Article

Object-Based Mapping of Gullies Using Optical Images: A Case Study in the Black Soil Region, Northeast of China

Biwei Wang [1,2], Zengxiang Zhang [1], Xiao Wang [1,*] , Xiaoli Zhao [1], Ling Yi [1] and Shunguang Hu [1]

[1]   Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100101, China;
     wangbw@radi.ac.cn (B.W.); zhangzx@radi.ac.cn (Z.Z.); zhaoxl@radi.ac.cn (X.Z.); Yiling@radi.ac.cn (L.Y.);
     husg@radi.ac.cn (S.H.)
[2]   University of Chinese Academy of Sciences, Beijing 100049, China
*   Correspondence: wangxiao98@radi.ac.cn

check for updates

Abstract: Gully erosion is a widespread natural hazard. Gully mapping is critical to erosion monitoring and the control of degraded areas. The analysis of high-resolution remote sensing images (HRI) and terrain data mixed with developed object-based methods and field verification has been certified as a good solution for automatic gully mapping. Considering the availability of data, we used only open-source optical images (Google Earth images) to identify gully erosion through image feature modeling based on OBIA (Object-Based Image Analysis) in this paper. A two-end extrusion method using the optimal machine learning algorithm (Light Gradient Boosting Machine (LightGBM)) and eCognition software was applied for the automatic extraction of gullies at a regional scale in the black soil region of Northeast China. Due to the characteristics of optical images and the design of the method, unmanaged gullies and gullies harnessed in non-forest areas were the objects of extraction. Moderate success was achieved in the absence of terrain data. According to independent validation, the true overestimation ranged from 20% to 30% and was mainly caused by land use types with high erosion risks, such as bare land and farm lanes being falsely classified as gullies. An underestimation of less than 40% was adjacent to the correctly extracted gullied areas. The results of extraction in regions with geographical object categories of a low complexity were usually more satisfactory. The overall performance demonstrates that the present method is feasible for gully mapping at a regional scale, with high automation, low cost, and acceptable accuracy.

Keywords: gully; google earth images; object-based image analysis; machine learning methods; the black soil region of northeast China

1. Introduction

Gully erosion, one type of soil erosion, occurs due to the removal of the top soil along the drainage channels through surface water runoff [1]. Gully erosion is recognized as a key indicator of land degradation [2], which threatens agricultural production, economic development, and the ecological environment [3–5]. Identifying gully erosion and vulnerable areas is thus of great importance to soil and water conservation and land use planning and concerns scientists, farmers, and land managers [6–8].

Field investigation methods were initially used to map gullies. However, these methods are time-consuming and laborious, making them infeasible in large areas. With more readily available remote sensing images and the development of GIS, image-based interpretation methods have been widely employed. There are two directions for these techniques: One is visual interpretation, and the other is automatic extraction. Visual interpretation relies on the image's characteristics and expert knowledge to extract gullies [9–12], so low efficiency and the subjectivity of manual work

remains problems for this method. Several studies have attempted to develop automatic methods to map gullies with good productivity and repeatability. Pixel-based image analysis of spectral similarities was studied first, but this method encountered the inevitable problem of spectrally similar false positives [13–16]. Compared to pixel-based methods, object-based image analysis (OBIA) can capture more high-resolution images because it employs characteristic textures and geometry and an optimized feature space [17,18]. A number of researchers have used OBIA to map gully erosion by high-resolution remote sensing images (HRI), with the auxiliaries of ground information and expert knowledge [19–21]. However, the establishment of extraction rules, such as the selection of object features and the determination of thresholds, can only be done in a semi-automatic way, so these methods retain a certain amount of subjectivity.

With the development of computational statistical models, recent work on the combination of object-oriented analysis and the Random Forest (RF) algorithm has improved the degree of automation and objectivity to a large extent [22–25]. The RF based on bootstrap aggregation and ensembles of classification trees is a well-accepted performance optimization algorithm. Compared to other traditional models, such as Regression Tree Analysis (RTA), Support Vector Machine (SVM), and Nave Bayes (NB), RF has better abilities for feature mapping [13,26,27]. However, some algorithms developed in the last few years, including extreme gradient boosting (Xgboost) and the Light Gradient Boosting Machine (LightGBM) [28,29], have been shown to outperform popular machine learning algorithms like RF in some scientific fields [30,31]. However, the potential of these algorithms to map gullies from HRI has been rarely explored.

In terms of the data used for gully extraction, most studies employ integrated images with terrain information based on OBIA [14,24,25] to ensure a strong connection between terrain features and gully distribution [32,33]. However, terrain data are not readily available for a broad range of users. Even though some open-source and downloadable DEM data are available worldwide, such as ALOS DEM (12.5 m resolution), ASTER DEM (the highest resolution 30m), and SRTM DEM (the highest resolution 90 m), the DEM data with a coarse resolution remain too inaccurate for gully extraction. Some studies have attempted to solve this problem by obtaining a terrain skeleton from ASTER DEM [23]. However, all these DEM data were produced before 2010, indicating that they do not reflect the current situation and changes in gully erosion. Research has begun to extract gullies in the absence of topographic data [19], but there are still some problems that need further study, such as insufficient accuracy and poor automation.

In addition, when HRI is used for image data (from, for example, QuickBird, WorldView, and GeoEye satellite images), it is generally expensive and frequently unavailable. Google Earth (GE) can provide global high spatial resolution images for free, allowing its successful application in the field of geoscience [34,35]. One must, however, consider the absence of near-infrared (NIR) bands, which are usually required to calculate vegetation indexes to distinguish between plant biomass and soil or residue backgrounds. Excess Green minus Excess Red (VI' or ExG–ExR), based on the RGB band, has been proven to be effective in extracting plant targets, especially fresh wheat straw backgrounds [36]. Thus, GE images are suggested to offer a promising alternative to map gullies and should be applied more commonly.

The extraction of gullies has been studied all over the world, and the characteristics of gullies vary from place to place. In the black soil region of Northeast China, gully erosion is widely distributed and common, according to the first national water conservancy census [37]. Studies have demonstrated that gullies mainly occur on sloping farmland [38]. Since the development of the black soil region mainly occurred over the past hundred years, most gullies were formed in a short time with a small width and depth ranging from meters to tens of meters [39]. These gullies, then, have had a very fast development speed [40]. At present, most of the determination of gully erosion in this region still relies on visual interpretation, which is laborious, time-consuming, and cannot meet the needs for the timely monitoring of gully erosion [12,41]. Hence, automatic or semi-automatic gully extraction methods applicable to the black soil region of Northeast China are required.

Considering automation, data availability for users, and the performance of extraction, the main objective of this study is to pursue a superior machine learning method with OBIA for the extraction of gully erosion using GE images with no terrain data in the black soil region of Northeast China. Addressing these objectives requires the following: (1) Deciding the optimal segmentation parameters for OBIA, (2) selecting the optimal machine learning method and revising the preliminary results in the absence of terrain data, (3) and validating the extraction by comparing it with the manually digitized gully system.

2. Study Area and Data

2.1. Study Area

The black soil region is a significant production base for commodity grain in China. However, soil erosion has become increasingly more serious in recent years, especially gully erosion, which disconnects and devours farmland, threatening food security [42]. The study area lies in the west of Baiquan County in the black soil region of Northeast China (Figure 1), including one training area (2.66 km^2), one testing area (0.95 km^2), and three validation areas (2.69 km^2, 1.10 km^2 and 2.00 km^2), each of which is a complete watershed response system. In the study area, several common stratigraphic compositions in the black soil area can be seen. Black soil, chernozem, and meadow soil form the main components of the substrate, and other dark brown soil is sporadically distributed. Its topographic characteristics feature long and gentle slopes. Influenced by its topographic factors, the water system of this area is relatively developed, with a considerable number of tributaries. Regarded as the earliest part of the water system to develop, gullies are widely distributed. The annual rainfall ranges from 450 mm to 790 mm (mainly occurring in June to September), and the annual average runoff depth is about 70 mm. Affected by human activities, cultivated land is the major land use type for maize, soybean, and rice cultivation, and natural forest land is only distributed in areas with steep slopes that are unsuitable for crop growth.

The natural environment of the study area is typical for the black soil region of northeast China. As the key area of gully erosion control, the density of the gullies in this area is high, showing a significant increase [43]. Moreover, the research data on gully erosion in this area are sufficient, which helped ensure a smooth study [44–46]. Thus, it was necessary, representative, and feasible to study the extraction of gully erosion in the chosen study area.

