



OPEN Research on water body information extraction and monitoring in high water table mining areas based on Google Earth Engine

Anya Zhong✉, Zhen Wang & Yulong Gen

The extensive and intensive exploitation of coal resources has led to a particularly prominent issue of water accumulation in high groundwater table mining areas, significantly impacting the surrounding ecological environment and directly threatening the red line of cultivated land and regional food security. To provide a scientific basis for the ecological restoration of water accumulation areas in coal mining subsidence, a study on the extraction of water body information in high groundwater level subsidence areas is conducted. The spectral characteristics of land types within mining subsidence areas were analyzed through the application of the Google Earth Engine (GEE) big data cloud platform and Landsat series imagery. This study addressed technical bottlenecks in applying traditional water indices in mining areas, such as spectral interference from coal slag, under-detection of small water bodies, and misclassification of agricultural fields. An Improved Normalized Difference Water Index (INDWI) was proposed based on the analysis of spectral characteristics of surface objects, in conjunction with the OTSU algorithm. The effectiveness of water body extraction using INDWI was compared with that of Normalized Difference Water Index (NDWI), Enhanced Water Index (EWI), and Modified Normalized Difference Water Index (MNDWI). The results indicated that: (1) The INDWI demonstrated the highest overall accuracy, surpassing 89%, and a Kappa coefficient exceeding 80%. The extraction of water body information in mining areas was significantly superior to that achieved by the other three prevalent water indices. (2) The extraction results of the MNDWI and INDWI water Index generally aligned with the actual conditions. The boundaries of water bodies extracted using MNDWI in mining subsidence areas were somewhat ambiguous, leading to the misidentification of small water accumulation pits and misclassification of certain agricultural fields. In contrast, the extraction results of INDWI exhibited better alignment with the imagery, with no significant identification errors observed. (3) Through the comparison of three typical areas, it was concluded that the clarity of the water body boundary lines extracted by INDWI was higher, with relatively fewer internal noise points, and the soil ridges and bridges within the water bodies were distinctly visible, aligning with the actual situation. The research findings offer a foundation for the formulation of land reclamation and ecological restoration plans in coal mining subsidence areas.

Coal is acknowledged as the primary energy source and a critical raw material in the country, thereby ensuring national energy security and facilitating the rapid growth of the national economy^{1,2}. Nonetheless, coal mining activities are associated with a range of adverse ecological and environmental consequences, directly influencing regional land use and sustainable development. The incidence of surface subsidence attributable to coal extraction is escalating, resulting in the degradation of farmland, impairment of public infrastructure, and, in extreme instances, endangering the life safety of local inhabitants³. The eastern region of China with a high groundwater table constitutes a significant farmland protection zone within the country, marked by a dense population, extensive cultivated land, and elevated soil fertility⁴. Owing to the elevated groundwater levels in the region, the occurrence of surface subsidence results in the formation of numerous depressions within the farmland, characterized by central low-lying areas that are susceptible to waterlogging. The resultant subsidence-induced

School of Geoscience and Surveying Engineering, China University of Mining and Technology (Beijing), No. 11, Xueyuan Road, Haidian District, Beijing 100083, China. ✉email: bqt2300204054@student.cumtb.edu.cn

water accumulation extensively damages agricultural lands and necessitates the relocation of villages, thereby engendering a multitude of environmental and social challenges^{5,6}. Consequently, the accurate extraction of information pertaining to subsidence water bodies within mining zones, coupled with targeted land reclamation initiatives, serves as a critical safeguard for the sustainable and environmentally-friendly development of regional mining operations.

The ongoing evolution of remote sensing technology has led to a continuous enhancement in the temporal, spatial, and spectral resolutions of satellite imagery, thereby facilitating the precise, comprehensive, dynamic, and near real-time monitoring of surface water resources⁷. Traditional water body extraction methods can be classified into visual interpretation, automatic classification and semi-automatic classification⁸. The visual interpretation method involves the manual differentiation of water bodies from other land features on remote sensing images, based on the experience of researchers, to delineate the boundaries of the water body sections. While the visual interpretation method is highly reliable, it requires substantial manpower support, making it difficult to extract water body information from large volumes of remote sensing data⁹. Secondly, the automatic classification method employs programs to establish rules for land feature classification. Based on the performance differences of various land features, algorithms are utilized to analyze these features, achieving automatic classification. Common automatic classification methods include machine learning, ISODATA, and K-means clustering, among others^{10,11}. To a certain extent, the automatic classification method reduces the influence of human judgment factors, with a moderate improvement in universality. However, its classification accuracy is still inferior to that of the semi-automatic classification method when compared, resulting in less extensive application in remote sensing image classification. The semi-automatic classification method takes advantage of the reflectance differences between water bodies and other land features in remote sensing images to set classification rules, and extracts water body information through various software and algorithms. At present, commonly used semi-automatic classification methods mainly include the water body index method, the inter-band relationship method, and the object-oriented method, among others^{12,13}. Water index methodologies employ algebraic operations between spectral bands to accentuate the distinctions between water bodies and background features. Characterized by their simplicity, these methods facilitate rapid and effective data processing, yielding elevated classification accuracies. Consequently, they demonstrate significant advantages in the large-scale extraction of water bodies^{14,15}. The Normalized Difference Water Index (NDWI) was first proposed by the foreign scholar Mcfeeters. On the basis of the NDVI, he analyzed the spectral ranges of different bands of water bodies on TM images and constructed the NDWI using the green band and the near-infrared band. Although the NDWI can suppress vegetation information, it is affected by various environmental factors when dealing with images of complex land cover compositions, resulting in a reduction of water body extraction accuracy¹⁶. In addressing this challenge, Xu introduced an enhanced version of the Modified Normalized Difference Water Index (MNDWI) replaced the near-infrared band with the mid-infrared band. Despite its advancements, the MNDWI method was found to be susceptible to the effects of mountain shadows¹⁷. Feyisa et al. proposed a new automated extraction method, termed the Automated Water Extraction Index (AWEI), which employed TM imagery to enhance spectral contrast, thereby improving the accuracy of mapping. This method has shown excellent effectiveness in eliminating mountain shadows¹⁸. Jiang et al.¹⁹ proposed the Shadow Water Index (SWI), utilizing Sentinel-2 imagery, which exhibited effective extraction performance for pure water, turbid water, saltwater, and floating ice. Despite the endeavors of researchers to develop optimal water indices customized for specific regions through the utilization of various data sources, a specialized index for the extraction of water bodies within mining subsidence areas remains undeveloped.

Mining areas characterized by high groundwater levels predominantly occur in the eastern region of China. Over the past two decades, Chinese scholars have engaged in extensive research on the monitoring of mining subsidence water bodies, encompassing a series of investigations into the methodologies for identifying such water bodies and the utilization of relevant data^{20,21}. Since the 1980s, the delineation of mining subsidence water bodies by Chinese scholars was predominantly achieved through field surveys, employing geodetic and leveling techniques to precisely determine the water body area. The rapid advancement of remote sensing technology resulted in an increasing number of scholars initiating the use of remote sensing data for research purposes, thereby facilitating the direct interpretation of water body areas within mining regions from imagery. Thereafter, the emergence of high-resolution images facilitated a concomitant enhancement in monitoring accuracy. Peng et al.²² employed Principal Component Analysis (PCA) to extract the extent of subsidence water bodies from remote sensing images of various periods, thereby aiding in the dynamic monitoring and remediation of these subsidence water bodies. Li et al. developed a comprehensive method for identifying water bodies, which combined the strengths of the MNDWI method and the HIS spatial water body model, utilizing Landsat TM imagery. This approach was specifically tailored for Jining City and exhibited notable efficacy in detecting scattered surface water accumulations resulting from surface subsidence²³. Wang et al.²⁴ found that the integration of the improved normalized difference water index with GIS spatial overlay technology, utilizing TM imagery, could effectively facilitate the dynamic monitoring of water bodies in mining areas.

The water index extraction method, grounded in remote sensing technology, is recognized for its simplicity and convenience. Nonetheless, as of now, no index has been specifically devised for the extraction of water bodies in mining areas characterized by high groundwater levels. This study addresses the technical issues associated with the application of traditional water indices in mining areas characterized by high groundwater levels. These issues include spectral interference from coal slag, the omission of small water bodies, and the misclassification of agricultural land. An INDWI was proposed based on the MNDWI to eliminate the influence of extraneous factors on water bodies and to accurately capture information regarding subsiding water bodies in these areas. The findings are of practical significance for land reclamation and ecological protection in mining regions.

Materials and methods

Study area

The Guqiao Coal Mine, positioned in the northwest of Fengtai County, Huainan City, Anhui Province, encompasses an area of 92.68 km² (Fig. 1). The study area, characterized as a subtropical monsoon climate zone, witnesses frequent convergence of cold and warm air masses. The climate is temperate with moderate rainfall. It features distinct seasons, having long summers and winters, short springs and autumns, ample sunlight, and significant influence from monsoons. The mining operations are conducted underground. The geomorphological classification of the area is the Huaibei accumulation-erosion plain, characterized by a gentle slope from the northwest to the southeast. The predominant soil types within the region are fluvo-aquic soils and sandy ginger soils, with their distribution exhibiting the following traits: in riverine areas, these soils are developed on loess-like sediments, predominantly consisting of fluvo-aquic and yellow fluvo-aquic soils; in the interfluvial plains, the soil parent material is akin to that found in riverine areas, with sandy ginger black soils being formed due to elevated groundwater levels. The study area has been mined for many years since its commencement in 2007. Most of the subsided land has been altered from its original topography due to human activities. As a result of coal mining, large areas of the surface have gradually formed accumulations of water, leading to the relocation of villages, inundation of farmland, and the conversion of extensive building plots, roads, and farmland into subsided water bodies, among other issues.

