetation indexes (VIs) are the most commonly used
variable parameter to achieve the estimation of crop growth in remote sensing. Meanwhile, there are studies
exhibits that the value of above ground biomass (AGB), soil and plant analyzer development (SPAD), and
leaf area index (LAI) have a strong relationship with the growth of crops as well as the yield, which makes
them become important tools for agricultural monitoring (Villoslada et al. 2021; Zhao et al. 2021). Nowadays,
the images acquired by cutting-edge hyperspectral data contain richer spatial, radiant and spectral information,
and the spectral resolution has been greatly improved, but the cost is high (Chen et al. 2019). Furthermore,
as the number of bands increase, the amount of data increase exponentially, which leads to information
redundancy and difficulty in post-processing. Therefore, in this study, the relatively low price and simple-
processed multispectral data was used to construct the corresponding VIs to complete crop growth and yield
analysis.

At present, UAV remote sensing technology based on multispectral data has been widely used in land
surveys, land damage evaluation, and crop production evaluation. UAVs have already played an important
role in the China third land survey (Li, 2021). Ren et al. (2020) estimated maize biomass using UAV multi-
spectral remote sensing data to assess the rapid, effective and non-destructive method of underground coal
mining damage to the land. Jełowicki et al. (2020) used the multispectral camera to evaluate the frost-
damaged area, in addition, the OBIA method was implemented to remove the impact of roads. In order to
assess the impact of climate disasters on crops and reasonably determine agricultural insurance, Vlachopoulos
et al. (2020) recommended pixel-based and object-based random forest algorithms to extract boundaries using
multispectral data obtained by UAV. Although many scholars have conducted a lot of related researches on
the problems of land damage, there are few studies directly on extraction of boundaries with various damage
levels. Threshold segmentation is a commonly used when it comes to image processing (Singh et al. 2015;
Qiu and Wang et al. 2010; Xie et al. 2010), the principle of this simple method is to select an appropriate
threshold, determining whether each pixel meets the threshold requirements, then dividing each pixel into
respective intervals that meet the requirements. Since the elevation difference of affected crops within
subsidence areas is obvious, out-of-production boundary (OB) extraction can be studied according to this
clue theoretically. The OTSU segmentation has been considered to be another one of robust and popular
algorithms. It is convenient to calculate and is not affected by image brightness and contrast, so it has been
widely used in image processing (Li, 2019). However, the traditional OTSU algorithm is to calculate the gray
value of the image, and is generally applied to discrete data, hence the traditional OTSU algorithm cannot be
directly used for threshold segmentation, which needs to be improved. Therefore, in this study, an improved OTSU algorithm was introduced to extract reduced-production boundary (RB).

Based on the above analysis, this study took a representative CGOA with high groundwater levels, the Dongtan coal mine in Shandong Province, using a DJI M100 UAV, equipped with a Parrot Sequoia multispectral camera and a ZenmuseX3 camera for the aerial photography. Moreover, field samplings were also achieved during the study period. The purpose was to: (1) Construct digital orthophoto map (DOM) and digital surface model (DSM) of the study area, analyze the distribution of the surface elevation, and perform elevation threshold segmentation to extract OB, (2) Use four regression methods, power regression (PR), exponential regression (ER), multiple linear regression (MR) and random forest (RF) to estimate the value of AGB, SPAD and LAI. After optimizing the inversion model, the reduced-production map (RM) was constructed, and finally RBs were extracted by performing an improved OTSU algorithm. In a word, this study is useful for subsequent reclamation work and it also provides solid technical foundation for compensation determination research of damaged crops.

Materials and methodology

Study area

The study was conducted at Dongtan coal mine in Jining city (116° 50′ 49″–116° 56′ 56″, 35° 24′ 11″–35° 31′ 25″). A typical CGOA with high groundwater levels, the summer cornfield of A1 workface was selected as the study area (Fig.1). Ground elevation of the targeted region is about +42.66~+54.58 m, gradually decreasing from northeast to southwest with relatively gentle inclination. The water table is quite shallow, about 2 m throughout the year. It has a warm temperate monsoon climate with four distinct seasons. The average annual temperature within the area is 13.3 ℃~14.1 ℃, and the average frost-free period is 199 days, the average annual precipitation is around 597-820 mm.