2.2. Data

2.2.1. Image Data

In the study area, GE images are available with sub-meter and meter resolution acquired at different times. Studies on the identification of gully erosion conducted in an environment similar to the black soil region of Northeast China indicated that late spring, autumn, and early summer images were the best choices, as they experienced minimal effects from vegetation cover or snowfall [47]. Another point to consider is spatial resolution. The higher the spatial resolution, the greater the probability that gully erosion will appear as pure pixels on the images and the clearer the outline, shape, and internal details will be. On the contrary, an excessively large spatial resolution makes the internal details overly detailed, which increases the noise information for object extraction [48]. Thus, we used GE images of a 1.19 m resolution acquired on 9 May 2018 (downloaded with the software Rivermap) to extract the gullies.
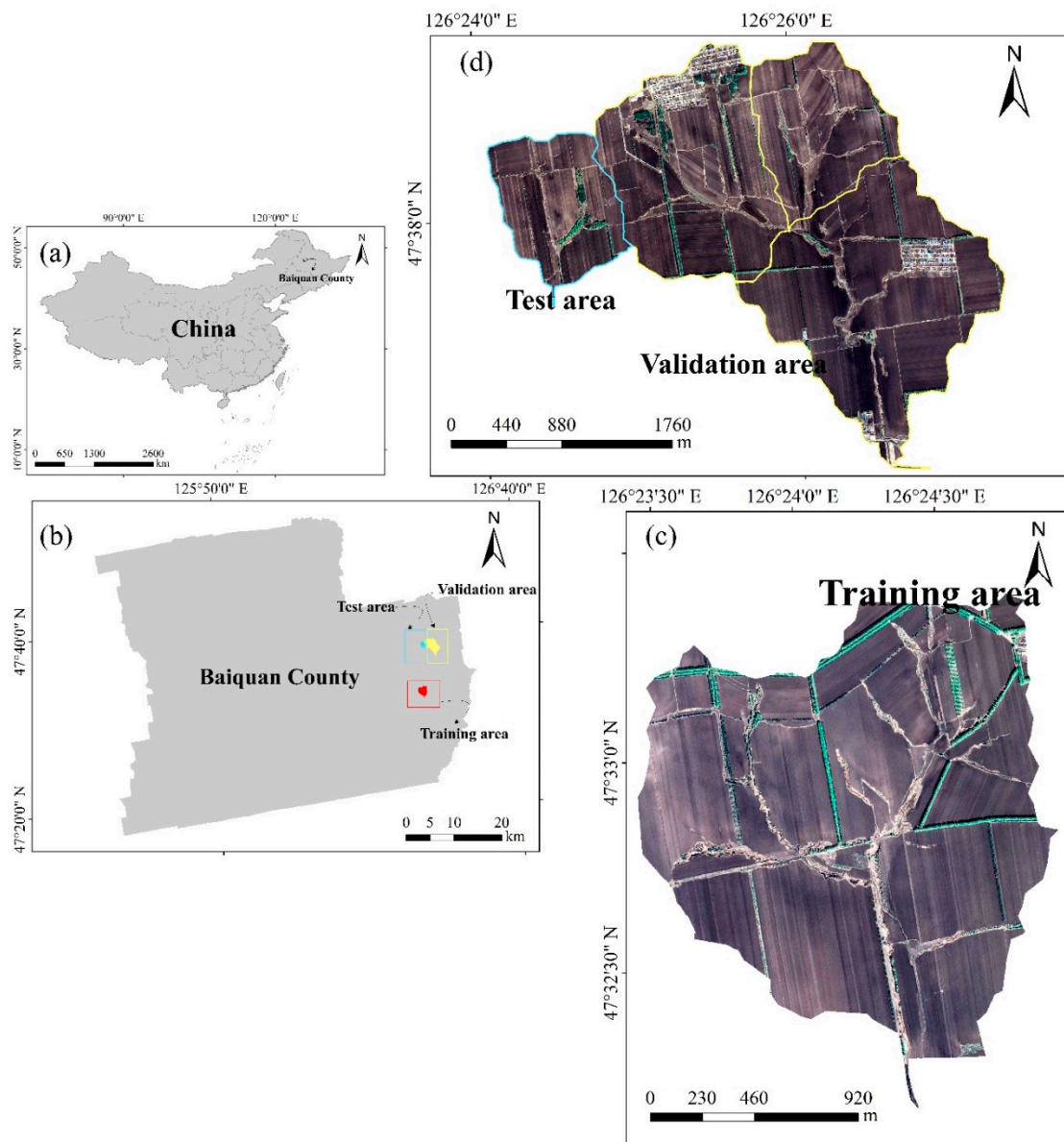
Figure 1. Location of the study area. (**a**) China. (**b**) Baiquan County. The areas of training, testing, and independent validation are shown in red, blue, and yellow rectangles. (**c**) Training area. (**d**) Test area and validation areas.

2.2.2. Reference Data from the Field Survey

The study sites were investigated during two field trips that lasted one week each for the establishment of interpretation marks and the verification of polygons. During the field investigations, the Global Positioning System (GPS) was the main instrument used to precisely locate gully erosion. Since ephemeral gullies with small widths can be easily filled with normal tillage, we did not work on them. Therefore, the 1.19 m spatial resolution images were reasonable for gully mapping. In addition, as the penetration ability of the optical images is weak, the gully erosion in forest areas was beyond the scope of this study. Based on field investigations and expert knowledge, the manually polygonised reference datasets of gully erosion were drawn for training and validating the gully extraction model within the ArcGIS software. The reference data for training are shown below as an example (Figure 2).
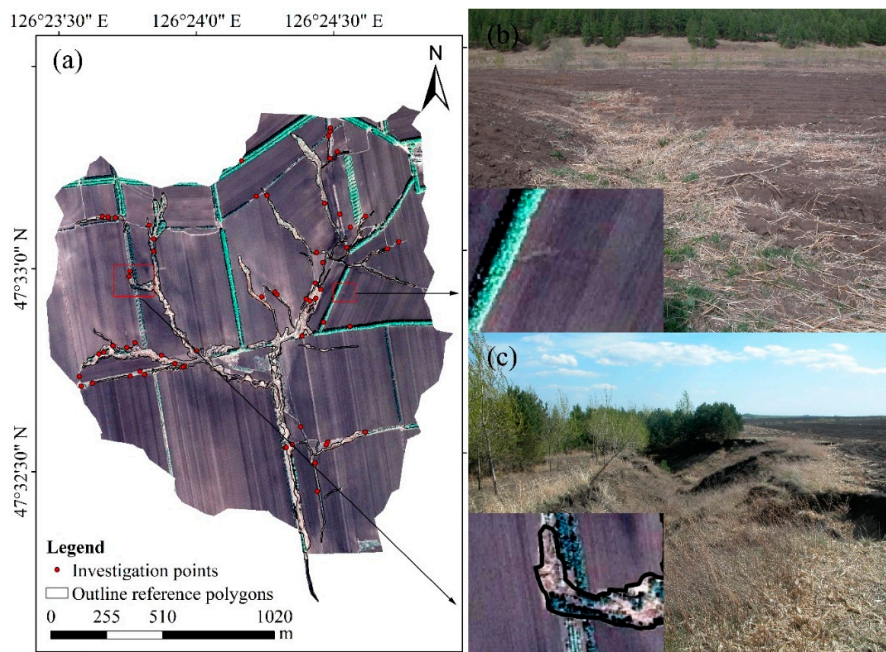
Figure 2. Reference data from the field survey for training. (**a**) Investigation points and manually delineated reference polygons of gully erosion for training. (**b**) An ephemeral gully that was not considered in this study. (**c**) An example of a gully that was extracted in this study.

3. Method

The presented method for the extraction of gully erosion was implemented by employing OBIA and the optimal method based on machine learning algorithms on GE images. Figure 3 shows a survey of the workflow, which can be summed up as follows: (1) Creating the sample dataset within the eCognition and ArcGIS software, (2) determining the optimal machine learning method for preliminary extraction of gullies, (3) revising the extraction results using eCognition, and (4) validating the method presented. These steps are explained in further detail below.
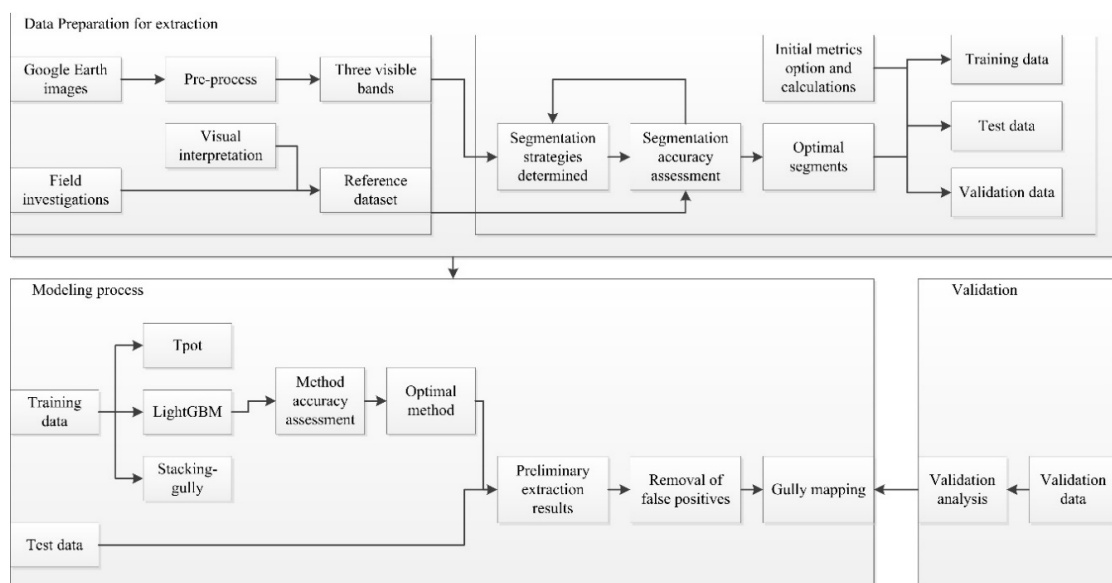

Figure 3. Overview of the research method.