Data resources and pre-processing

Based on the GEE cloud platform, this study conducts a before-and-after comparison of coal mining activities. In the process, image data sources significantly affected by cloud cover are eliminated. Ultimately, six Landsat series (Landsat5 TM and Landsat8 OLI) remote sensing images during the wet season (from June to September) in 2005, 2007, 2010, 2013, 2018, and 2021 are selected. Among them, the data from 2005 serve as a control before coal mining. With 2007 as the reference, images are selected at an interval of three years. Due to the poor quality of the data in 2016, the data from 2018 is used as a substitute to study the changes in the area of subsidence water bodies throughout the mining area (Fig. 2). Following the exclusion of remote sensing images characterized by significant cloud cover and diminished visibility, a final selection of 55 instances of remote sensing imagery that fulfilled the fundamental criteria for water body extraction was made. The selected images underwent preprocessing, which included radiometric calibration, atmospheric correction, and image trimming, resulting in the acquisition of the definitive imagery. The relevant details of the imagery are recorded in Table 1.

Methodology

INDWI model construction

The subsided zones at Guqiao Coal Mine, predominantly comprising arable land, are a consequence of coal extraction. The water-filled depressions resulting from surface subsidence exhibit unmanaged edges, integrating indistinguishably with the adjacent farmland, which is heavily vegetated. The soil along the peripheries of

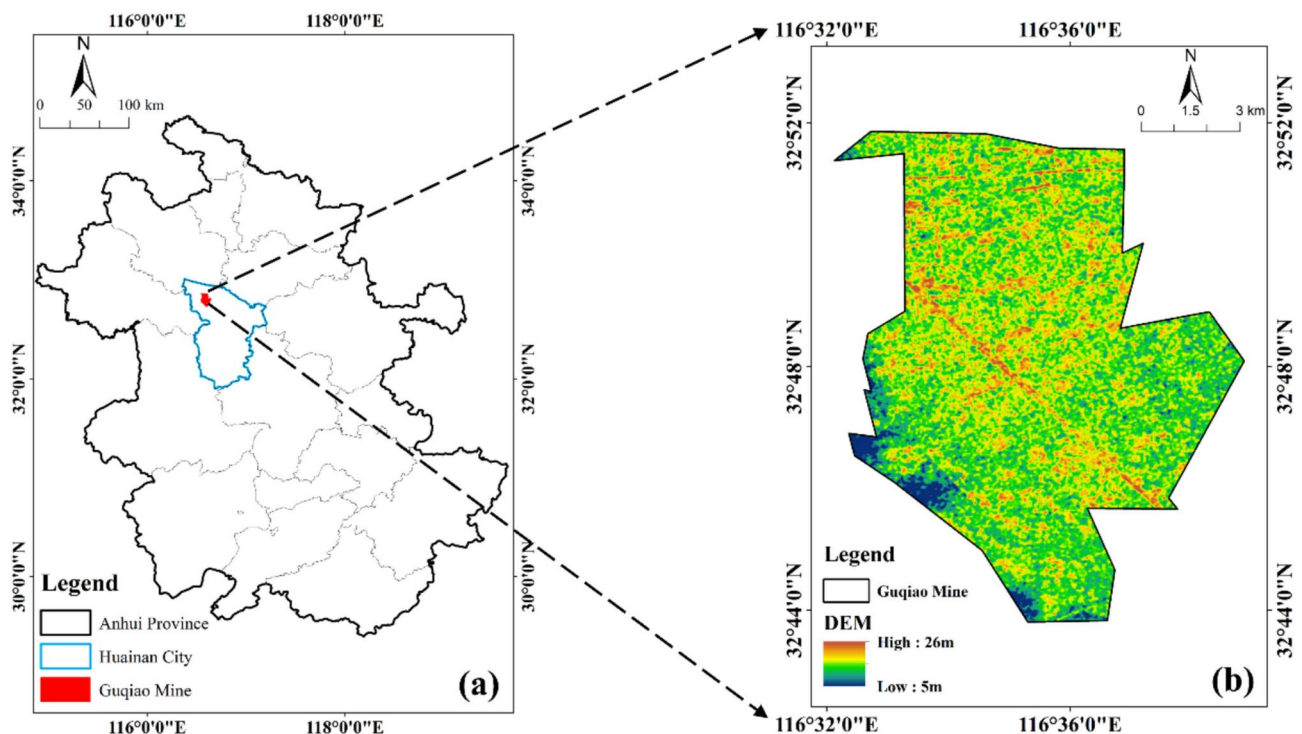


Fig. 1. Map of study area; (a) Location of the study area in Anhui Province, China, (b) Topographic characteristics of the study area.

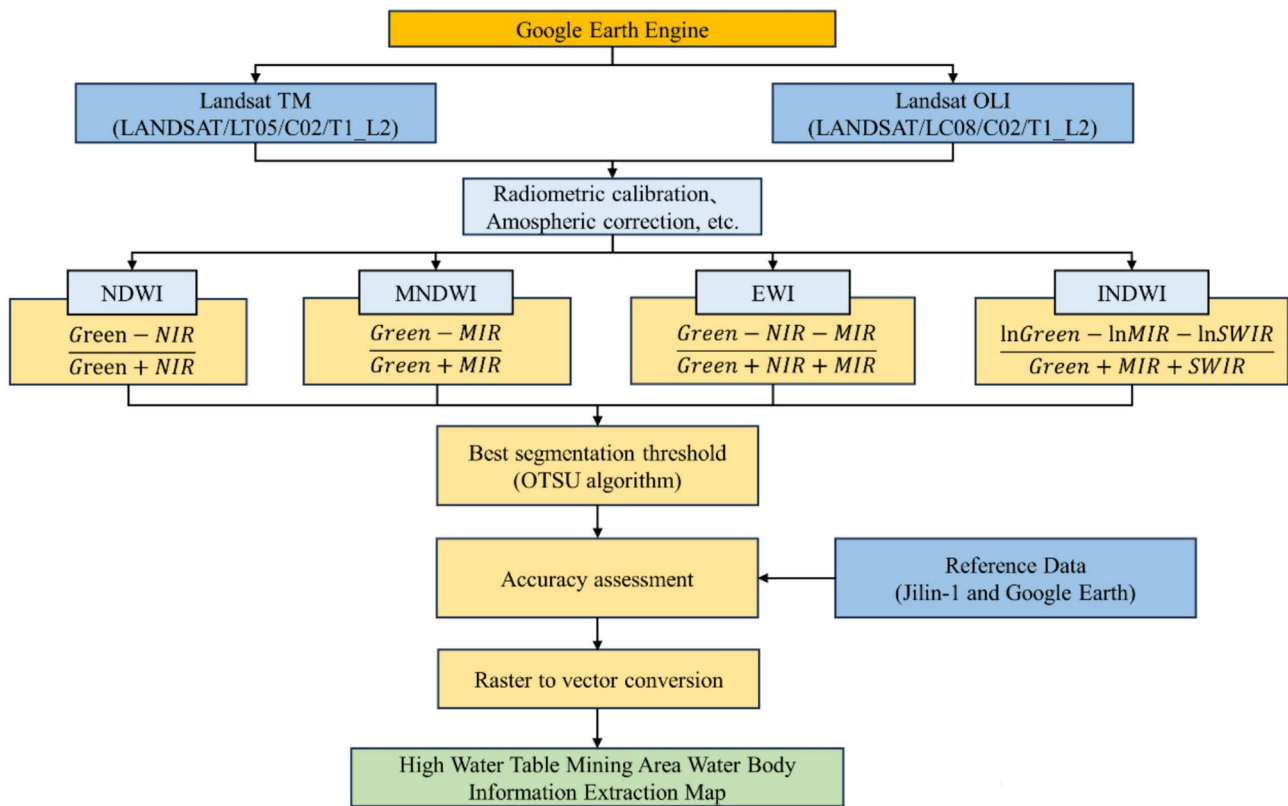


Fig. 2. Technical flow chart of the study.

Satellite	Band	Sensitivity spectrum (μm)	Spatial resolution (m)
Landsat 5 TM	Blue	0.45–0.52	30
	Green	0.52–0.60	30
	Red	0.63–0.69	30
	Near IR	0.76–0.90	30
	SWIR	1.55–1.75	30
Landsat 8 OLI	Blue	0.45–0.52	30
	Green	0.53–0.60	30
	Red	0.63–0.68	30
	Near IR	0.85–0.89	30
	SWIR 1	1.56–1.66	30
	SWIR 2	2.10–2.30	30

Table 1. Landsat image parameters information.

these water pits is notably water-logged. Furthermore, coal and mining detritus are stored in the open within the mining zone, and substantial coal dust is dispersed during coal transport. These conditions introduce interference in the delineation of water body areas²⁵. Research indicated that the MNDWI performed poorly in distinguishing soils with high water content. During the experiment, MNDWI, relying solely on the Green and MIR bands, tended to blur the boundaries of extracted water bodies, leading to the misidentification of other mining area features as water bodies. Consequently, the obtained area of water bodies in the mining zone was significantly overestimated²⁶. Therefore, it is necessary to identify a more suitable method for extracting subsided water bodies in mining areas with high water tables.