The mining of the working face started in August 2014. During the research period, the average mining depth was 550m, and the maximum subsidence depth had reached 7 m. Because of high groundwater levels, the sinking center had flooded a large amount of high-quality cultivated lands, forming a waterlogged basin. The surface of subsidence within the study area had been stabilized during field campaigns. Meanwhile, in order to complete the field survey and sampling works, the study area was ensured that it was not repaired by manual restoration beforehand.
Fig. 1 Location of study area and distribution of sample lines, (a) sample lines within the study area, (b) location of Jining city, and (c) the map of China

Data collection

UAV remote sensing data collection

The data in the study were acquired by a DJI M100 UAV equipped with a Parrot Sequoia multispectral camera. The four bands of the Parrot Sequoia multispectral camera are green (530–570 nm), red (640–680 nm), red edge (730–740 nm), and near infrared (770–810 nm), each with 1.2 million pixels (1280×960). In addition, the UAV was also equipped with a ZenmuseX3 digital camera with 16 million pixels, which realized the simultaneous collection of visible light data and terrestrial spectrum information. The flight was performed on 12 August 2017, at the altitude of 110 m above the ground, with speed of 9 m/s, and spatial resolution was 13 cm. The area taken during the flight mission was about 1.1 km², and a total of 4980 photos were collected, completely covering the entire study area.

The procession of images was performed in Pix4DMapper software (Pix4D SA, Switzerland). One excellent feature of this software is that the image stitching, geometric and radiometric corrections are all
automatically completed, meanwhile, DOM and DSM can be acquired quickly without extra human knowledge.

Field data collection

Field samplings were carried out on 13 August 2017. Based on the mining subsidence theory and the probability integration method, combined with a variety of surface movement parameters (sink coefficient, horizontal movement coefficient, etc.) and coal seam information (coal seam thickness, inclination, etc.), the mining subsidence prediction software (MSPS) was applied to predict mining subsidence. Meanwhile, according to the traditional definition, the 10-mm subsidence contour was used as the basis of disturbed boundaries. Combine the two approaches above, the subsidence contour of the working face corresponding to the study area was achieved. Three sampling lines, namely L1 (550 m), L2 (540 m), and L3 (620 m) were laid out along the strike, dip, and angular bisector directions. The spacing of the sample points increased according to the decrease of the sinking value, which was 5 m, 10 m, 20 m, 30 m, and 60 m in sequence, generating a total of 54 sampling points (Fig.1).

For each sampling point, a 1 m×1 m standard plot was designed to collect AGB, SPAD and LAI values.

1. AGB measurement

The AGB was determined by “the field harvest method” (Liu et al. 2020). Randomly select two maize (Maize 1 and Maize 2) in the plot as the test samples. In the laboratory, the two samples were stored for 10 h at 105 °C and dried at 80 °C for 24 h until reaching their constant weights. The final AGB was calculated based on the equation below.

\[
AGB = \frac{(AGB_1 + AGB_2) \times N}{2S}
\]  

where \( N \) represents the number of maize in each sample plot, \( S \) denotes the area of the sample plot (m\(^2\)), \( AGB_1 \) and \( AGB_2 \) are the maize AGB value of sample 1 and sample 2 respectively.

2. Chlorophyll contents measurement

This parameter was measured with SPAD-502 using “five-point method”. Select five uniformly growing maize, measure the leaves along the tip, middle, and base of each plant from the top leaf to the secondary leaf twice. Finally, the measured values of the five plants in the plot were averaged and used as the SPAD value for the sample point.

3. LAI measurement

For the determination of LAI, the most traditional “length-width coefficient method” was selected (Kross et al. 2015). The length of all leaves and the maximum width of the corresponding leaves of five plants were measured using a steel ruler. The leaf area correction coefficient of the north China plain 0.7017 was used (Wang, et al., 2015). The formula is as follows:
Where $k$ is the leaf area correction coefficient, $L$ is the leaf length, $B$ is the width of leaf, $N$ is the total number of leaves of the plant $j$, $m$ stands for the number of the selected plants, $\rho$ is the planting density.