3.1. Data Preparation

3.1.1. GE Image Segmentation

For the initial step, we segmented GE images to derive objects in eCognition for the subsequent analysis; this step is a vital part of OBIA processing. The multiresolution segmentation method, an optimal method for high-quality multi-scale image segmentation [49], was chosen to achieve segmentation in this study. This process began with an individual pixel and further amalgamated the most similar neighbor pixel by a minimum heterogeneity increment until the heterogeneity of the generated object exceeded that of the preset threshold. Three visible bands (RGB) were involved in the segmentation process, each weighted with a value of 1.

There were three input parameters: Scale, shape, and compactness. Segmentation parameter determination was an iterative process. In the first iteration, the ranges of values for shape and compactness were 0 to 0.9 and 0 to 1.0, respectively. To obtain the appropriate ranges of shape and compactness, a variable-controlling approach was adopted. In this process, the scale parameter was temporarily set to 50 by visual assessment. After determining these ranges, the optimal combination of both was determined. The appropriate scale parameter, a key value that heavily affects the emergence of objects and facilitates further analysis [50], was explored with the help of the Estimation of Scale Parameter 2 (ESP2), proposed by Drăguţ et al. [51]. Regarded as a credible tool, the ESP2 can determine suitable scale parameters based on an object's shape, compactness, and other parameters (e.g., the number of loops and starting scales), which are input based on Local Variance (LV) values [52]. The scale parameter obtained in the first iteration was used as the initial parameter in the second iteration to find the appropriate combination of shape and compactness within their respective ranges obtained by the first iteration. Iterations continued until the results no longer changed.

In order to evaluate the goodness of the image segmentation's potential parameter combinations, over-segmentation (OS), under-segmentation (US), and Euclidian distance (ED) were computed, as follows [53]:

$$OS = 1 - \frac{\sum |r_i \cap s_k|}{\sum |r_i|} \tag{1}$$

$$US = 1 - \frac{\sum |r_i \cap s_k|}{\sum |s_k|} \tag{2}$$

$$ED = \left(\frac{OS^2 + US^2}{2} \right)^{1/2} \tag{3}$$

where r_i is part of the reference data, and s_k is one of the corresponding segmentation objects (a segment that overlaps r_i by no less than 50% is considered as s_k, which can provide valuable information for training the subsequent extraction models). OS and US indicate how well the segments match the reference data. As a composite index, ED is explained by its proximity to perfect segmentation. Ideally, the value of each index will be 0.

3.1.2. Object-Based Explanatory Features Generation

In this study, the spectral, textural, and geometrical features of the segments were selected and used for gully mapping, as shown in Table 1. To obtain as many variables as possible, 37 variables were derived within eCognition. The spectral features comprised the mean band value, band ratios, mean brightness, maximum difference index, standard deviation of the band, and VI'. Among them, the band ratios and VI' were customized features. The calculation formulas are as follows [36]:

$$\text{ratio\_RG} = \frac{R}{G} \tag{4}$$

$$\text{ratio\_RB} = \frac{R}{B} \tag{5}$$

$$\text{ratio\_GB} = \frac{G}{B} \tag{6}$$

$$\text{VI}' = (2 \times G' - R' - B') - (1.4 \times R' - G') \tag{7}$$

$$\text{R}' = \frac{R}{R + G + B} \tag{8}$$

$$\text{G}' = \frac{G}{R + G + B} \tag{9}$$

$$\text{B}' = \frac{B}{R + G + B} \tag{10}$$

where R, G, and B are the mean values of the red, green, and blue bands, respectively.

Table 1. Data source, object features, and 37 relevant variables.

Image Data	Object Features	Variables Extracted
	Spectral features	Red, green, blue, ratio_RG, ratio_RB, ratio_GB, brightness, max. diff, standard deviation of red; green; blue, and VI';
GE images	Textural features	GLCM homogeneity, contrast, mean, stdDev, correlation, dissimilarity, entropy, and ang. 2nd moment;
	Geometrical features	Area, length, width, length/width, border length, rel. border to image border, border index, asymmetry, compactness, roundness, density, main direction, elliptic fit, rectangular fit, shape index, radius of largest enclosed ellipse, and radius of smallest enclosing ellipse.

The built-in geometrical features for planar objects in eCognition were all selected, including the area, length, width, length/width, border length, rel. border to image border, border index, asymmetry, compactness, roundness, density, main direction, elliptic fit, rectangular fit, shape index, radius of largest enclosed ellipse, and radius of smallest enclosing ellipse.

The image texture expresses the connection of the adjacent pixels, which has been used to improve the classification accuracy [54]. Eight texture features based on the Grey Level Co-occurrence Matrix (GLCM), which is a tabulation illustrating the frequency of various combinations of pixel brightness values (gray levels) on the image [55], were computed for all directions using eCognition. Thus, the training, testing, and validation data were generated.

Acronyms: R: Red, G: Green, B: Blue, max. diff: Maximum difference index, ang. 2nd moment: Angular second moment, stdDev: Standard deviation, rel. border to image border: Relative border to image border.

3.2. Gully Mapping Using Machine Learning Methods

In this section, an optimal method based on machine learning algorithms is used for gully mapping, and the evaluation indexes, precision, recall, and F-score are introduced. These factors all have an important role to play in the validation of the results, as detailed in Section 3.4. We selected as many features as possible, but an overabundance of features may have caused data redundancy and noise interference. Thus, the importance of the features was explored in the machine learning model mentioned below.

Compared to machine learning methods, the popular methods for deep learning require large datasets (with millions of scales) to achieve high performance. Moreover, these methods are too complex to establish a model with good properties, so this study did not consider deep learning methods.

3.2.1. Tree-Based Pipeline Optimization Tool (TPOT)

TPOT was implemented to determine the optimum method simply and rapidly [56,57]. TPOT is a tool for automated machine learning in Python, which intelligently determines the best pipeline for

data among thousands of possible pipelines using genetic programming (Figure 4). Multiple machine learning methods can be integrated into one pipeline by taking into account the data preprocessing steps (normalization, missing value imputation, transformation, scaling, etc.), the hyper parameter settings, and the ways to stack or ensemble the methods within the pipeline. However, although TPOT can function normally with almost all machine learning methods, it cannot function with LightGBM, which is a fairly efficient method.
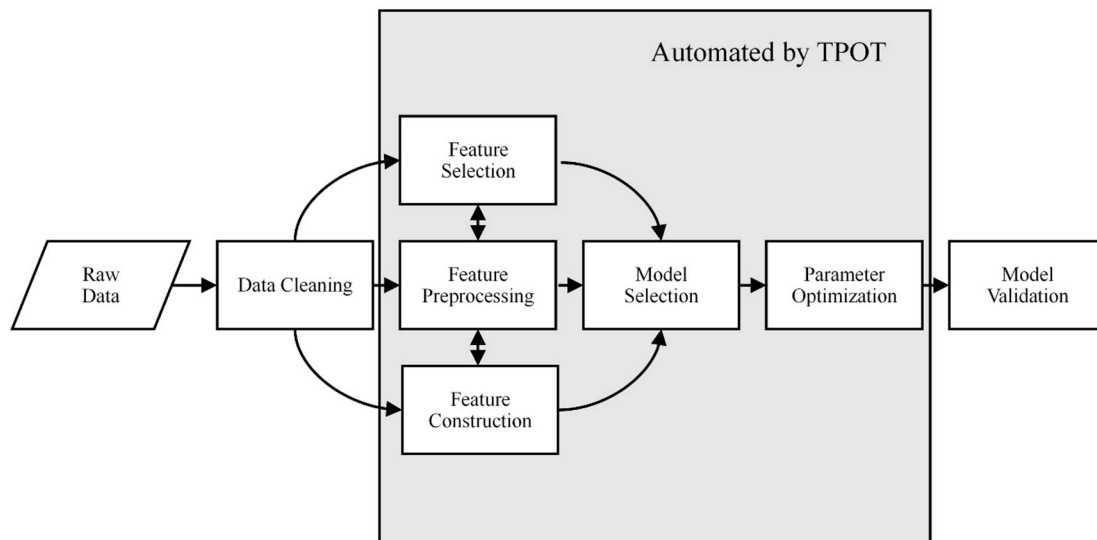

Figure 4. Illustration of Tree-Based Pipeline Optimization Tool (TPOT).

TPOT needs to run for enough time to achieve convergence. In this study, we installed XGBoost to supplement TPOT and ran TPOT until the same pipeline was recommended for the same dataset by two TPOT runs.

3.2.2. LightGBM

For the above-mentioned reasons, LightGBM alone was analyzed. Using tree-based learning methods, LightGBM offers a gradient boosting framework that includes two novel techniques: Gradient-based One-Side Sampling (GOSS) and Exclusive Feature Bundling (EFB) [28]. Due to the elimination of a high percentage of data instances with few gradients using GOSS, GOSS only employs the remainder to assess the information gain. With EFB, mutually exclusive features are packed so that the number of features decreases. The implementation of GOSS and EFB increased the speed of the method by dozens of times. To achieve accuracy optimization, LightGBM grows trees by a leaf-wise (best-first) approach that grows leaves with a max delta loss (Figure 5) [58]. In addition, considering the possible leaf-wise over-fitting phenomenon with a small dataset, LightGBM restricts tree depth using a parameter. However, the trees still grow leaf-wise even when this parameter is identified.