Within remote sensing imagery, water bodies are characterized by an overall lower reflectance relative to other land features, especially in the infrared spectral range. In the shortwave infrared (SWIR) band, water bodies exhibit the capacity to absorb almost all incident energy, resulting in a precipitous drop in their reflectance values. Conversely, the reflectance values of other land features are significantly higher than those of water bodies²⁷. A comparison of reflectance values among different features on the imagery reveals that water bodies, pit edges, and vegetation along the water's edge display comparable reflectance values in the visible light spectrum. In

the infrared band, their brightness trends are generally similar; however, the disparities are evident, with water bodies demonstrating the lowest reflectance values, while all other features present higher values (Fig. 3). The pronounced absorption by water bodies in the infrared spectrum, particularly at wavelengths extending beyond the mid-infrared, provides a clear advantage for distinguishing water bodies from other land features. This is particularly evident for features that are prone to confusion with water bodies, including collapsed pit edges, aquatic vegetation, and bare land.

Analysis of Fig. 3 demonstrates that the SWIR band is characterized by substantial absorption by water bodies, with vegetation, pit edges, and bare land all exhibiting higher reflectance values in this band. Studies have shown that SWIR offers notable advantages in the identification of vegetation cover, moist soil, and minerals^{28,29}. On the basis of the MNDWI, by subtracting the SWIR from the numerator and adding it to the denominator, the differentiation between water bodies and other land features becomes more pronounced due to the minimal reflectance values of water bodies at this stage. The Preliminary Water Index (PWI) constructed accordingly is as follows:

$$PWI = \frac{\beta_G - \beta_{MIR} - \beta_{SWIR}}{\beta_G + \beta_{MIR} + \beta_{SWIR}} \quad (1)$$

where β_G represents the green band, β_{MIR} denotes the mid-infrared band, and β_{SWIR} signifies the shortwave infrared band. In Landsat TM, the green band corresponds to the second band, and the mid-infrared band is the fifth. For Landsat OLI data, the green band is the third, and the mid-infrared band is the sixth. In both Landsat TM and OLI, the shortwave infrared band is the seventh.

Furthermore, the opposing reflectance change trends between water bodies and background features are essential for differentiating water bodies from background features. To accentuate this trend, differential enhancement processing is applied to the water information within mining areas and the surrounding land types. For the numerator of *PWI* in Eq. (1), using a monotonically increasing function with a slope greater than 1 can enhance the downward trend of water body reflectance. Considering that the reflectance of various background features in both bands ranges from 0 to 1, four commonly used monotonically increasing functions, $y = \sin x$, $y = \tan x$, $y = e^x$, and $y = \ln x$, are selected as enhancement functions³⁰. As the function $y = x$ with a slope of 1 accurately reflects the inherent reflectance differences of background features between bands, it is employed as the control function. The curves of the five functions within the reflectance range of 0–1 are depicted in Fig. 4.

Analysis reveals that the slope (growth rate) of the function $y = \ln x$ exceeds those of other functions when reflectance ranges between 0 and 1. Furthermore, the growth trend of $y = \ln x$ becomes increasingly pronounced as reflectance diminishes below 0.3, thereby enhancing the reflectance disparity between the green and infrared bands for water bodies. Therefore, the function $y = \ln x$ is selected to enhance the reflectance difference of water

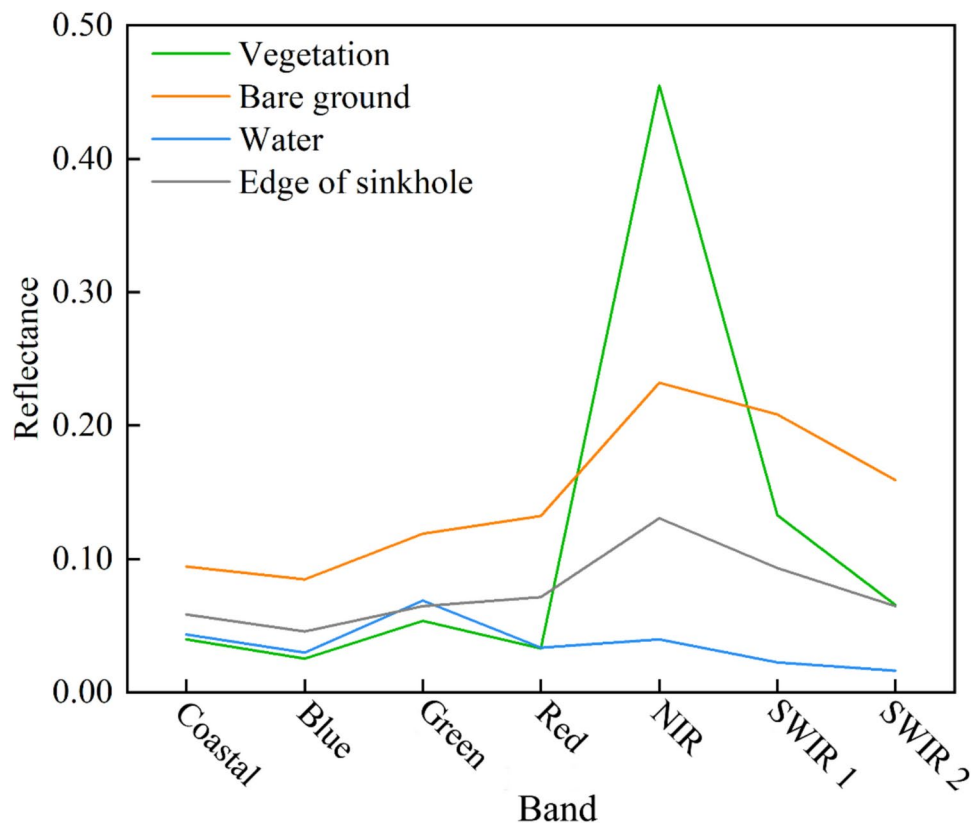


Fig. 3. Spectral characterization curves of typical features in the mining area.

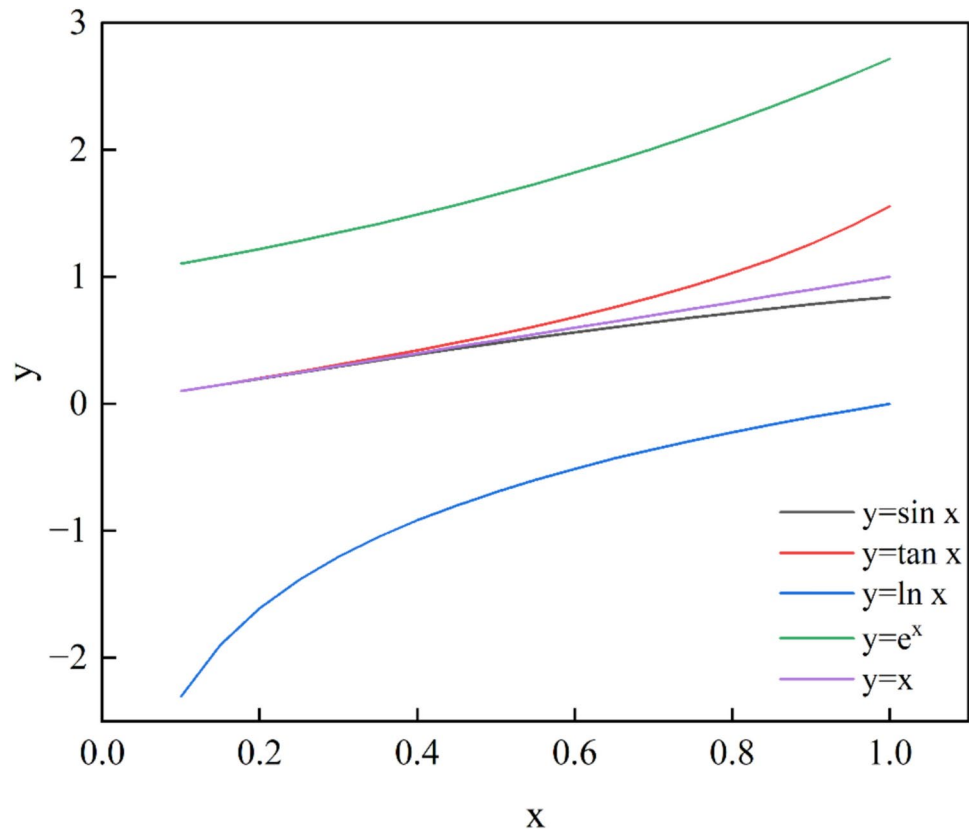


Fig. 4. Curves of each function as a function of reflectance between 0 and 1.

bodies in the red edge and near-infrared bands. For the denominator term of *PWI* in Eq. (1), considering that an increase in the denominator would lead to a decrease in the overall value, thereby offsetting the effect of reflectance difference, the denominator term is maintained in its original form without stretching. The final result is *INDWI*:

$$INDWI = \frac{\ln \beta_G - \ln \beta_{MIR} - \ln \beta_{SWIR}}{\beta_G + \beta_{MIR} + \beta_{SWIR}} \quad (2)$$

where the expression's meaning is as defined in Eq. (1).

Otsu algorithm

The Otsu method (Nobuyuki Otsu method, OTSU) is an algorithm used for determining the binary threshold of an image^{31,32}. OTSU is based on the gray-level characteristics of image pixels and measures the uniformity of the gray-level distribution by calculating the interclass variance. A greater interclass variance between the foreground and background indicates a larger difference between the two components of the image, thereby enhancing the distinction between the foreground and background³³.