**Fig. 2** The scheme of the study performed

**Elevation threshold segmentation**
Flooded crops and unaffected crops show significant elevation differences. Derived from this idea, the desired
OB could be achieved. Threshold segmentation is to achieve the function of segmenting the background and
the target by determining a specific threshold, it is a simple type of segmentation algorithms (Kaur et al.
2016). The threshold segmentation function is defined as follows:

$$ g(i, j) = \begin{cases} 1, & f(i, j) > T \\ 0, & f(i, j) < T \end{cases} $$ (3)

Where $T$ represents the predetermined threshold, the pixel of the set target is $g(i, j) = 1$, the others
would be $g(i, j) = 0$. Thus, the pixel with the larger given feature belongs to the foreground group while the
one with the smaller value is thrown into the background group.

For the OB identification task, ArcGIS 10.2 was used to make multiple profile lines on the DSM of the
study area, then discontinuity points of each profile line were recorded, using the average value of these
points as the threshold for elevation segmentation. The data with elevation higher than the threshold was
treated as non-out-of-production area, while the data below the threshold was classified as the out-of-
production area.

**Optimization of vegetation index**

There is a strong correlation between the characteristic bands of spectral reflectance and vegetation
photosynthetic pigments. Therefore, the VIs based on this relationship are a good tool to help test the health
of plants and solve the problem of defining the scope of damaged cultivated lands (Jelowicki et al. 2020;
Kross et al. 2015). Some commonly used VIs in existing research, which met the requirements of this study
were selected, and on this basis, because the red-edge band is more sensitive to plants chlorophyll, the red or
green band was replaced by the red-edge band to improve and expand the existing VIs during study.

Through correlation analysis of AGB and VIs, the red-edge microwave vegetation index ($\text{MVI}_{\text{redge}}$), the
normalized difference vegetation index ($\text{NDVI}$) and the red-edge modified simple ratio index ($\text{MSR}_{\text{redge}}$)
were selected as the optimal VIs for biomass estimation. For the study of SPAD, the red-edge normalized
difference vegetation index ($\text{NDVI}_{\text{edge}}$), as well as the green normalized difference vegetation index ($\text{GNDVI}$)
and the red-edge chlorophyll index ($\text{CI}_{\text{redge}}$) were chosen. Meanwhile, the red-edge modified simple ratio
index ($\text{MSR}_{\text{redge}}$), the green normalized difference vegetation index ($\text{GNDVI}$) and the green chlorophyll index
($\text{CI}_{\text{green}}$) were put into the estimation of LAI.

**Table 1** Selected VIs of $\rho_{\text{NIR}}$, $\rho_{\text{R}}$, $\rho_{\text{G}}$, and $\rho_{\text{redge}}$ represent the $\rho_{\text{Reflectance}}$ (%) of the near-infrared (790
nm), red (660 nm), green (550 nm), and red-edge (735 nm) bands respectively

| Vegetation Index | Expression | Reference |
|------------------|------------|-----------|
| NDVI             | $\text{NDVI} = (\rho_{\text{NIR}} * \rho_{\text{R}}) / (\rho_{\text{NIR}} + \rho_{\text{R}})$ | Rouse (1974) |
| $\text{CI}_{\text{redge}}$ | $\text{CI}_{\text{redge}} = (\rho_{\text{NIR}} / \rho_{\text{redge}}) - 1$ | Gitelson et al. (2003) |
| $\text{MVI}_{\text{redge}}$ | $\text{MVI}_{\text{redge}} = [(\rho_{\text{NIR}} * \rho_{\text{redge}}) / (\rho_{\text{NIR}} + \rho_{\text{redge}}) + 0.5]^{1/2}$ | Shi et al. (2008) |
Modeling and evaluation

In this study, four popular regression models, power regression (PR), exponential regression (ER), multiple linear regression (MR) and random forest (RF) were applied for the AGB, SPAD, and LAI analysis. PR and ER were utilized as pioneers for testing the responding of each VI. Furthermore, the leave-one-out cross-validation (LOOCV) method was utilized for the model training and validation (Maimaitijiang et al. 2019). The root mean squared error (RMSE) and the coefficient of determination ($R^2$) were selected as the support for the model accuracy verification.

The $R^2$ represents the degree of fitting between the simulated value and the measured value. The closer the value of $R^2$ approaches 1, the better the degree of fitting. The RMSE can reflect the degree of deviation between the simulated value and the measured value. The smaller the value of RMSE, the higher the accuracy of the representative.

In this study, above mentioned three models combined with DSM were used to perform equal-weight spatial calculations. As a result, the reduced-production map (RM) was generated by overlaying characteristics of these four sources. The value of RM, the reduced-production coefficient, was an imperative criterion to gauge the crops growth. The lower the coefficient value, the more serious the corn yield reduction degree, and the higher the coefficient value, the weaker the yield reduction degree.