We used the LightGBM Python-package established by Microsoft Corporation for statistical computing to associate gully presence with explanatory variables. Parametric adjustment was the most important step affecting accuracy. A large max_bin, small learning_rate, and large num_leaves were possibilities considered to improve accuracy.
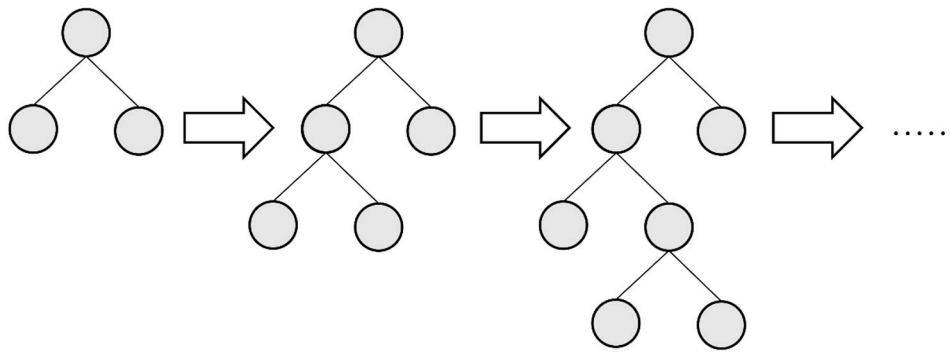
Leaf-wise tree growth

Figure 5. Illustration of a leaf-wise tree.

3.2.3. Stacking

Considering multiple machine learning methods, including LightGBM, stacking was applied. The idea of this approach is derived from stackedgeneralization, proposed by Wolpert [59]. Stacking is a hierarchical fusion model (Figure 6). For example, when we fuse three machine learning models trained with different characteristics, we use them as the basic models. Then, we train a sub-learner on the basic models; the sub-learner is used to organize the basic learner's predictions or the predictions of the basic models as input, so the sub-learner learning organization assigns weight to the predictions of the basic models to obtain the final prediction. The following points should be noted in the selection of each model: First, basic models should be strong models, while the sub-learner of the second layer may use a simple classifier to prevent over-fitting. Second, the number of the basic models can be neither too small nor too big. Third, the basic models need to be accurate and have little correlation.
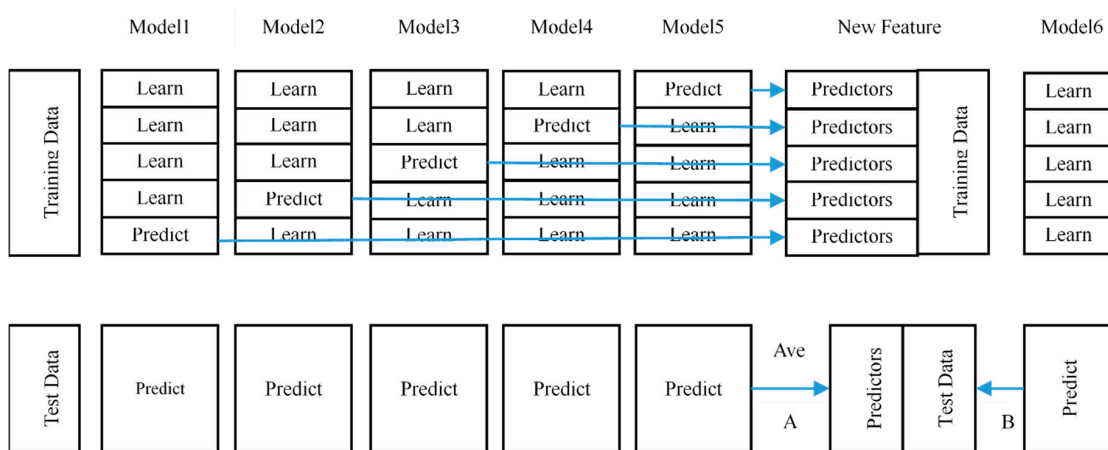

Figure 6. Illustration of stacking.

On the basis of the above considerations, Xgboost, LogisticRegression, and GradientBoost were selected as the basic models, and LightGBM assumed the role of the sub-learner in this study. The whole process was carried out in a Python environment.

3.3. Revision of Preliminary Gully Extraction

To improve the results of the machine learning method, false positives were removed. In the absence of topographic data, the main idea of this part of the study was to remove residential areas, roads, arable land, and forest areas from the above extraction results, which have their own characteristics distinct from gully erosion; further, some of their segmentation objects are easily

misclassified when extracting gullies based on OBIA. The principle of this step was to remove as many of the abovementioned ground objects as possible without removing other objects not addressed. There are many similar features between unused land and gully erosion that are relatively difficult to distinguish. Moreover, unused land is vulnerable to erosion and is often distributed near gully erosion; thus, its removal was not considered.

Residential areas include houses with blue tiled roofs and courtyards. Each independent residential area appears as a regular rectangle. Roads are strip-like and belong to one of two types: Cement roads or asphalt roads. The cultivated land presents a regular, uniform texture, with no vegetation on the surface. Forest land appears as contiguous or striped green areas.

On the basis of the optimal segmentation, target features, and the preliminary results above, the four types of objects used the nearest neighbor rule or membership function rule to extract as many corresponding pure objects as possible within eCognition. The houses of residential areas and cement roads were identified by higher values of the blue band and brightness, respectively. In addition, one obvious difference between houses and cement roads is the border length that roads have along the border length (houses have the opposite). The minimum boundary geometry of each housing agglomeration area was considered a residential area. Forest areas were distinguished mainly based on high values of the VI′, whereas arable land was detected with low values of the VI′ and brightness and a homogeneous texture. Next, these four types of objects were removed from the results above in ArcGIS to map the gullies.

3.4. Validation

Validation of the extraction is essential for judging the fitness of the model presented in this study for gully extraction. Validation was implemented by an overlay analysis between the gully map obtained from the model and the inventory map of gully erosion provided by field surveys and expert knowledge. Adopting different approaches to assess the extraction results highlighted the added value [60]. In this study, we used two strategies to estimate the accuracy of the extraction: (1) The conventional method, whereby the classification results were divided into two types—correct or wrong compared to the reference data. (2) Inspired by similar studies [19,61], there were two kinds of errors of commission. One was a limited error of commission, which bordered on the correctly extracted areas and displayed similar spectral, textural, and geometric properties. The remaining error of commission was regarded as a true error of commission, which was entirely falsely extracted.

The precision, recall, and F-score were used to process the quantitative analysis of the extraction results. The effects of these indicators are as follows. Precision indicated how many areas that the model determined to be positive were true positive areas. Recall indicated how many positive samples were judged to be positive by the model. F-score was an index balancing precision against recall. Ideally, all three indexes were 0.

$$\text{precision} = \frac{|A \cap B|}{|A|} \tag{11}$$

$$\text{recall} = \frac{|A \cap B|}{|B|} \tag{12}$$

$$\text{F} - \text{score} = \frac{2 \times precision \times recall}{precision + recall} \tag{13}$$

Here, A is the area of extracted gully erosion and B is the gully area in the reference data.

4. Results

The determination of the shape and compactness parameters was based on the value of ED. When the ED value reached its minimum, the corresponding parameters were defined as best. Take the result of the first iteration as an example, the goodness assessment is shown in Tables 2–4. The primary range of the parameters of compactness and shape are obtained based on Tables 2 and 3, respectively:

(0.8, 1.0) and (0.1, 0.4). Each combination of the compactness and shape parameters was screened as a candidate to obtain the optimum value (compactness: 0.8; shape: 0.3) shown in Table 4.

Table 2. Segmentation accuracy evaluation for selecting the compactness parameter.

Scale	Shape	Compactness	OS (%)	US (%)	ED (%)
50	0.1	0	19.3	20.2	19.8
50	0.1	0.1	17.1	20.0	18.6
50	0.1	0.2	18.7	19.0	18.8
50	0.1	0.3	20.3	19.3	19.8
50	0.1	0.4	16.9	20.4	18.7
50	0.1	0.5	17.7	18.8	18.2
50	0.1	0.6	20.1	18.2	19.2
50	0.1	0.7	18.0	18.7	18.4
50	**0.1**	**0.8**	**17.5**	**18.6**	**18.1**
50	**0.1**	**0.9**	**17.4**	**18.6**	**18.0**
50	**0.1**	**1.0**	**18.7**	**17.1**	**17.9**

Table 3. Segmentation accuracy evaluation for selecting the shape parameter.

Scale	Shape	Compactness	OS (%)	US (%)	ED (%)
50	0	0.8	20.8	19.5	20.1
50	**0.1**	**0.8**	**17.5**	**18.6**	**18.1**
50	**0.2**	**0.8**	**18.3**	**16.6**	**17.4**
50	**0.3**	**0.8**	**17.9**	**16.9**	**17.4**
50	**0.4**	**0.8**	**17.8**	**17.5**	**17.7**
50	0.5	0.8	18.9	17.8	18.3
50	0.6	0.8	20.4	18.9	19.5
50	0.7	0.8	20.9	22.2	21.5
50	0.8	0.8	26.7	23.5	25.2
50	0.9	0.8	40.9	28.7	35.4

Table 4. Segmentation accuracy evaluation for selecting the optimal combination of compactness and shape.