Some researchers applied the Otsu method to water extraction studies based on water indices, and the results demonstrated the ability to obtain reliable water information, significantly improving the accuracy of water extraction^{34,35}. In this study, the Otsu method was employed to determine the optimal segmentation threshold. For an image $I(x, y)$, the parameter θ is designated as the segmentation threshold between the foreground (water body) and background (non-water body). Initially, a gray level from the water index is established as the initial segmentation threshold θ , and the corresponding interclass variance g is computed. Subsequently, the gray levels are iteratively processed, with the corresponding g values calculated until all gray levels within the water index image are exhausted. The segmentation threshold θ associated with the maximum g value is then identified as the optimal segmentation threshold. The algorithm for determining the g value is delineated as follows:

$$g = w_0 w_1 (u_0 - u_1)^2 \quad (3)$$

$$w_0 = \frac{N_0}{M} \quad (4)$$

$$w_1 = \frac{N_1}{M} \quad (5)$$

$$u_0 = \frac{\sum_{i=1}^{N_0} u_{0i}}{N_0} \quad (6)$$

$$u_1 = \frac{\sum_{j=1}^{N_1} u_{1j}}{N_1} \quad (7)$$

where w_0 denotes the ratio of the number of pixels occupied by the water body to the total number of pixels in the image; w_1 signifies the ratio of the number of pixels occupied by non-water bodies to the total number of pixels in the image; u_0 represents the average gray level of the foreground, namely the water body, within the image; u_{0i} indicates the gray level of the i th pixel within the water body; u_1 denotes the average gray level of the background, which encompasses all non-water bodies, within the image; u_{1j} signifies the gray level of the j th pixel within the non-water bodies; N_0 is the count of pixels constituting the water body; N_1 is the count of pixels constituting the non-water bodies; M represents the aggregate number of pixels in the selected image data.

Comparative water body index

To validate the effectiveness of the INDWI water index, the NDWI, MNDWI, and EWI water indices are selected for comparison. The formulas for these three water indices are as follows:

$$NDWI = \frac{\beta_G - \beta_{NIR}}{\beta_G + \beta_{NIR}} \quad (8)$$

$$MNDWI = \frac{\beta_G - \beta_{MIR}}{\beta_G + \beta_{MIR}} \quad (9)$$

$$EWI = \frac{\beta_G - \beta_{NIR} - \beta_{MIR}}{\beta_G + \beta_{NIR} + \beta_{MIR}} \quad (10)$$

where β_G , β_{NIR} , and β_{MIR} denote the reflectance of the Landsat series images in the green, near-infrared, and shortwave infrared bands, respectively.

Accuracy evaluation index

The Overall Accuracy (OA), Kappa coefficient, User's Accuracy (UA), and Producer's Accuracy (PA) are used to evaluate the precision of water body extraction for each water index. The OA is defined as the ratio of the number of correctly classified category pixels (samples) to the total number of category pixels. The Kappa coefficient quantifies the proportionate reduction in error relative to a purely random classification. UA is defined as the ratio of the number of samples correctly classified as a specific land cover type to the total number of samples classified as that type. PA is defined as the ratio of the number of samples accurately classified as a specific feature to the total number of actual samples for that feature. The formula for each evaluation indicator is shown below:

$$OA = \sum_{k=1}^n p_{kk}/p \quad (11)$$

$$Kappa = \frac{p \sum_{i=1}^n p_{ii} - \sum_{i=1}^n (p_{i+} p_{+i})}{p^2 - \sum_{i=1}^n (p_{i+} p_{+i})} \quad (12)$$

$$UA = \frac{P_{ii}}{P_{i+}} \times 100\% \quad (13)$$

$$PA = \frac{P_{ii}}{P_{+i}} \times 100\% \quad (14)$$

where P_{kk} signifies the count of samples accurately classified into category k ; P represents the aggregate sample count; P_{ii} denotes the count of samples correctly categorized into category i ; P_{i+} is the total number of samples assigned to category i through classification; P_{+i} indicates the total number of samples in category i derived from visual interpretation.

Results and analysis

Accuracy verification

The extraction accuracy of the constructed INDWI index was validated, and the classification results of NDWI, EWI, MNDWI, and INDWI were evaluated using a confusion matrix. In this study, image data from the year 2021 was selected, and optical remote sensing data with a resolution of 0.75 m from Jilin-1 was utilized to choose the same 50 random points as the validation sample areas. The overall accuracies of the classification results obtained from Landsat images for NDWI, EWI, MNDWI, and INDWI were 80.85%, 90.84%, 91.49%, and 93.62% respectively. The Kappa coefficients were 71.65%, 81.72%, 82.82%, and 87.08% respectively.

To address the impact of image changes in different years on water body extraction results, visual interpretation characteristics were summarized based on the 2021 high-resolution image data. Water body areas were extracted from Landsat images in 2007, 2010, 2013, 2018, and 2021 using the expert interpretation method as verification

data. The same random points were selected from the verification data of each year as validation samples to evaluate the accuracy of the classification results of various water body indices (Table 2).

The findings revealed that the NDWI index provided the lowest accuracy in water body extraction, ranging from approximately 60% to 70%. The INDWI index, developed in this study, attained the highest overall accuracy, surpassing 89%, accompanied by a Kappa coefficient exceeding 80%. The accuracy of both UA and PA indicators is above 80%. Conclusively, the INDWI index exhibited superior efficacy in extracting water information within mining areas compared to the other three widely utilized water indices, demonstrating elevated accuracy.

The preceding analysis of water extraction overall accuracy demonstrated that the MNDWI and INDWI indices exhibited the highest precision. Accordingly, the subsequent investigation employed these two indices to extract water bodies from Landsat imagery and superimposed the outcomes with Jilin-1 data at a resolution of 0.75 m (Fig. 5). The map in Fig. 5 was created using ArcGIS Desktop (version 10.8, <https://www.esri.com/>). Upon comparative analysis, it was ascertained that the overall extraction outcomes of the two water indices were largely congruent with the actual circumstances. Nonetheless, nuances in detail emerged. Specifically, the MNDWI-derived boundaries of subsided water bodies appeared indistinct, especially proximate to the mining site, encompassing several small-scale water accumulation pits and occasional misclassification of farmland, thereby inflating the water area estimate. Conversely, the INDWI extraction results fitted better with the images, and the boundaries of collapse water accumulation areas and roads were recognized more clearly, without obvious recognition errors.

Additionally, area accuracy was applied in this paper to compare MNDWI and INDWI. Area accuracy denotes the proportion of the area of subsided water bodies in mining regions acquired via the water index method in relation to the actual area. The on-site measured water body area data were utilized as the true values, and the water body areas obtained from MNDWI and INDWI were compared with them respectively (Table 3).

The findings revealed that the calculated area accuracy of INDWI was as high as 96.41%, surpassing that of MNDWI by 6.61%. This index exhibited superior efficacy in extracting water body information from mining areas with high groundwater levels, characterized by elevated overall accuracy. Conversely, while MNDWI was capable of generally capturing the water body area within mining regions, its area accuracy was inferior. The water body area extracted by MNDWI exceeded the actual, manifesting substantial discrepancies even at a minor scale. Therefore, its application to broader study areas would exacerbate these errors, compromising the precision of the extraction outcomes. The principal cause may have been the misclassification of water-adjacent vegetation, farmland, and collapse pit edges as water bodies by MNDWI, resulting in an overestimation of the extraction outcomes. In mining areas characterized by high groundwater levels, the precise delineation of water body extents is pivotal for subsequent land reclamation efforts. The Landsat data, when utilized with the INDWI index for water body extraction, exhibited congruity with the 0.75m resolution high-resolution imagery data employing supervised classification, maintaining comparable accuracy levels. The integration of the INDWI index with Landsat data could thus be more effectively deployed in the investigation of subsided water bodies within areas of elevated groundwater levels.

Year	Index	OA (%)	Kappa (%)	UA (%)	PA (%)
2007	NDWI	66.14	60.97	63.58	62.47
	EWI	76.32	68.25	65.34	67.32
	MNDWI	87.17	78.63	75.21	73.88
	INDWI	92.36	87.21	85.96	86.63
2010	NDWI	65.89	60.83	62.39	61.15
	EWI	75.78	69.06	68.22	69.11
	MNDWI	89.24	75.75	74.13	73.52
	INDWI	95.08	88.62	86.67	87.48
2013	NDWI	68.28	63.41	64.91	63.89
	EWI	74.56	68.37	66.72	66.43
	MNDWI	88.63	76.24	77.36	77.32
	INDWI	90.74	80.32	81.47	80.06
2018	NDWI	61.66	52.64	55.28	53.67
	EWI	77.42	63.68	61.35	63.93
	MNDWI	83.61	70.27	71.20	69.38
	INDWI	89.71	81.10	82.34	80.82
2021	NDWI	80.85	71.65	73.37	72.66
	EWI	90.84	81.72	79.55	80.53
	MNDWI	91.49	82.82	83.99	84.08
	INDWI	93.62	87.08	86.37	88.76

Table 2. Classification accuracy of different water body indices by year.