OTSU threshold segmentation

OTSU algorithm was proposed by the Japanese scholar Nobuyuki Otsu in 1979. It is an efficient algorithm for image binarization. This algorithm can maximize the between-class variance between the foreground image and the background image (Otsu N, 1979). Among all the segmentation methods, Otsu method is one of the most successful methods for image thresholding because of its simple calculation (Vala et al. 2013).

The basic idea is as follows:

The total number of pixels of the given image is $N$ with gray levels $[0, L-1]$. The number of pixels $n_i$ corresponding to the gray level is $i$, its probability can be expressed as:

$$ p_i = \frac{n_i}{N} \quad (4) $$

There is a threshold $T$, and the gray values are divided into two categories $C_0, C_1$ according to $T$. The average value of the distribution probability based on the gray levels of the entire image is expressed as:

$$ \text{MSR}_{\text{redge}} = \left[ \frac{\rho_{\text{NIR}}}{\rho_{\text{redge}}} \right] - 1 \left[ \frac{\rho_{\text{NIR}} + \rho_{\text{redge}}}{} \right] $$

Roujean et al. (1995)

$$ \text{CI}_{\text{green}} = \frac{\rho_{\text{NIR}}}{\rho_{\text{G}} - 1} $$

Gitelson et al. (2003)

$$ \text{NDVI}_{\text{redge}} = \left( \frac{\rho_{\text{NIR}}}{\rho_{\text{redge}}} \right) \left( \frac{\rho_{\text{NIR}} + \rho_{\text{redge}}}{} \right) $$

Rouse (1974)

$$ \text{GNDVI} = \left( \frac{\rho_{\text{NIR}}}{\rho_{\text{G}}} \right) \left( \frac{\rho_{\text{NIR}} + \rho_{\text{G}}}{} \right) $$

Gitelson et al. (1996)
The average values of $C_0, C_1$ are:

$$
\bar{\mu} = \frac{\sum_{i=0}^{T-1} \bar{ip}_i}{\bar{w}_0}
$$

(5)

where

$$
\bar{w}_0 = \sum_{i=0}^{T} p_i, \quad \bar{w}_1 = \sum_{t=1}^{T+1} p_i = 1 - \bar{w}_0
$$

(6)

(7)

from (5)-(7), it can be deducted:

$$
\bar{\mu} = \mu_0 \bar{w}_0 + \mu_1 \bar{w}_1
$$

(8)

it can be concluded that the between-class variance of the image is:

$$
\sigma_B^2 = w_0(\mu_0 - \bar{\mu})^2 + w_1(\mu_1 - \bar{\mu})^2
$$

(9)

combine (7)-(9):

$$
\sigma_B^2 = w_0 \mu_0^2 + w_1 \mu_1^2 - \bar{\mu}^2
$$

(10)

Continue to substitute (9) and (10):

$$
\sigma_B^2 = w_0 \mu_0^2 + w_1 \mu_1^2 - (\mu_0 \bar{w}_0 + \mu_1 \bar{w}_1)^2
$$

(11)

$T$ is sequentially selected in the range of $[0, L-1]$, and the value of $T$ that maximizes the between-class variance $\sigma_B^2$ is the optimal threshold of the OTSU algorithm.

However, the image of production reduction in the study area was raster data, as well as a double-type discrete data. In this case, the traditional OTSU algorithm could not be directly used for threshold segmentation, it needed to be improved:

(1) Endow the minimum value and maximum value of the given image with $MinData$ and $MaxData$ respectively, and select the initial segmentation number $breaknum_0$, and the corresponding unit interval $Step_0$ was obtained.

$$
Step_0 = \frac{MaxData - MinData}{breaknum_0}
$$

(12)

(2) On the basis of the traditional OTSU algorithm, take the value between $MinData$ and $MaxData$ using $Step_0$, find the maximum inter-cluster variance $\sigma_0$ under $breaknum_0$ and the corresponding value of the reduction-production degree coefficient $h_0$, and select a new $NN_l = 2 \times N_0$ at the same time.