Scale	Shape	Compactness	OS (%)	US (%)	ED (%)
50	0.1	0.8	17.5	18.6	18.1
50	0.1	0.9	17.4	18.6	18.0
50	0.1	1.0	18.7	17.1	17.9
50	0.2	0.8	18.3	16.6	17.4
50	0.2	0.9	17.6	17.1	17.4
50	0.2	1.0	17.4	17.7	17.6
50	**0.3**	**0.8**	**17.6**	**16.7**	**17.2**
50	0.3	0.9	18.1	16.6	17.3
50	0.3	1.0	17.8	17.7	17.7
50	0.4	0.8	17.8	17.5	17.7
50	0.4	0.9	17.4	17.5	17.5
50	0.4	1.0	18.7	17.9	18.3

Based on the acquired shape and compactness parameter, estimation of the scale parameter is provided to help determine the most proper scale (Figure 7). Because over-segmentation produces fewer threats to classification accuracy than under-segmentation [62], and based on an exploration of the scale parameters in alternating regions, 34 was chosen as the value of the scale parameter.

Finally, the segmentation parameter set (scale, shape, and compactness) was established as (34, 0.3, 0.8).
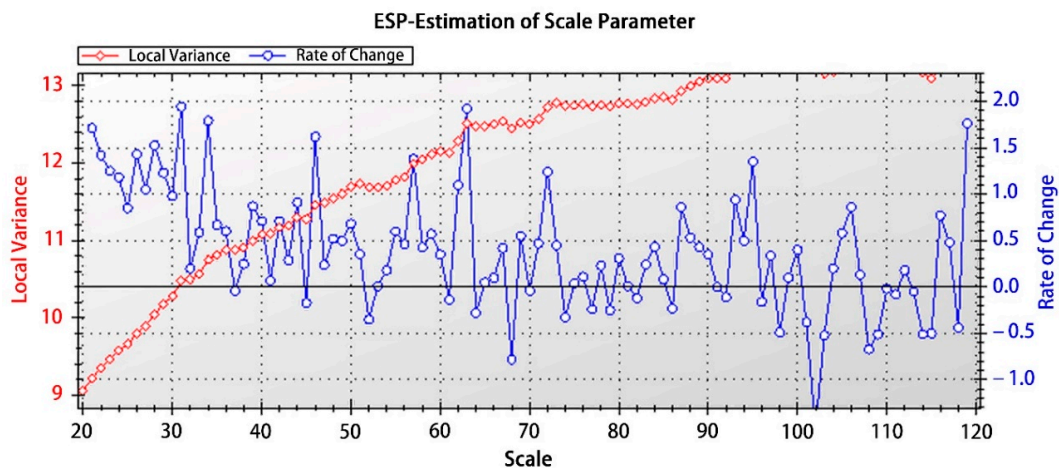
Figure 7. Local Variance (LV) graph of the scale parameter.

The segmentation results were obtained. Figure 8 illustrates some generated objects that were divided into objects of gully and objects of non-gully, which depended on the relationship between the percentage of gully erosion area and 50%, together with a gully system boundary that was obtained by digitizing the gully systems using visual image interpretation based on field investigations and expert knowledge. Objects with a percentage of gully erosion area greater than or equal to 50% were considered as objects of gully. On the contrary, other objects are objects of non-gully.
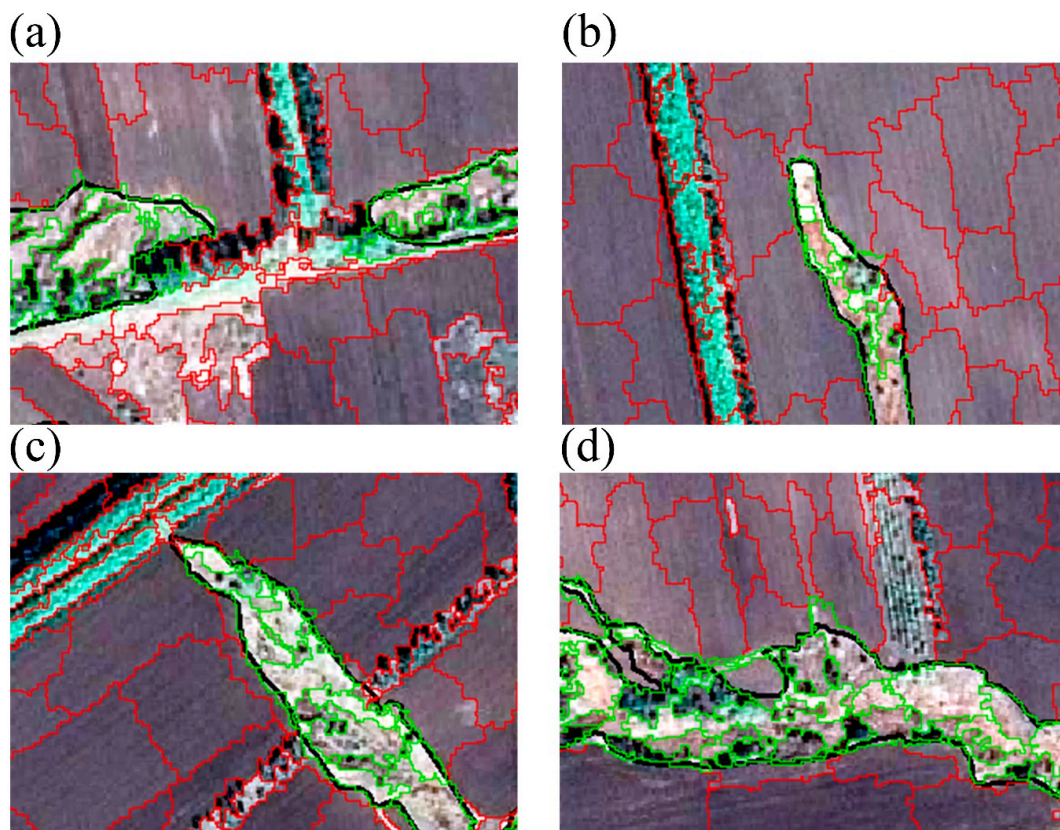
Figure 8. Segmentation results including objects of gully in green and objects of non-gully in red, with manually polygonized reference data of gully. As four examples, (**a**–**d**) show the segmentation results of different regions.

The training data were randomly categorized into the training subset or testing subset with a ratio of 3. By setting and adjusting the parameters of each method based on machine learning algorithms to fit them with the training subset, the initial models were constructed. Table 5 reveals the goodness-of-fit results. TPOT and LightGBM obtained similar performance, and both were better than stacking-gully. In addition, TPOT was a time-consuming procedure; TPOT required almost a week to finish its search, while LigthGBM could finish its search within minutes. Therefore, LightGBM was determined to be the best choice for extracting gully erosion.

Table 5. Goodness-of-fit results for the three machine-learning methods.

Model	Data	Precision (%)	Recall (%)	F-score (%)
TPOT	Training data	93.4	94.6	94.0
	Training subset	99.8	99.8	1.0
	Test subset	73.3	77.4	75.3
LightGBM	Training data	91.2	96.1	93.6
	Training subset	98.6	1.0	99.3
	Test subset	70.1	83.1	76.0
Stacking-gully	Training data	68.1	79.7	73.4
	Training subset	68.3	76.7	72.2
	Test subset	67.7	89.5	77.1

Figure 9 shows that the extraction accuracy first increased with an increment of the number of features and remained basically stable when the number exceeds 15. Thus, it was unnecessary to select and remove features in this study, thereby simplifying the extraction process of gullies. In addition, VI' was sixth in feature importance, which demonstrates the significance of its introduction to the extraction of gullies.
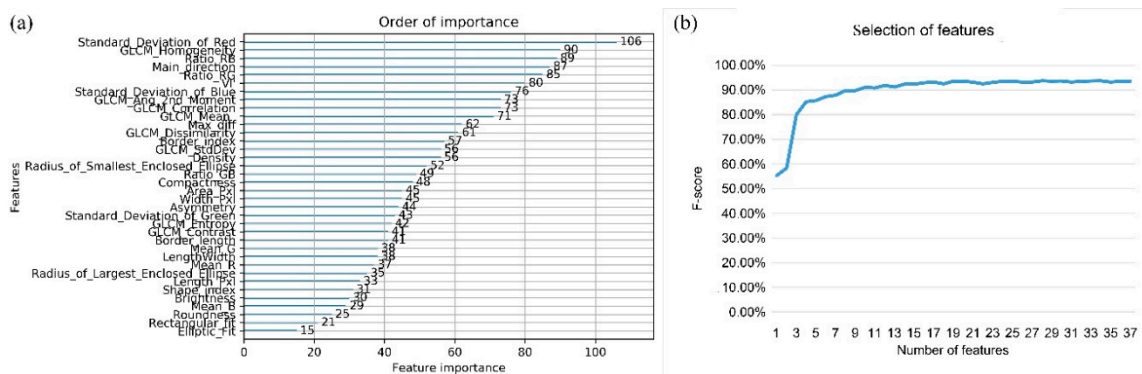

Figure 9. Evaluation of the features. (**a**) Order of importance. (**b**) Selection of features based on feature importance.