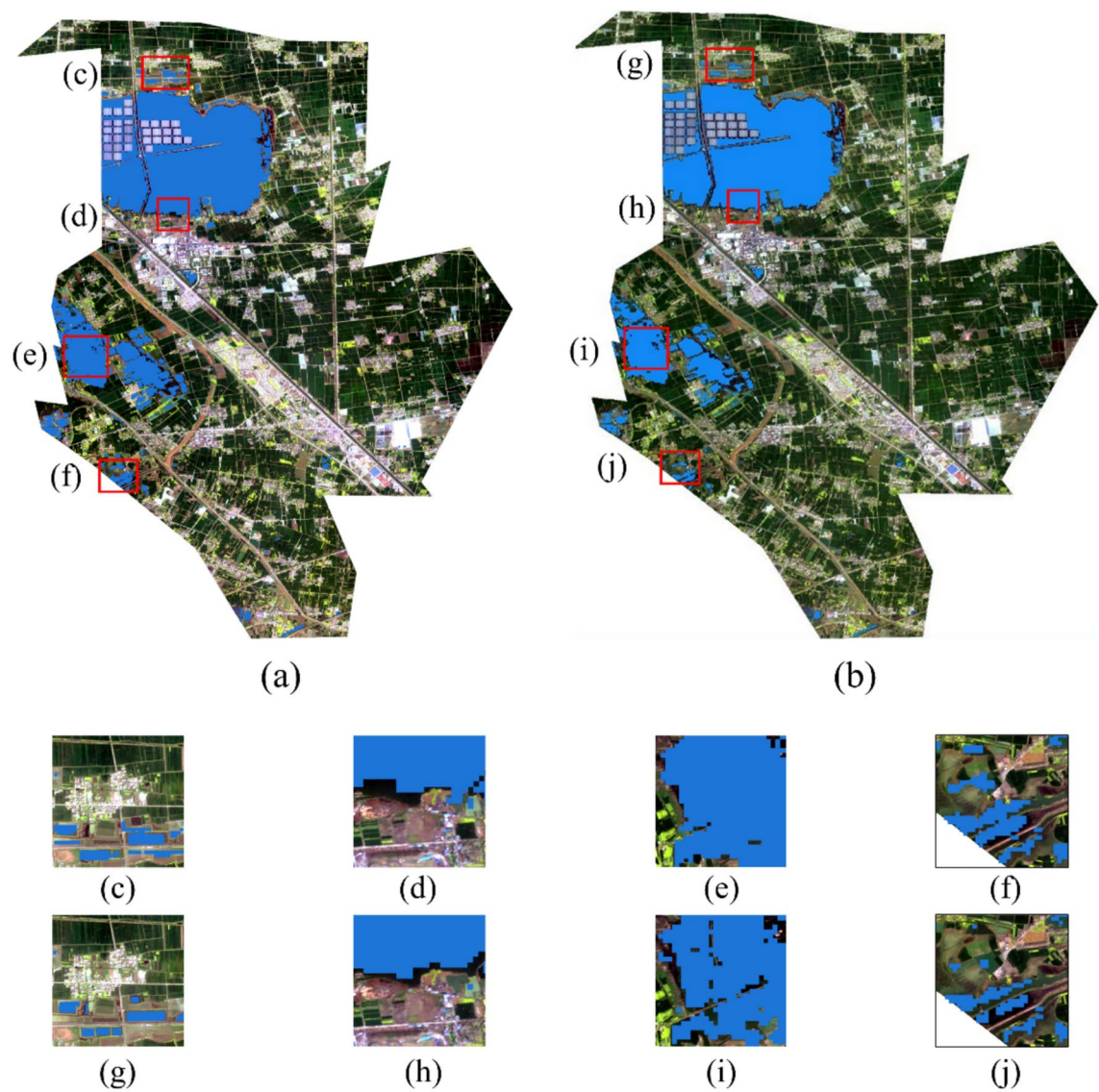


Fig. 5. Comparison of MNDWI and INDWI waterbody information overlay high score data. **(a)** MNDWI superimposed high score images; **(b)** INDWI superimposed high score images; **(c)** MNDWI overlay map in study area 1; **(d)** MNDWI overlay map in study area 2; **(e)** MNDWI overlay map in study area 3; **(f)** MNDWI overlay map in study area 4; **(g)** INDWI overlay map in study area 1; **(h)** INDWI overlay map in study area 2; **(i)** INDWI overlay map in study area 3; **(j)** INDWI overlay map in study area 4.

Water body extraction data types	Acquisition Time	Area of water body (km ²)	Precision (%)
Measured data on-site	2021-03-20	10.37	–
Jilin-1 Oversight classification	2021-03-28	10.04	96.82
Landsat 8 MNDWI	2021-03-24	11.18	89.80
Landsat 8 INDWI	2021-03-24	9.68	96.41

Table 3. Comparison of MNDWI and INDWI extracted water body areas.

Results of water body classification

The study area, endowed with abundant coal resources, constitutes a pivotal energy foundation in Anhui Province. The region's elevated groundwater levels, compounded by the consequences of underground coal mining, have progressively given rise to expansive zones of standing water. This development has subsequently led to the displacement of villages, the flooding of agricultural lands, and the conversion of extensive tracts of construction land, roadways, and farmlands into subsided water bodies. Analysis of Landsat series data pre- and post-underground coal mining indicated that prior to extraction, the mining area's water bodies were exclusively

confined to the old Xifei River segment, with no detection of additional aquatic regions. Consequently, the river area was systematically omitted during the subsequent delineation of subsided water bodies within the mining area.

Remote sensing imagery obtained post-mining from the study area was chosen for the analysis of dynamic subsidence changes in water bodies. The MNDWI and INDWI methodologies were utilized to delineate the subsided water bodies within the mining zone spanning from 2005 to 2021 (Fig. 6), with the Otsu algorithm employed to establish an optimal water threshold. Subsequently, ArcGIS 10.8 was leveraged for statistical analysis, culminating in the quantification of the area encompassed by the subsided water bodies in the mining region (Fig. 7, Table 4).

The statistical analysis revealed that prior to mining, no subsidence was observed on the surface of the mining area, resulting in a water body area of 0. Following the commencement of mining, the area of subsided water bodies within the mining zone exhibited a sustained increase over the subsequent decade. Notably, the area delineated by MNDWI was substantially larger than that by INDWI, with discrepancies in the annual proportions of subsided water body areas amounting to 0.85%, 1.15%, 0.35%, 1.58%, and 1.50%, respectively. The areas of subsided water bodies within the study region, delineated by the MNDWI and INDWI methodologies, escalated from 1.54 and 0.75 km² in 2007 to 10.36 and 8.97 km² by 2021, respectively. Corresponding growth rates were quantified at 85.14% and 91.64%, respectively. These findings underscore the substantial influence of underground mining operations on the surface landscape configuration. Additionally, it was discernible that the surface subsidence induced by the initial phase of mining within the mining area exerted a minimal impact, with the subsidence depth failing to reach the water table, thereby facilitating the identification of subsided water pits. As a result, the extraction efficacy of the two water indices was equivalent during the period from 2007 to 2013. Subsequent to this, the expansion of underground mining operations led to an escalating discrepancy in the areas delineated by the two water indices between 2018 and 2021.

Discussion

Comparison of extraction results for different water body indices

The Guqiao Coal Mine area exemplifies a high water table coal mining region, where decades of mining activities have culminated in the development of extensive water bodies. Employing the Jilin-1 imagery from March 2021 as validation, the Landsat OLI data were utilized to delineate the extent of water bodies within the mining area through the application of four distinct water indices. In order to assess the extraction efficacy of INDWI across diverse regions of the study area, three representative zones were chosen: river bridges, edges of subsidence pit waters, and urban water bodies. The outcomes of the various water indices' extraction across these regions are illustrated in Fig. 8.

Examination of Fig. 8a–f indicated that NDWI demonstrated the least effective extraction of river bridges, predominantly classifying them as water bodies, with EWI being the next least effective. The extraction outcomes of MNDWI and INDWI showed negligible discrepancies, the only variation being that INDWI yielded more distinct water body boundary delineations and fewer erroneous internal classifications.

Analysis of Fig. 8g–l revealed that NDWI and EWI erroneously identified certain residential areas as water bodies, leading to an inflated water area. Conversely, MNDWI and INDWI circumvented this misclassification. Nevertheless, MNDWI exhibited a tendency to incorrectly categorize dry farmland as water bodies. INDWI, on the other hand, showcased enhanced extraction precision, characterized by more distinct and smoother water boundaries, and clear delineation of soil ridges within the water bodies, thereby accurately reflecting the actual scenario.

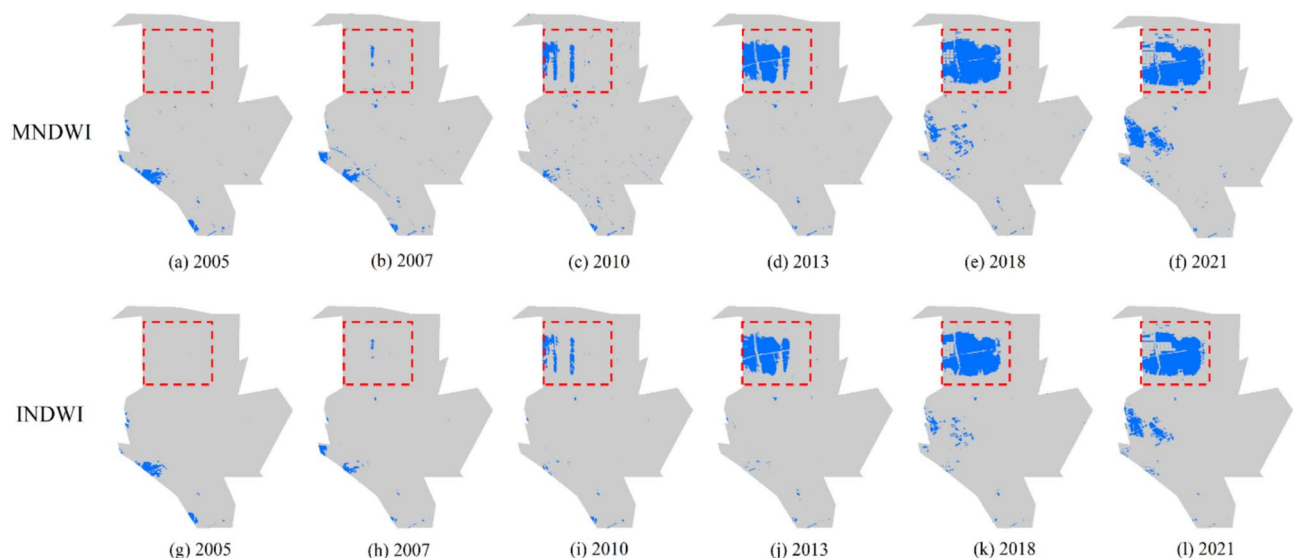


Fig. 6. Extraction results of subsidence water area in the study area.