(3) Recalculate (12), repeat the step (6) to generate a new $\sigma_l$ and corresponding $h_l$, when $h_l - h_{l-1}$ is less than the threshold, the iteration stops.
Result

OB extraction

The values of discontinuity points of four selected profile lines a, b, c, d (Fig. 4a) were 46.41 m, 46.52 m, 46.98 m, 46.01 m (Fig. 3). The calculated average 46.48 m was perceived as the threshold of the elevation segmentation. And then in the wake of some optimized operations, mathematical morphology, boundary tracking, manual adjustments, OB was successfully extracted (Fig. 4).
Common and popular regression algorithms, power regression (PR), exponential regression (ER), were applied to test and select optimal VIs according to the value of $R^2$. Meanwhile, multiple linear regression (MR), random forest (RF), were tested to evaluate prediction results of AGB, SPAD and LAI. The single VI for stimulating each parameter result was illustrated in Fig. 5a. Then LOOCV was devoted into assessing those models by further training and validation processes. From Table 2, it was concluded that RF was proved to be more effective than other competitors for estimating AGB (Using MVI$_{redge}$, NDVI, MSR$_{redge}$ with $R^2=0.83$, RMSE=0.114 kg·m$^{-2}$) and SPAD (Using NDVI$_{redge}$, GNDVI, CI$_{redge}$ with $R^2=0.84$, RMSE=0.152 SPAD), which showed higher $R^2$ and lower RMSE during the test. In addition, MR (Using MSR$_{redge}$, GNDVI,
CI green with $R^2=0.88$, RMSE=1.070) exhibited more strong agreement with LAI estimation, although its RMSE was higher than RF, the relatively lower $R^2$ emphasized its exemplary performance. The comparison of best single VI estimation (Fig. 5a) and combined VIs estimation were showed below (Table 2 and Fig. 5b).

Fig.5 Relationship between UAS-derived and field measured parameters, (a-c) Results of selected VIs, (d-f) modeling results of AGB, SPAD, LAI
### Table 2: Accuracy analysis of inversion model (For PR and ER, only shows the highest result in each category)

| Parameter | Vegetation Index | Model | Modeling set | | | | Cross-validation set | | | |
|-----------|------------------|-------|--------------|---|---|-----------------|---|---|---|
|           |                  |       | Size         | R² | RMSE | Size | R² | RMSE | |
| AGB       | MVIs_edge        | Exponential function | y = 0.0372e^{-0.9x} | 26 | 0.81 | 0.15 | 36 | 0.78 | 0.2 |
|           | MVI             | Multiple linear regression | y = 6.3584 + 7.8281x1 + 0.4408x2 - 0.5715x3 | 26 | 0.81 | 0.15 | 36 | 0.77 | 0.19 |
|           | NDVI             | Random forest | / | 10 | 0.83 | 0.11 | 10 | 0.72 | 0.18 |
|           | MSR_edge         | Power function | y = 5.4226x^{0.88} | 25 | 0.71 | 6.62 | 35 | 0.7 | 6.92 |
| SPAD      | NDVI_edge        | Multiple regression | y = -26.3245 + 519.7154x1 - 6.6237x2 - 87.3451x3 | 25 | 0.77 | 5.96 | 35 | 0.75 | 6.17 |
|           | GNDVI            | Random forest | / | 10 | 0.83 | 0.15 | 10 | 0.89 | 0.01 |
|           | CI_edge          | Power function | y = 148.77x^{0.05} | 26 | 0.87 | 1.22 | 36 | 0.85 | 1.18 |
| LAI       | MSR_edge         | Multiple regression | y = 12.9006 + 14.9724x1 - 32.5171x2 + 1.3229x3 | 26 | 0.88 | 1.07 | 36 | 0.84 | 1.11 |
|           | GNDVI            | Random forest | / | 10 | 0.84 | 0.15 | 10 | 0.81 | 0.14 |
The estimation results showed that the values of AGB, SPAD, LAI in the study ranged between 0.01~1.85 kg/m$^2$, 6.25~63.57 SPAD and 0.15~14.41 respectively (Fig. 5), which were consistent with the measured results collected from field samples.

RM demonstrated that the reduced-production coefficient of corns fluctuated from 0.18 to 2.79. It could be clearly observed that the coefficient values from the center of the subsidence basin towards the outside were gradually increasing (Fig. 6d). Moreover, compared with the modeling results of a single parameter, the spatial difference in the distribution of multiple parameters were more significant (Fig. 6).