Based on initial extraction by LightGBM and subsequent modification, an overview of the extraction result in the test area is displayed in Figure 10. The dark green polygons indicate correctly extracted areas of gully erosion. The light green and blue polygons indicate the limited and true error of omission, respectively, and the light red polygons show the error of commission. The reference data are represented as colorless polygons with black boundaries.
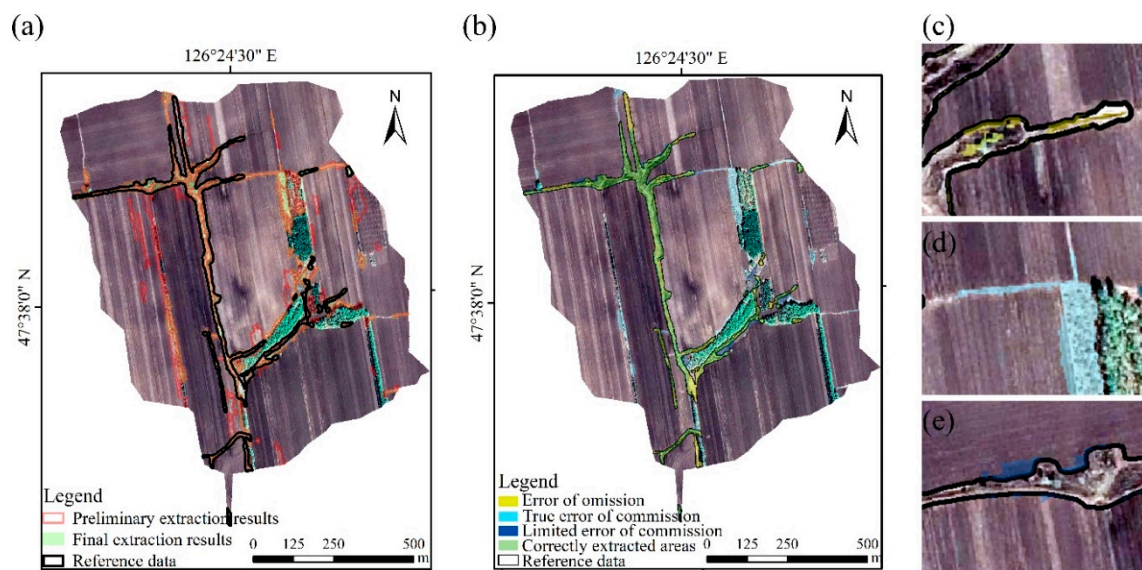
Figure 10. (**a**) Results of initial extraction and revised extraction. (**b**) Mapping of gully in the test area. (**c**) Limited error of commission. (**d**) True error of commission. (**e**) Error of omission.

Table 6 shows the accuracy evaluation of the extraction results. The comparison between the results of preliminary extraction and revised extraction showed that although the area of gully erosion extracted correctly decreased, revised extraction significantly improved the situation that non-gully was wrongly classified into gully erosion, and the performance of gully extraction was better in general. Finally, this study had a recall of 75.6%. A total of 24.4% of gully areas were not correctly extracted, which constitutes an omission error that mainly occurs at inflection points, bifurcations, and small-scale separate erosion gullies. A total of 38.4% of the extracted results were mistakenly extracted as gullies, while almost half of the errors were in the vicinity of gully and exhibited similar spectral, textural, and geometric properties to gully erosion, which was regarded as the limited error of commission. The true error of commission accounted for 21.1% of the extracted results. These errors occurred mainly in areas containing roadsides, farm paths, weed land, bare land, and sparse woodland. Overall, most gullies were correctly identified, and this method enabled an appropriate extraction of the gullies.

Table 6. Evaluation of the extraction results of the test area.

Process	Layer	Area(m^2)	Evaluation Index	Ratio (%)
Initial extraction	Reference data	47,938.7	Precision	40.7
	Gully area extracted	98,641.1	Recall	83.7
	Correctly extracted area	40,127.0	F-score	54.8
	False negatives	7811.8	Error of omission	16.3
	Truly false positives	35,428.4	True error of commission	35.9
	Limited false positives	23,085.7	Limited error of commission	23.4
Revised extraction	Reference data	47,938.7	Precision	61.6
	Gully area extracted	58,787.9	Recall	75.6
	Correctly extracted area	36,234.8	F-score	67.8
	False negatives	11,703.9	Error of omission	24.4
	Truly false positives	12,389.6	True error of commission	21.1
	Limited false positives	10,163.4	Limited error of commission	17.3

To examine the performance of this approach, validation datasets were used (Figure 11). Table 7 shows the validation results. The gullied areas had an underestimation of less than 40%. These areas were mostly adjacent to the correctly extracted gullied areas and almost certainly did not indicate the absence of a whole gully. A total of 20% to 30% of the areas were entirely classified falsely and thus assigned as the overestimation. That is, more than 70% of the extraction results were correct or

important indicators for the location of gullies. The range of underestimation and overestimation varied as a result of considering areas with different complexities of geographical object categories for validation. The extraction results in the watershed with residential areas and more roads were worse, as they were either underestimated or overestimated.
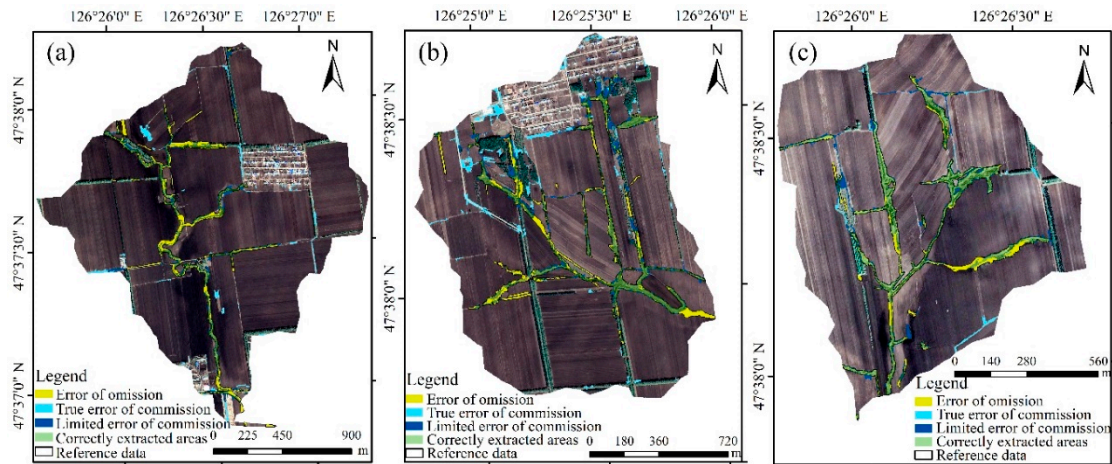

Figure 11. Mapping of the gullies in the validation areas. (**a**) v_1. (**b**) v_2. (**c**) v_3.

Table 7. Evaluation of the extraction results of the validation areas.

Subset	Gullied Area Digitized (m²)	Gullied Area Extracted (m²)	Precision (%)	Recall (%)	Error of Omission (%)	True Error of Commission (%)	Limited Error of Commission (%)
v_1	97,832.3	98,016.4	60.2	60.3	39.7	23.2	16.6
v_2	68,479.2	72,372.6	76.3	82.0	18.0	5.6	18.1
v_3	111,655.2	137,558.1	55.9	68.8	31.2	23.3	20.8

5. Discussion

GE imagery, a type of optical imaging technology, is the only data source used for gully mapping in this paper. Although there is an unquestionable link between topographic attributes and the occurrence of gullies, which helps to improve the extraction of gullies (as suggested by recent studies) [23,25,63], the acquisition and update of high-precision topographic data remain difficult for a vast number of users compared to optical imagery, especially when conducting regional gully erosion surveys and mapping. Hence, this type of data selection allows more people to use the method presented and facilitates large-scale gully erosion investigations at the sacrifice of some accuracy. Considering the scope of application, the lack of topographic data is regarded as an advantage. In addition, other optical imaging techniques can be used as the data source due to the adequate spatial and spectral information provided, such as images captured by QuickBird, WorldView, and GeoEye (if affordable). The approach developed using GE imagery provides good cross-data portability, which increases our ability to study the dynamic changes of gully erosion because there are more data to choose from at different points in time.

This study combines direct extraction and indirect extraction by removing non-eroded gully areas from the preliminary results of direct extraction to extract gully erosion; this is described as two-end extrusion. Various methods based on machine learning algorithms were evaluated and selected to obtain an optimal model for the direct extraction of gully erosion. This step explored the effectiveness of methods based on machine learning algorithms other than the random forest methods commonly used in previous studies for the extraction of gully erosion. We determined that LightGBM may be a better choice. Based on the preliminary extraction results, the final extraction results were obtained, via eCognition, by removing the areas identified not to be eroded by the gully. Compared with the studies on direct extraction [24,25], two-end extrusion has the capacity to improve the extraction results when

the experimental conditions are consistent. In addition, the inclusion of a machine learning method simplifies the extraction process and enhances the automaticity of the approach by establishing rule sets within eCognition [19,20].

Although the extraction accuracy is not completely satisfactory, it is still acceptable considering that this method uses no topographic data. The lack of topographic data may lead to inadequate object features, which could have contributed to the extraction errors. False negatives are mainly isolated fragments and directly adjacent to the correctly extracted areas of gully erosion (true positives), which indicates that the main body of gullies can achieve good initial recognition. Almost all false positives are segments of roadsides, farm paths, weed land, bare land, and sparse woodland. Studies have shown that gully erosion develops seriously in areas with greater disturbances and lower vegetation coverage. Indeed, one of the primary causes of gully erosion is roads and paths [39,64,65]. Hence, false positives are mainly concentrated in high risk areas of gully erosion, which deserve attention. For the above-mentioned reasons, the method presented in this paper is a viable option for initially determining gully erosion, with acceptable accuracy to provide information for the pertinent stakeholders.