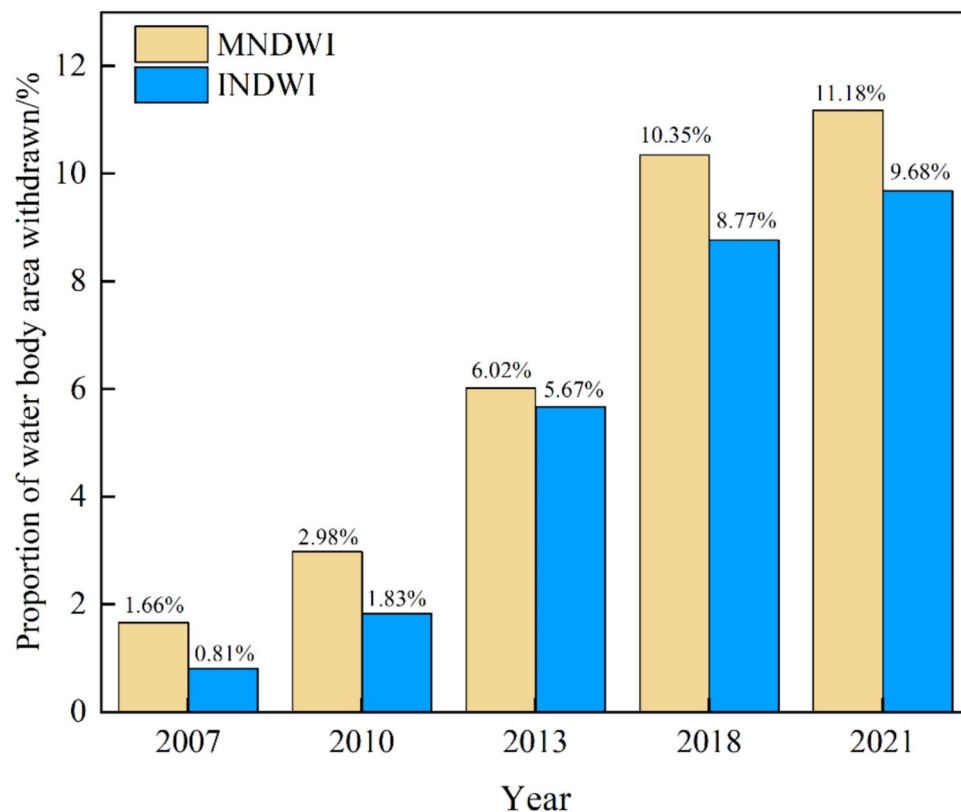


Fig. 7. Percentage of area of sunken water bodies in mining areas.

Water body index	Area of water body subsidence by year/ km ²					
	2005	2007	2010	2013	2018	2021
MNDWI	0	1.54	2.76	5.58	9.59	10.36
INDWI	0	0.75	1.70	5.25	8.13	8.97

Table 4. Statistics on the area of sunken water bodies in mining areas.

Examination of Fig. 8m–r during the urban water body extraction comparison revealed that the red-boxed areas corresponded to urban construction zones. INDWI was capable of extracting the fragmented water bodies in the city with minimal misclassification, in contrast to NDWI, EWI, and MNDWI, which suffered from severe misclassification of background features. Furthermore, the green-boxed areas, representing the water body edge regions, showed that NDWI, EWI, and MNDWI were significantly affected by noise, whereas INDWI effectively delineated the water body boundaries.

In summary, three representative areas including river bridges, collapse pit water edges, and urban water bodies were selected to verify the water body extraction effect of INDWI. Compared with the other three water body indices, INDWI was able to extract water body boundaries more clearly, reduce internal misclassification points, and produce clear and smooth boundaries that better matched the actual situation. Moreover, it could accurately extract fragmented water bodies with very few misclassifications and effectively suppress noise in the water body edge areas. These findings indicate that INDWI has significant advantages in water body extraction accuracy and boundary recognition capability in complex ground object backgrounds. However, it is worth noting that misclassification still occurs at the edges of collapse pit water bodies when using INDWI, which requires further investigation in subsequent studies.

In mining areas characterized by high groundwater levels, underground coal extraction leads to extensive surface subsidence, subsequently resulting in waterlogging, alterations in the land use structure, soil pollution, and disruption of the groundwater system. These factors collectively hinder the sustainable development of mining areas³⁶. Furthermore, high groundwater level mining areas are primarily situated in the Huang-Huai-Hai Plain region of China, characterized by a high proportion of basic farmland. The shallow groundwater is susceptible to surface waterlogging as a result of ground subsidence. The ongoing expansion of waterlogging zones leads to the submersion of a considerable amount of high-quality arable land, thereby impacting farmland protection and food security³⁷. Consequently, there exists a necessity for the precise identification and extraction

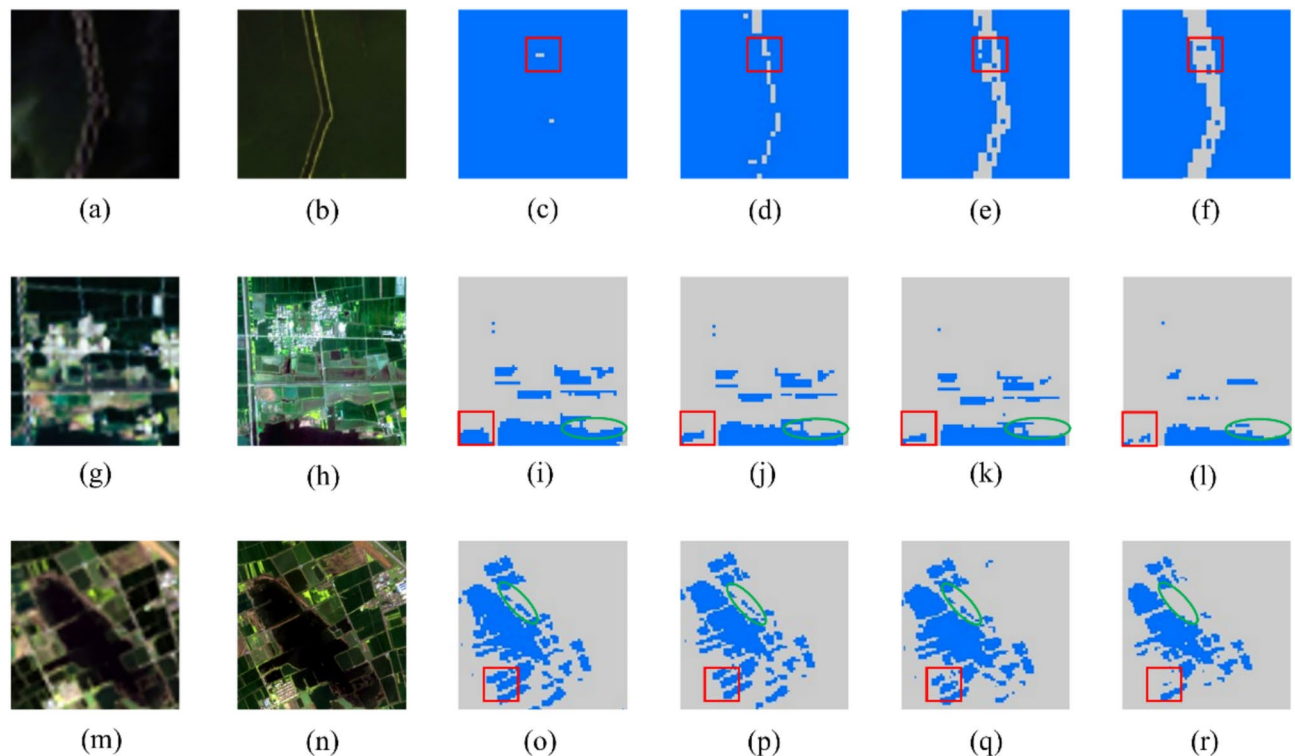


Fig. 8. Comparison of the effectiveness of NDWI, EWI, MNDWI and INDWI in mining water extraction. (a) Landsat OLI false colour composite image (band 432) identified river bridges; (b) Jilin-1 high-resolution image (0.75 m resolution) identified river bridges; (c) Binary image from the NDWI method; (d) Binary image from the EWI method; (e) Binary image from the MNDWI method; (f) Binary image from the INDWI method; (g) Landsat OLI false colour composite image (band 432) identified the edge of the water body in the collapse zone; (h) Jilin-1 high-resolution image (0.75 m resolution) identified the edge of the water body in the collapse zone; (i) Binary image from the NDWI method; (j) Binary image from the EWI method; (k) Binary image from the MNDWI method; (l) Binary image from the INDWI method; (m) Landsat OLI false colour composite image (band 432) identified urban water body; (n) Jilin-1 high-resolution image (0.75 m resolution) identified urban water body; (o) Binary image from the NDWI method; (p) Binary image from the EWI method; (q) Binary image from the MNDWI method; (r) Binary image from the INDWI method.

of information regarding subsided water bodies, thereby facilitating the timely implementation of interventions such as land reclamation and ecological restoration.