![Fig. 6](image)

Fig. 6 results of modeling (From left to right), (a) the AGB inversion map, (b) the SPAD map, (c) the LAI map, (d) RM

The improved OTSU algorithm was used to perform threshold segmentations on RM. Then severe RB and mild RB in the study area were obtained only after the performing of mathematical morphology, boundary tracking methods. The two different levels of RBs and OB were put together on the DOM, and total results of OB and RBs extraction were shown in Fig. 7e.
**Fig. 7** RBs extraction process (From left to right), (a) results of first OTSU segmentation, (b) result of second OTSU segmentation, (c) results of post-optimized mild RB, (d) results of post-optimized severe RB, (e) final RBs and OB on DOM

**Discuss**

**Performance of the elevation threshold segmentation for OB**

The study area is a typical CGOA with high groundwater levels, and OB is located between the flooded tidal flat and the unflooded land. In this case, the flooded crops and the unaffected crops show significant elevation differences in DSM. As a result, using this special characteristic based on the simple segmentation algorithm, OB can be accurately extracted. The effect of extraction can be directly tested from Fig. 6. The outer edge of the red line is obviously darker than that of the normal cultivated land. In this scenario, the dark area represents the direct result of abnormal growth of crops. The OTSU segmentation algorithm can also be used to extract OB, Zhao et al. (2020) provided a valid reference by implementing OTSU segmentation based on RGB images. However, this trait on the elevation is obvious, and good results can be achieved by simple threshold segmentation, the most important component of the extraction results should be the accuracy of the DSM. In addition to utilize RGB images, this study also added multi-spectral bands to construct DSM, so the accuracy of the extraction result was greatly improved.

**Performance comparison of models for AGB, SPAD, LAI estimation**

This study used multispectral and visible light images collected by a UAV, combining with field-measured data to extract boundaries of disturbed cultivated lands in different levels. Three important parameters that
are closely related to vegetation growth (AGB, SPAD and LAI) were selected to establish a RM before final extracting results. Commonly used regression models (PR, ER, MR, RF) were employed during this step, outperformance of RF in the estimation of AGB and SPAD compared to MR and PR (Table 2) can be proved by many extant studies. Multiple linear regression, support vector machine, artificial neural network, and random forest were used to create a suitable model for AGB estimation, and the random forest model gave the most balanced results, with low error and a high ratio of the explained variance for both the training set and the test set during the study (Han et al. 2019). Random forest, partial least square method, BP neural network and support vector machine were applied to estimate the chlorophyll content of apple leaves, and random forest was regarded to be the best modeling choice (Feng et al. 2018). Additionally, MR showed a promising result as an estimator for LAI, which is supported by LAI estimation of a natural forest study (Pu 2012) and one winter wheat growth study (Han et al. 2021).

Despite ideal results have been showed, some limitations need to be further considered. Using only one stage data, summer corns, as the final overall crop growth inversion result has certain limitations, since many uncertain variables throughout the year can interfere with crop growth. Moreover, the study cannot eliminate the influences of differences in internal growth of plants (Siebers et al. 2018).

Performance of the improved OTSU segmentation algorithm for RBs

In the RM, the spatial distribution of coefficients all showed a certain law (Fig. 5), that was, the closer to the out-of-production area, the lower value it appeared, and the increase gradually ranged from the inner area of the subsidence basin to the outside along the direction of the three sampling lines. Based on this theory, RM could be obtained by exerting the improved OTSU segmentation twice. Although many scholars have invented miscellaneous improved OTSU methods to date (Sun et al. 2016; Bo Lei et al. 2015; Rajinikanth et al. 2015), the application of extracting cultivated land boundaries is rarely traceable.

After the completion of the field sampling, Shandong Province was affected by a severe typhoon, Tiange (Which started in August 24, 2017), in the following week, making gauge the ground-based yield become an impossible mission. Using actual yield data as a standard to test the accuracy of boundary extraction should be further investigated. Moreover, regarding the problem that there was lesser richness of vegetation on the road, which led to a relatively low reduced-production coefficient within a certain scope of the road and its surroundings, so the accuracy of the directly extracted boundary could be compromised, follow-up research will consider how to automatically eliminate these influencing factors.