This study was conducted in a typical black soil region of Northeast China. Our results can be generalized to other areas of the black soil region for two reasons: (1) There are similarly distributed rules for gully erosion across the entire black soil region, despite variations in density and shape. (2) The results apply to a study area with the most serious and complex gully erosion, which demonstrates its gully mapping capabilities in areas with lighter and simpler gully erosion. However, when attempting to extend the present approach to areas other than the black soil region, an elaborate field investigation should be carried out. Based on the geographic environmental information collected, as well as expert knowledge, this approach should be adjusted accordingly.

Gullies, after harnessing, show the characteristics of woodland in the optical images, as the management of gully erosion in Northeast China has been vigorously put into practice since 2002 by the Chinese Government [66]. In light of the weak penetration of optical images and the removal of woodland, gullies after harnessing, and other gullies in forest areas, cannot be extracted. However, the stable state of gullies with less damage makes this issue less serious and less conspicuous. In this paper, unmanaged gullies and gullies harnessed in non-forest areas were extracted, but the distinction between these two types of gullies for the control and monitoring of gully erosion cannot be achieved and requires further study.

6. Conclusions

In this paper, we successfully built a two-end extrusion method integrating methods based on machine learning algorithms and OBIA using GE images as the data source to obtain the initial extraction of a gully in the black soil region. Several conclusions can be summarized: (1) This highly automatic method simplifies the extraction process and is applicable to diverse image data, which makes it broadly available and transferable to a large extent. Among mainstream methods based on machine learning algorithms, LightGBM is regarded as the best choice to preliminarily extract a gully due to its fast speed and high accuracy. (2) The performance of this method was modest but acceptable, as it was verified to have a true overestimation in the range of 20% to 30%, which was mainly caused by land use types with a high erosion risk, such as bare land and farm lanes, being falsely classified as gullies, and an underestimation of less than 40%, which was largely adjacent to the correctly extracted areas of the gully. Further, the lower the complexity of the geographical object categories of the study area, the more satisfactory its accuracy of extraction. (3) Due to the weak penetration of optical images and the removal of woodland, the extracted gullies consisted of unmanaged gullies and gullies harnessed in the non-forest areas, while it excluded gullies after harnessing and other gullies in the forest areas.

Meanwhile, this method has various limitations that must be overcome in future studies. On the one hand, the extraction accuracy of this method is not completely satisfactory. On the other hand, it is not possible to extract and distinguish all types of gully erosion, and each gully cannot be extracted as

an individual. In addition, the applicability of this method needs to be verified in more areas in the black soil region and beyond.

Author Contributions: Methodology, Data curation, Formal analysis & Writing—original draft, B.W.; Conceptualization & Writing—review & editing, X.W.; Project administration, Z.Z.; Supervision, X.Z.; Validation, L.Y.; Investigation, S.H. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Poesen, J.; Nachtergaele, J.; Verstraeten, G.; Valentin, C. Gully erosion and environmental change: importance and research needs. *Catena* **2003**, *50*, 91–133. [CrossRef]
2. Ma, H.; Zhao, H. Uuited nations: convention to combat desertification in those countries experiencing serious drought and/or desertification, particularly in Africa. *Int. Leg. Mater.* **1994**, *33*, 1328–1382.
3. Ibrahim, A.H.; Yaro, N.A.; Adebola, A.O. Assessing the socio-economic impact of gully erosion in Chikun Local Government Area, Kaduna State, Nigeria. *Sci. World J.* **2017**, *5*, 30–41.
4. Jahantigh, M.; Pessarakli, M. Causes and effects of gully erosion on agricultural lands and the environment. *Commun. Soil Sci. Plant Anal.* **2011**, *42*, 2250–2255. [CrossRef]
5. Li, H.; Cruse, R.M.; Bingner, R.L.; Gesch, K.R.; Zhang, X. Evaluating ephemeral gully erosion impact on *Zea mays* L. yield and economics using AnnAGNPS. *Soil Tillage Res.* **2016**, *155*, 157–165. [CrossRef]
6. Li, Y.; Zhang, S.; Yang, P. Soil erosion and its relationship to the spatial distribution of land use patterns in the Lancang River Watershed, Yunnan Province, China. *Agric. Sci.* **2015**, *6*, 823–833. [CrossRef]
7. Zaimes, G.N.; Schultz, R.C. Assessing riparian conservation land management practice impacts on gully erosion in Iowa. *Environ. Manag.* **2012**, *49*, 1009–1021. [CrossRef] [PubMed]
8. Zakerinejad, R.; Maerker, M. An integrated assessment of soil erosion dynamics with special emphasis on gully erosion in the Mazayjan basin, southwestern Iran. *Nat. Hazards* **2015**, *79*, 25–50. [CrossRef]
9. Kheir, R.B.; Hansmann, B.; Abdallah, C. Detecting major environmental parameters influencing the gully erosion occurrence in Mediterranean karst landscapes of Lebanon using remote sensing and GIS statistical correlations. In Proceedings of the Egu General Assembly Conference, Vienna, Austria, 2–7 May 2010.
10. Pu, L.; Zhang, S.; Wang, R.; Chang, L.; Yang, J. Analysis of erosion gully information extraction based on multi-resource remote sensing images. *Geogr. Geo-Inf. Sci.* **2016**, *32*, 90–94. (In Chinese)
11. Wang, R.; Zhang, S.; Pu, L.; Yang, J.; Yang, C.; Chen, J.; Guan, C.; Wang, Q.; Chen, D.; Fu, B. Gully erosion mapping and monitoring at multiple scales based on multi-source remote sensing data of the Sancha River Catchment, Northeast China. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 200. [CrossRef]
12. Yan, Y.; Zhang, S.W.; Yue, S. Classification of erosion gullies by remote sensing and spatial pattern analysis in black soil region of eastern Kebai. *Acta Geogr. Sinica* **2007**, *27*, 193–199. (In Chinese)
13. Garosi, Y.; Sheklabadi, M.; Conoscenti, C.; Pourghasemi, H.R.; Van Oost, K. Assessing the performance of GIS-based machine learning models with different accuracy measures for determining susceptibility to gully erosion. *Sci. Total Environ.* **2019**, *664*, 1117–1132. [CrossRef] [PubMed]
14. Garosi, Y.; Sheklabadi, M.; Pourghasemi, H.R.; Besalatpour, A.A.; Conoscenti, C.; Oost, K.V. Comparison of differences in resolution and sources of controlling factors for gully erosion susceptibility mapping. *Geoderma* **2018**, *330*, 65–78. [CrossRef]
15. Metternicht, G.I.; Zinck, J.A. Evaluating the information content of JERS-1 SAR and Landsat TM data for discrimination of soil erosion features. *ISPRS J. Photogramm. Remote Sens.* **1998**, *53*, 143–153. [CrossRef]
16. Rahmati, O.; Tahmasebipour, N.; Haghizadeh, A.; Pourghasemi, H.R.; Feizizadeh, B. Evaluation of different machine learning models for predicting and mapping the susceptibility of gully erosion. *Geomorphology* **2017**, *298*, S0169555X16308418. [CrossRef]
17. Blaschke, T.; Strobl, J. What's wrong with pixels? Some recent developments interfacing remote sensing and GIS Interfacing Remote Sensing and GIS. *Z. Geoinf.* **2001**, *6*, 12–17.

18. Karami, A.; Khoorani, A.; Noohegar, A.; Shamsi, S.R.F.; Moosavi, V. Gully erosion mapping using object-based and pixel-based image classification methods. *Environ. Eng. Geosci.* **2015**, *21*, 101–110. [CrossRef]

19. D'Oleire-Oltmanns, S.; Marzolff, I.; Tiede, D.; Blaschke, T. Detection of gully-affected areas by applying object-based image analysis (OBIA) in the region of Taroudannt, Morocco. *Remote Sens.* **2014**, *6*, 8287–8309. [CrossRef]

20. Shruthi, R.B.V.; Kerle, N.; Jetten, V. Object-based gully feature extraction using high spatial resolution imagery. *Geomorphology* **2011**, *134*, 260–268. [CrossRef]

21. Shruthi, R.B.V.; Kerle, N.; Jetten, V.; Abdellah, L.; Machmach, I. Quantifying temporal changes in gully erosion areas with object oriented analysis. *Catena* **2015**, *128*, 262–277. [CrossRef]

22. Eustace, A.H.; Pringle, M.J.; Denham, R.J. A risk map for gully locations in central Queensland, Australia. *Eur. J. Soil Sci.* **2011**, *62*, 431–441. [CrossRef]

23. Liu, K.; Ding, H.; Tang, G.; Song, C.; Liu, Y.; Jiang, L.; Zhao, B.; Gao, Y.; Ma, R. Large-scale mapping of gully-affected areas: An approach integrating Google Earth images and terrain skeleton information. *Geomorphology* **2018**, *314*, 13–26. [CrossRef]

24. Liu, K.; Ding, H.; Tang, G.; Zhu, A.; Yang, X.; Jiang, S.; Cao, J. An object-based approach for two-level gully feature mapping using high-resolution DEM and imagery: A case study on hilly loess plateau region, China. *Chin. Geogr. Sci.* **2017**, *27*, 415–430. [CrossRef]

25. Shruthi, R.B.V.; Kerle, N.; Jetten, V.; Stein, A. Object-based gully system prediction from medium resolution imagery using Random Forests. *Geomorphology* **2014**, *216*, 283–294. [CrossRef]

26. Díaz-Uriarte, R.; Andrés, S.A.D. Gene selection and classification of microarray data using random forest. *BMC Bioinform.* **2006**, *7*, 3. [CrossRef]

27. Prasad, A.M.; Iverson, L.R.; Liaw, A. Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems* **2006**, *9*, 181–199. [CrossRef]

28. Chen, T.; Carlos, G. XGBoost: A scalable tree boosting system. In Proceedings of the ACM Sigkdd International Conference on Knowledge Discovery & Data Mining, San Francisco, CA, USA, 13–17 August 2016.

29. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T. LightGBM: A highly efficient gradient boosting decision tree. In Proceedings of the 31st Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017.

30. Babajide, M.I.; Saeed, F. Bioactive molecule prediction using Extreme Gradient Boosting. *Molecules* **2016**, *21*, 983. [CrossRef]

31. Sheridan, R.P.; Wang, W.M.; Liaw, A.; Ma, J.; Gifford, E.M. Extreme Gradient Boosting as a method for quantitative structure-activity relationships. *J. Chem. Inf. Model.* **2016**, *56*, 2353. [CrossRef]

32. Castillo, C.; Taguas, E.V.; Zarco-Tejada, P.; James, M.R.; Gómez, J.A. The normalized topographic method: an automated procedure for gully mapping using GIS. *Earth Surf. Process. Landf.* **2015**, *39*, 2002–2015. [CrossRef]

33. Evans, M.; Lindsay, J. High resolution quantification of gully erosion in upland peatlands at the landscape scale. *Earth Surf. Process. Landf.* **2010**, *35*, 876–886. [CrossRef]

34. Guo, Z.; Shao, X.; Xu, Y.; Miyazaki, H.; Ohira, W.; Shibasaki, R. Identification of village building via google earth images and supervised machine learning methods. *Remote Sens.* **2016**, *8*, 271. [CrossRef]

35. Ludwig, A.; Meyer, H.; Higginbottom, T.; Nauss, T. Classifying Google Earth images as training sites for application to a larger scale monitoring of bush encroachment in South Africa. In Proceedings of the Egu General Assembly, Vienna, Austria, 17–22 April 2016.

36. Meyer, G.E.; Neto, J.C. Verification of color vegetation indices for automated crop imaging applications. *Comput. Electron. Agric.* **2008**, *63*, 282–293. [CrossRef]

37. Ministry of Water Resources of the People's Republic of China (MWR). *Bulletin on Soil and Water Conservation of the First National Water Conservancy Census*; China Water Power Press: Beijing, China, 2013. (In Chinese)

38. Tang, L.; Meng, L.; Zhang, F. Mechanism of gully development on slope farmland in black soil area, China. *J. Anhui Agric. Sci.* **2012**, *40*, 819–821. (In Chinese)

39. Fan, H.; Gu, G.; Wang, Y.; Zhong, Y. Gully erosion development and environmental characteristics in the black soil region of Northeast China. *Soil Water Conserv. China* **2013**, *10*, 75–78. (In Chinese)

40. Bai, J.; Hui, L. Preliminary investigation on the development and harm of erosion gully in the black soil area of Northeast China. *Soil Water Conserv. China* **2015**, *8*, 68–70. (In Chinese)

41. Du, G.; Lei, G.; Zong, X. Analysis on spatial pattern of gully erosion across typical black hilly region of Northeast China. *Res. Soil Water Conserv.* **2011**, *18*, 94–97. (In Chinese)

42. Yan, Y.; Zhang, S.; Li, X. Temporal and spatial variation of erosion gullies in Kebai black soil region of Heilongjiang during the past 50 years. *Acta Geogr. Sinica* **2005**, *60*, 1015–1020. (In Chinese)

43. Yan, Y.; Zhang, S.; Yue, S. Application of Corona and Spot imagery on erosion gully research in typical black soil regions of Northeast China. *Resour. Sci.* **2006**, *6*, 154–160. (In Chinese)

44. Wang, W.; Deng, R.; Zhang, S. Preliminary research on risk evaluation of gully erosion in typical black soil area of Northeast China. *J. Nat. Resour.* **2014**, *29*, 2058–2067.

45. Wang, W.; Zhang, S.; Fang, H. Coupling mechanism of slope-gully erosion in typical black soil area of Northeast China. *J. Nat. Resour.* **2012**, *27*, 2113–2122.

46. Xu, X.Z.; Xu, Y.; Chen, S.C.; Xu, S.G.; Zhang, H.W. Soil loss and conservation in the black soil region of Northeast China: a retrospective study. *Environ. Sci. Policy* **2010**, *13*, 793–800. [CrossRef]

47. Platoncheva, E.V. Investigation of the dynamics of ephemeral gully erosion on arable land of the forest-steppe and steppe zone of the East of the Russian Plain from remote sensing data. In *Iop Conference Series: Earth & Environmental Science*; IOP Publishing: Bristol, UK, 2018; p. 012019.

48. Wang, Q.; Xu, J.; Cheng, Y.; Li, J.; Wang, X. Influence of the varied spatial resolution of remote sensing images on urban and rural residential information extraction. *Resour. Sci.* **2012**, *34*, 159–165. (In Chinese)

49. Baatz, M.; Schäpe, A. Multiresolution segmentation—An optimization approach for high quality multi-scale image segmentation. *Angew. Geogaraphische Inf. Sev.* **2000**, *12*, 12–23.

50. Martha, T.R.; Kerle, N.; Van Westen, C.J. Segment optimization and data-driven thresholding for knowledge-based landslide detection by object-based image analysis. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 4928–4943. [CrossRef]

51. Drăguţ, L.; Csillik, O.; Eisank, C.; Tiede, D. Automated parameterisation for multi-scale image segmentation on multiple layers. *ISPRS J. Photogramm. Remote Sens.* **2014**, *88*, 119–127.

52. Woodcock, C.E.; Strahler, A.H. The factor of scale in remote sensing. *Remote Sens. Environ.* **1987**, *21*, 311–332. [CrossRef]

53. Clinton, N.; Holt, A.; Scarborough, J.; Li, Y.; Peng, G. Accuracy assessment measures for object-based image segmentation goodness. *Photogramm. Eng. Remote Sens.* **2010**, *76*, 289–299. [CrossRef]

54. Blaschke, T.; Hay, G.J.; Kelly, M.; Lang, S.; Hofmann, P.; Addink, E.; Feitosa, R.Q.; Meer, F.V.D.; Werff, H.V.D.; Coillie, F.V. Geographic object-based image analysis—Towards a new paradigm. *ISPRS J. Photogramm. Remote Sens.* **2014**, *87*, 180–191. [CrossRef]

55. Haralick, R.M.; Shanmugam, K.; Dinstein, I. Textural features for image classification. *Stud. Media Commun.* **1973**, *3*, 610–621. [CrossRef]

56. Olson, R.S.; Bartley, N.; Urbanowicz, R.J.; Moore, J.H. Evaluation of a tree-based pipeline optimization tool for automating data science. In Proceedings of the Genetic & Evolutionary Computation Conference, Denver, CO, USA, 20–24 July 2016.

57. Olson, R.S.; Urbanowicz, R.J.; Andrews, P.C.; Lavender, N.A.; Kidd, L.C.; Moore, J.H. Automating biomedical data science through tree-based pipeline optimization. In Proceedings of the European Conference on the Applications of Evolutionary Computation, Porto, Portugal, 30 March–1 April 2016.

58. Shi, H. *Best-First Decision Tree Learning*; The University of Waikato: Hamilton, New Zealand, 2007.

59. Wolpert, D.H. Stacked generalization. *Neural Netw.* **1992**, *5*, 241–259. [CrossRef]

60. Drăguţ, L.; Eisank, C. Automated object-based classification of topography from SRTM data. *Geomorphology* **2012**, *141*, 21–33. [CrossRef] [PubMed]

61. Congalton, R.G.; Green, K. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*; CRC Press: Boca Raton, FL, USA, 2002.

62. Liu, Y.; Bian, L.; Meng, Y.; Wang, H.; Zhang, S.; Yang, Y.; Shao, X.; Wang, B.O. Discrepancy measures for selecting optimal combination of parameter values in object-based image analysis. *ISPRS J. Photogramm. Remote Sens.* **2012**, *68*, 144–156. [CrossRef]

63. Gómez-Gutiérrez, Á.; Conoscenti, C.; Angileri, S.E.; Rotigliano, E.; Schnabel, S. Using topographical attributes to evaluate gully erosion proneness (susceptibility) in two mediterranean basins: advantages and limitations. *Nat. Hazards* **2015**, *79*, 291–314. [CrossRef]

64. Lesschen, J.P.; Kok, K.; Verburg, P.H.; Cammeraat, L.H. Identification of vulnerable areas for gully erosion under different scenarios of land abandonment in Southeast Spain. *Catena* **2007**, *71*, 110–121.

65. Showers, K.B. Soil erosion in the Kingdom of Lesotho and development of historical environmental impact assessment. *Ecol. Appl.* **1996**, *6*, 653–664. [CrossRef]

66. Qin, W.; Zuo, C.; Fan, J.; Xu, X.; Xu, J. Control measures for gully erosion in black soil areas of Northeast China. *China Water Resour.* **2014**, *20*, 37–41. (In Chinese)