Upon the extension of coal mining impacts to the surface, the surface undergoes subsidence from its original elevation, thereby forming a subsided area above the mined-out zone, known as a subsidence basin or surface movement basin. In areas characterized by elevated groundwater levels, even minor surface subsidence can result in water accumulation within the movement basin, leading to a significant reduction in the arable land area within the mining zone³⁸. Ground subsidence inevitably results in alterations to water depth, surface area, and the spatial location of the water body, subsequently affecting water area, reservoir capacity, and flood storage capabilities. Guqiao Coal Mine, situated in a high groundwater level region in eastern China with a shallow groundwater table, experiences the unavoidable formation of subsided areas as coal mining activities expand. The accumulation of water in mining areas due to coal mining subsidence, to a certain extent, enlarges the water area within the mining zone, concurrently instigating a series of ecological and production challenges. Extensive water accumulation inundates crops, leading to the destruction of substantial quantities of high-quality farmland and a dramatic decline in local grain production. Furthermore, this situation deprives numerous farmers of their essential land for livelihood.

Utilizing the INDWI constructed in this study, it was possible, to a certain extent, to accurately extract water bodies in high groundwater level mining areas. The extraction results and spatial distribution information served as foundational data and references, facilitating the subsequent scientific management of high groundwater level mining areas. Ecological restoration and land reclamation are required for the damaged land in mining areas. In 2011, the State Council implemented the Land Reclamation Regulations, aimed at improving the management of land reclamation and standardizing associated activities. These regulations mandated that land reclamation must adhere to principles of scientific planning, suitability to local conditions, comprehensive management, economic feasibility, and rational utilization. Emphasis was placed on the enhancement of mine ecological environment restoration and management, adhering to the principles of "whoever damages, reclaims, mines while restoring, and subsides while managing," thereby ensuring synchronization between mine ecological environment damage and restoration management³⁹. The extraction of thick coal seams constitutes the majority of coal seam mining

in the study area. Since 2007, mining depths have ranged from -550 to -780 m, with cumulative mining thicknesses of 1.29 to 7.88 m. Post-extraction, surface subsidence occurs, resulting in the formation of subsidence areas exhibiting varying degrees of subsidence. Furthermore, owing to the high groundwater level in the study area, the bottoms of these subsidence areas are susceptible to water accumulation, leading to the formation of extensive water-logged areas. For the management of subsidence land in this region, it is recommended to employ a flexible approach by utilizing various subsidence land management techniques to improve the efficacy of such interventions. The specific measures should be tailored to the unique subsidence conditions of each coal mine, adhering to the principle of "agriculture where suitable for agriculture, fishery where suitable for fishery, forestry where suitable for forestry, construction where suitable for construction." This approach ensures the rational utilization of subsidence land in accordance with local conditions. For areas with seasonal water accumulation and shallow subsidence, techniques such as coal gangue filling, ditching and drainage, and cutting high and filling low can be implemented. Subsequently, the leveled land can be utilized for continued cultivation or partially reclaimed as forest land, grassland, etc. In areas with perennial water accumulation, measures such as excavating fish ponds and constructing wetland parks can be adopted.

Limitations and future research

To address the spectral characteristics of easily confusable land types in high groundwater level mining areas, this study introduced a method that integrates the INDWI water extraction index with the Otsu algorithm. When compared to three alternative indices, this integrated approach effectively mitigated the influence of mining area vegetation, pit edges, and coal and mining waste, thereby enhancing the clarity of the water body delineation and the accuracy of the extracted area. Nonetheless, this study exhibits areas requiring enhancement. Primarily, during image selection, due to the influence of factors such as clouds and cloud shadows on some images, it is difficult to ensure that high-quality TM/OLI images acquired during the plant growing season are available for every year. Additionally, there are differences in the actual acquisition times of the images, which can still affect the accuracy of the results to a certain extent^{40,41}. To address this issue, in the future, image accuracy can be improved through multi-source data integration, time series analysis, and cloud detection and repair. Secondly, the extraction of water bodies fails to adequately address the issue of mixed pixels, resulting in a small degree of misclassification and omission in the constructed water body index. Here, the accuracy of water body extraction can be improved through a multi-source data fusion method, utilizing the spatial information of high-resolution images and the spectral information of medium-resolution images^{42,43}. For example, the introduction of DEM terrain data can aid in better handling the issue of water body extraction in areas with undulating terrain⁴⁴. Alternatively, the introduction of deep learning models allows for automatic learning of the differences between water body and non-water body features, thereby more accurately identifying water body information in mixed pixels. Deep learning models can learn complex spectral and spatial characteristics through a large amount of training data, thus improving the accuracy of water body extraction^{45,46}. Lastly, while the GEE platform has automated the extraction of long-term water body sequences in the identification of subsided water bodies, the process of identifying subsided water bodies based on time-series water body results requires further automation. Moreover, the temporal evolution patterns of subsided water bodies could be subjected to additional analysis and validation, including the precise identification of individual subsided water bodies and the observation of the gradual evolution of multiple sub-subsidence water bodies into a single subsided water body.

In recent years, the rapid advancement of high-resolution Earth observation programs has led to the deployment of numerous high-resolution remote sensing satellites, including ESA's Sentinel series and China's Gaofen series. The imagery from these satellites boasts ground resolutions at the decimeter and meter scales, with some even achieving sub-meter resolution at nadir, exemplified by the Jilin-1 and Gaojing-1 satellite imagery. This level of detail significantly aids in the precise delineation of water body boundaries^{47,48}. The formation of subsided water bodies progresses from small to large. Over time, high-resolution remote sensing imagery develops into a time series. In future research, employing higher-resolution remote sensing imagery for the extraction of high-resolution water bodies in a time-series context will facilitate a more detailed portrayal of the formation process of subsided water bodies, thereby enhancing the precision of their identification.

Conclusions

In mining areas with high groundwater levels, coal mining subsidence causes surface sinking, leading to the formation of waterlogged collapse zones. These zones inundate large areas of arable land and trigger ecological and environmental issues. To address the challenges of accurately extracting water body information in mining areas with high groundwater levels, an innovative water body index, INDWI, was developed using the OTSU algorithm.

In this study, the influence of extensive vegetation along the edges of waterlogged areas, as well as the presence of large quantities of fly ash and coal gangue resulting from coal mining activities, was considered during water body extraction. Through comparative experiments involving various indices and enhancement functions, the SWIR band was incorporated, and the natural logarithm (ln) function was introduced to improve the MNDWI index, leading to the development of the INDWI index. Compared to NDWI, EWI, and MNDWI, INDWI demonstrated a superior ability to clearly delineate the edges of subsidence water bodies in mining areas, reducing misclassification and omission errors. The index effectively minimized the impact of vegetation near water bodies, resulting in more accurate classification outcomes, with an overall accuracy exceeding 89% and a Kappa coefficient above 80%. This study aims to provide decision support for the sustainable development of mining areas with high groundwater levels, the efficient utilization of water and soil resources, and ecological restoration efforts.

Data availability

Correspondence and requests for materials should be addressed to A.Z.