Conclusion

This study presents an efficient approach for utilizing low-cost, time-saving, high resolution RGB and multispectral images obtained from UAV to accurately extract disturbed boundaries in OGCA with high groundwater levels. The demonstrated methodology includes analysis of elevation from DSM to exert a simple threshold segmentation, selecting optimal models for AGB, SPAD, LAI estimations, and approaches for using an improved OTSU method, which can be perceived as an effective alternative for disturbed boundaries extraction during the refurbishment and reclamation process.
The non-crop area and the crop area within the scope of subsidence were obviously different in elevation. The elevation threshold could be set to segment DSM to achieve the purpose of extracting OB accurately in the study area.

Furthermore, during the process of establishing RM, three parameters (AGB, SPAD, LAI) highly related to crops growth were used. The model of RF with NDVI_{redge}, GNDVI and CI_{redge} index as variables proved to be the best estimator for AGB. The model of RF with NDVI_{redge}, GNDVI and CI_{green} index outperformed other combinations within SPAD estimation. In the inversion test of LAI, the MR model with MSR_{redge}, GNDVI and CI_{green} index tended to be a superior choice.

Based on RM, two different RBs were extracted successfully by implementing an improved OTSU segmentation. This study shall give an innovative insight into using simple physiological traits of corps to identify its disturbed boundaries swiftly and accurately as well as relief the heavy workload during the reclamation and compensation survey.

**Availability of data and materials** The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

**Funding** This work was supported by the Joint Funds of the National Nature Science Foundation of China (No. 41401609).

**Declarations**

**Ethics approval and consent to participate** Not applicable.

**Consent for publication** Not applicable.

**Competing interests** The authors declare that they have no competing interests.

**Acknowledgements** All authors would like to express heartfelt gratitude to people involved in field data collections.

**Author’s Contributions** Yunqi Guo: Data management, results visualization, writing and editing. Yanling Zhao: Conceptualization, Oversight the research activity planning and execution. Haoyue Yan: Software, Validation.

**References**

Bell F. G, Stacey T. R, Genske D. D (2000) Mining subsidence and its effect on the environment: some differing examples. Environ Geol 40(1)

Chen Y, Guerschman JP, Cheng Z, Guo L (2019) Remote sensing for vegetation monitoring in carbon capture storage regions: A review. Appl Energ 240(APR.15): 312-326

Feng H K, Yang F Q, Yang G J, Li Z H, Pei H J, Xing H M (2018) Estimation of chlorophyll content in apple leaves base on spectral feature parameters. Transactions of the CSAE 34(6): 182-188

Gitelson AA (2005) Remote estimation of canopy chlorophyll content in crops. Geophys Res Lett

Gitelson AA, Gritz Y, Merzlyak MN (2003) Relationships between leaf chlorophyll content and spectral
reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. J Plant Physiol 160(3): 271-282

Gitelson AA, Kaufman YJ, Merzlyak MN (1996) Use of a green channel in remote sensing of global vegetation from EOS-MODIS. Remote Sens Environ 58(3): 289–298

Gitelson AA, Kaufman YJ, Merzlyak MN (1996) Use of a green channel in remote sensing of global vegetation from EOS-MODIS. Remote Sens Environ 58: 289-298

Han X, Wei Z, Chen H, Zhang B, Li Y, Du T (2021) Inversion of Winter Wheat Growth Parameters and Yield Under Different Water Treatments Based on UAV Multispectral Remote Sensing. Front Plant Sci 12

Hu X, Li X (2019) Information extraction of subsided cultivated land in high-groundwater-level coal mines based on unmanned aerial vehicle visible bands. Environ Earth Sci. 78(14)

Hu Q, Yuan Z (2021) Research on several fundamental issues of coal mining subsidence control in plain coal mining area of the Lower Yellow River. J China Coal Soc 46(05): 1392-1403 (in Chinese with English abstract)

Hu Z, Yang G, Xiao W, Li J, Yang Y, Yu Y (2014) Farmland damage and its impact on the overlapped areas of cropland and coal resources in the eastern plains of China. Resour Conserv Recycl 86: 1-8

JA Schnell (1974) Monitoring the vernal advancement and retrogradation (greenwave effect) of natural vegetation. Nasa/gsfct Type Final Report

Jełowicki Ł, Sosnowicz K, Ostrowski W, Osińska-Skotak K, Bakuła K (2020) Evaluation of Rapeseed Winter Crop Damage Using UAV-Based Multispectral Imagery. Remote Sens-Basel 12(16)