Received: 25 June 2024; Accepted: 1 April 2025

Published online: 09 April 2025

References

1. Lei, K., Pan, H. Y. & Lin, C. Y. A landscape approach towards ecological restoration and sustainable development of mining areas. *Ecol. Eng.* **90**, 320–325. <https://doi.org/10.1016/j.ecoleng.2016.01.080> (2016).
2. Xu, H. L. et al. A systematic review and comprehensive analysis on ecological restoration of mining areas in the arid region of China: Challenge, capability and reconsideration. *Ecol. Indic.* **154**, 110630. <https://doi.org/10.1016/j.ecolind.2023.110630> (2023).
3. Feng, Z. J. et al. Improving mine reclamation efficiency for farmland sustainable use: Insights from optimizing mining scheme. *J. Clean Prod.* **379**, 134615. <https://doi.org/10.1016/j.jclepro.2022.134615> (2022).
4. Feng, Z. J. et al. Integrated mining and reclamation practices enhance sustainable land use: A case study in Huainan coalfield, China. *Land* **12**, 1994. <https://doi.org/10.3390/land12111994> (2023).
5. Li, G. S. et al. A new approach to increased land reclamation rate in a coal mining subsidence area: A case-study of Guqiao Coal Mine, China. *Land Degrad. Dev.* **33**, 866–880. <https://doi.org/10.1002/ldr.4184> (2022).
6. Xiao, W., Hu, Z. Q., Chugh, Y. P. & Zhao, Y. L. Dynamic subsidence simulation and topsoil removal strategy in high groundwater table and underground coal mining area: A case study in Shandong Province. *Int. J. Min. Reclam. Environ.* **28**, 250–263. <https://doi.org/10.1080/17480930.2013.828457> (2014).
7. Farhadi, H., Ebadi, H., Kiani, A. & Asgari, A. Near real-time flood monitoring using multi-sensor optical imagery and machine learning by GEE: An automatic feature-based multi-class classification approach. *Remote Sens.* **16**, 4454. <https://doi.org/10.3390/rs16234454> (2024).
8. Sagan, V. et al. Monitoring inland water quality using remote sensing: potential and limitations of spectral indices, bio-optical simulations, machine learning, and cloud computing. *Earth-Sci. Rev.* **205**, 103187. <https://doi.org/10.1016/j.earscirev.2020.103187> (2020).
9. Chen, Y., Fan, R., Yang, X., Wang, J. & Latif, A. Extraction of urban water bodies from high-resolution remote-sensing imagery using deep learning. *Water* **10**, 585. <https://doi.org/10.3390/w10050585> (2018).
10. Zhao, B. et al. An improved surface water extraction method by integrating multi-type priori information from remote sensing. *Int. J. Appl. Earth Obs. Geoinf.* **124**, 103529. <https://doi.org/10.1016/j.jag.2023.103529> (2023).
11. Xu, Y., Lin, J., Zhao, J. & Zhu, X. New method improves extraction accuracy of lake water bodies in Central Asia. *J. Hydrol.* **603**, 127180. <https://doi.org/10.1016/j.jhydrol.2021.127180> (2021).
12. Nagaraj, R. & Kumar, L. S. Extraction of surface water bodies using optical remote sensing images: A review. *Earth Sci. Inform.* **17**, 893–956. <https://doi.org/10.1007/s12145-023-01196-0> (2024).
13. Rajeswari, S. & Rathika, P. Emerging methodologies in waterbody delineation: an In-depth review. *Int. J. Remote Sens.* **45**, 5789–5819. <https://doi.org/10.1080/01431161.2024.2379518> (2024).
14. Farhadi, H., Ebadi, H., Kiani, A. & Asgari, A. Introducing a new index for flood mapping using sentinel-2 imagery (SFMI). *Comput. Geosci.* **194**, 105742. <https://doi.org/10.1016/j.cageo.2024.105742> (2025).
15. Farhadi, H., Ebadi, H., Kiani, A. & Asgari, A. A novel flood/water extraction index (FWEI) for identifying water and flooded areas using sentinel-2 visible and near-infrared spectral bands. *Stoch. Environ. Res. Risk Assess* **38**, 1873–1895. <https://doi.org/10.1007/s00477-024-02660-z> (2024).
16. Mcfeeters, S. K. The use of the normalized difference water index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* **17**, 1425–1432. <https://doi.org/10.1080/01431169608948714> (1996).
17. Xu, H. Q. A study on the extraction of water body information using the modified normalised difference water body index (MNDWI). *J. Remote Sens.* <https://doi.org/10.11834/jrs.20050586> (2005).
18. Feyisa, G. L., Meilby, H., Fensholt, R. & Proud, S. R. Automated water extraction index: A new technique for surface water mapping using Landsat imagery. *Remote Sens. Environ.* **140**, 23–35. <https://doi.org/10.1016/j.rse.2013.08.029> (2014).
19. Jiang, W. et al. An effective water body extraction method with new water index for sentinel-2 imagery. *Water* **13**, 1647. <https://doi.org/10.3390/w13121647> (2021).
20. He, T. et al. Continues monitoring of subsidence water in mining area from the eastern plain in China from 1986 to 2018 using Landsat imagery and Google Earth Engine. *J. Clean Prod.* **279**, 123610. <https://doi.org/10.1016/j.jclepro.2020.123610> (2021).
21. Hu, X., Li, X., Min, X. & Niu, B. Optimal scale extraction of farmland in coal mining areas with high groundwater levels based on visible light images from an unmanned aerial vehicle (UAV). *Earth Sci. Inform.* **13**, 1151–1162. <https://doi.org/10.1007/s12145-020-00493-2> (2020).
22. Peng, S. P. et al. Application of remote sensing technology in dynamic monitoring of waterlogging and subsidence in coal mining areas—a case study of Huainan mining area. *J. China Coal Soc.* **04**, 374–378. <https://doi.org/10.3321/j.issn:0253-9993.2002.04.009> (2002).
23. Li, J. et al. Research on the identification method of water bodies in Jining City based on LandsatTM data. *Environ. Sci. Technol.* **36**, 175–179. <https://doi.org/10.3969/j.issn.1003-6504.2013.09.037> (2013).
24. Wang, Q. Y., Xiao, W., Li, S. C., Zhang, W. K. & Zhang, W. Dynamic monitoring of coal mining subsidence waters in Panxie mining area of Huainan Province based on time-series TM images. *China Coal* **44**, 37–43. <https://doi.org/10.19880/j.cnki.ccm.2018.07.007> (2018).
25. Nie, X., Hu, Z., Ruan, M., Zhu, Q. & Sun, H. Remote-sensing evaluation and temporal and spatial change detection of ecological environment quality in coal-mining areas. *Remote Sens.* **14**, 345. <https://doi.org/10.3390/rs14020345> (2022).
26. Bhunia, G. S. Assessment of automatic extraction of surface water dynamism using multi-temporal satellite data. *Earth Sci. Inform.* **14**, 1433–1446. <https://doi.org/10.1007/s12145-021-00612-7> (2021).
27. Lai, Y., Zhang, J., Song, Y. & Cao, Y. Comparative analysis of different methods for extracting water body area of Miyun reservoir and driving forces for nearly 40 years. *J. Indian Soc. Remote Sens.* **48**, 451–463. <https://doi.org/10.1007/s12524-019-01076-5> (2020).
28. Malahlela, O. E. Inland waterbody mapping: Towards improving discrimination and extraction of inland surface water features. *Int. J. Remote Sens.* **37**, 4574–4589. <https://doi.org/10.1080/01431161.2016.1217441> (2016).
29. Sun, F., Sun, W., Chen, J. & Gong, P. Comparison and improvement of methods for identifying waterbodies in remotely sensed imagery. *Int. J. Remote Sens.* **33**, 6854–6875. <https://doi.org/10.1080/01431161.2012.692829> (2012).
30. Sun, G. Y. et al. Research on water body index in the Yellow River Basin based on GEE platform. *Jen Min Huang Ho* **45**, 119–124. <https://doi.org/10.3969/j.issn.1000-1379.2023.03.022> (2023).
31. Pan, F., Xi, X. & Wang, C. A comparative study of water indices and image classification algorithms for mapping inland surface water bodies using Landsat imagery. *Remote Sens.* **12**, 1611. <https://doi.org/10.3390/rs12101611> (2020).
32. Yang, K. et al. River delineation from remotely sensed imagery using a multi-scale classification approach. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **7**, 4726–4737. <https://doi.org/10.1109/jstars.2014.2309707> (2014).
33. Rad, A. M., Kreitler, J. & Sadeh, M. Augmented normalized difference water index for improved surface water monitoring. *Environ. Modell. Softw.* **140**, 105030. <https://doi.org/10.1016/j.envsoft.2021.105030> (2021).

34. Liu, S., Wu, Y., Zhang, G., Lin, N. & Liu, Z. Comparing water indices for Landsat data for automated surface water body extraction under complex ground background: A case study in Jilin Province. *Remote Sens.* **15**, 1678. <https://doi.org/10.3390/rs15061678> (2023).
35. Xie, H., Luo, X., Xu, X., Pan, H. & Tong, X. Evaluation of Landsat 8 OLI imagery for unsupervised inland water extraction. *Int. J. Remote Sens.* **37**, 1826–1844. <https://doi.org/10.1080/01431161.2016.1168948> (2016).
36. Hu, Z. Q., Xiao, W. & Zhao, Y. L. Revisiting the ecological environment of coal mining area “restoration while mining”. *J. China Coal Soc.* **45**, 351–359. <https://doi.org/10.13225/j.cnki.jccs.YG19.1694> (2020).
37. Hu, Z. et al. Coupling of underground coal mining and mine reclamation for farmland protection and sustainable mining. *Resour. Policy* **84**, 103756. <https://doi.org/10.1016/j.resourpol.2023.103756> (2023).
38. Li, G. et al. Innovation for sustainable mining: Integrated planning of underground coal mining and mine reclamation. *J. Clean Prod.* **351**, 131522. <https://doi.org/10.1016/j.jclepro.2022.131522> (2022).
39. Hu, Z. & Xiao, W. Optimization of concurrent mining and reclamation plans for single coal seam: A case study in northern Anhui, China. *Environ. Earth Sci.* **68**, 1247–1254. <https://doi.org/10.1007/s12665-012-1822-9> (2013).
40. Hao, Q., Zheng, W. & Xiao, Y. Fusion information multi-view classification method for remote sensing cloud detection. *Appl. Sci.* **12**, 7295. <https://doi.org/10.3390/app12147295> (2022).
41. Xia, M. & Jia, K. Reconstructing missing information of remote sensing data contaminated by large and thick clouds based on an improved multitemporal dictionary learning method. *IEEE Trans. Geosci. Remote Sens.* **60**, 5605914. <https://doi.org/10.1109/tgrs.2021.3095067> (2022).
42. Li, M. et al. A deep learning method of water body extraction from high resolution remote sensing images with multisensors. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **14**, 3120–3132. <https://doi.org/10.1109/jstars.2021.3060769> (2021).
43. Yuan, C., Wang, F., Wang, S. & Zhou, Y. Accuracy evaluation of flood monitoring based on multiscale remote sensing for different landscapes. *Geomat. Nat. Hazards Risk* **10**, 1389–1411. <https://doi.org/10.1080/19475705.2019.1580224> (2019).
44. Sun, Q. & Li, J. A method for extracting small water bodies based on DEM and remote sensing images. *Sci. Rep.* **14**, 760. <https://doi.org/10.1038/s41598-024-51346-7> (2024).
45. Liu, J. & Wang, Y. Water body extraction in remote sensing imagery using domain adaptation-based network embedding selective self-attention and multi-scale feature fusion. *Remote Sens.* **14**, 3538. <https://doi.org/10.3390/rs14153538> (2022).
46. Yu, Y. et al. A self-attention capsule feature pyramid network for water body extraction from remote sensing imagery. *Int. J. Remote Sens.* **42**, 1801–1822. <https://doi.org/10.1080/01431161.2020.1842544> (2021).
47. Chen, J. et al. Remote sensing big data for water environment monitoring: current status, challenges, and future prospects. *Earth Future* **10**, 2289. <https://doi.org/10.1029/2021EF002289> (2022).
48. Zhao, S. et al. An overview of satellite remote sensing technology used in China’s environmental protection. *Earth Sci. Inform.* **10**, 137–148. <https://doi.org/10.1007/s12145-017-0286-6> (2017).

Acknowledgements

This study was funded by the Beijing Business Environment Reform and Support Program in the field of ecology and environment (2241STC60470).

Author contributions

A.Z. and Z.W. and Y. G. wrote the main manuscript text. All authors reviewed the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to A.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025