Jełowicki Ł, Sosnowicz K, Ostrowski W, Osińska-Skotak K, Bakuła K (2020) Evaluation of Rapeseed Winter Crop Damage Using UAV-Based Multispectral Imagery. Remote Sens-Basel 12(16): 2618

Kaur B, Kaur P (2017) Improving the color image segmentation using fuzzy-C-means. 2016 International Conference on Advanced Communication Control and Computing Technologies (ICACCCCT). IEEE

Komnitsas K, Guo X, Li D (2010) Mapping of soil nutrients in an abandoned Chinese coal mine and waste disposal site. Miner Eng. 23(8): 627-635

Kross A, McNairn H, Lapen D, et al (2015) Assessment of RapidEye vegetation indices for estimation of leaf area index and biomass in corn and soybean crops. Int J Appl Earth Obs 34(1):235-248

Lei B, Fan J L (2015) A Gradient Weighted Thresholding Method for Image Segmentation. ISCIDE 9242: 300-309

Li J (2021) Research on the application of UAV in the third land survey. Low Carbon World 11(04): 130-131 (in Chinese with English abstract)

Li N, Lv X, Li B, Xu S (2019) An Improved Otsu Method Based on Uniformity Measurement for Segmentation of Water Surface Images. International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), 2019, pp. 675-681
Liang H, Yang G J, Dai H Y, Xu B, Yang H, Feng H K, Li Z H, Yang X D (2019) Modeling maize above-ground biomass based on machine learning approaches using UAV remote-sensing data. Plant Methods 15(1)

Liu H, Zhang M, Su L, Chen X, Liu C, Sun A (2021) A boundary model of terrain reconstruction in a coal-mining subsidence waterlogged area. Environ Earth Sci 80(5)

Liu Y, Feng H K, Sun Q, Yang F Q, Yang G J (2020) Estimation of potato above-ground biomass based on fractional differential of UAV hyperspectral. Transactions of the Chinese Society for Agricultural Machinery 51(12): 202-211 (in Chinese with English abstract)

Lu H, Liu C, Li N, Fu X, Li L (2021) Optimal segmentation scale selection and evaluation of cultivated land objects based on high-resolution remote sensing images with spectral and texture features. Environ Sci Pollut Res 28: 27067-27083

Maimaitijiang M, Sagan V, Sidike P et al. (2019) Vegetation Index Weighted Canopy Volume Model (CVMVI) for soybean biomass estimation from Unmanned Aerial System-based RGB imagery. ISPRS J Photogramm 151: 27-41

Otsu N (1979) A Threshold Selection Method from Gray-Level Histograms. IEEE T Syst Man Cybern 9: 62-66

Pu R (2012) Mapping leaf area index over a mixed natural forest area in the flooding season using ground-based measurements and Landsat TM imagery. Int J Remote Sens 33(19-20): 6600-6622

Rajinikanth V, Couceiro MS (2015) Optimal Multilevel Image Threshold Selection Using a Novel Objective Function. Springer India, pp. 177-186

Rajinikanth V, Couceiro MS (2015) Optimal Multilevel Image Threshold Selection Using a Novel Objective Function. In: Mandal JK, Satapathy SC, Kumar Sanyal M, Sarkar PP, Mukhopadhyay A (Hrsg.), Information Systems Design and Intelligent Applications. Springer India, New Delhi, pp. 177-186

Ren H, Xiao W, Zhao Y, Hu Z (2020) Land damage assessment using maize aboveground biomass estimated from unmanned aerial vehicle in high groundwater level regions affected by underground coal mining. Environ Sci Pollut Res 27(23): 21666-21679

Roujean J L, Breon F M (1995) Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. Remote Sens Environ 51: 375-384

Shi J, Jackson T, Tao J, Du J, Bindlish R, Lu L, Chen K (2008) Microwave vegetation indices for short vegetation covers from satellite passive microwave sensor AMSR-E. Remote Sens Environ 112: 4285-4300

Siebers M, Edwards E, J. Berni JA, Thomas M, Salim M, Walker R (2018) Fast Phenomics in Vineyards: Development of GRover, the Grapevine Rover, and LiDAR for Assessing Grapevine Traits in the Field. Sensors 18(9): 2924

Siebers M, Edwards E, Jimenez-Berni J, Thomas M, Salim M, Walker R (2018) Fast Phenomics in Vineyards: Development of GRover, the Grapevine Rover, and LiDAR for Assessing Grapevine Traits in the Field. Sensors 18: 2924
