A New Method for Quantitative Analysis of Driving Factors for Vegetation Coverage Change in Mining Areas: GWDF-ANN

Jun Li 1, Tingting Qin 1, Chengye Zhang 1,* , Huiyu Zheng 1, Junting Guo 2, Huizhen Xie 1, Caiyue Zhang 1 and Yicong Zhang 1

1 College of Geoscience and Surveying Engineering, China University of Mining and Technology-Beijing, Beijing 100083, China; junli@cumtb.edu.cn (J.L.); qting@student.cumtb.edu.cn (T.Q.); zhenghy@student.cumtb.edu.cn (H.Z.); huizhenx@student.cumtb.edu.cn (H.X.); zcy_cumtb@student.cumtb.edu.cn (C.Z.); zhangyc@student.cumtb.edu.cn (Y.Z.)

2 State Key Laboratory of Water Resource Protection and Utilization in Coal Mining, Beijing 102209, China; junting.guo.a@chnenergy.com.cn

* Correspondence: czhang@cumtb.edu.cn

Abstract: Mining has caused considerable damage to vegetation coverage, especially in grasslands. It is of great significance to investigate the specific contributions of various factors to vegetation cover change. In this study, fractional vegetation coverage (FVC) is used as a proxy indicator for vegetation coverage. We constructed 50 sets of geographically weighted artificial neural network models for FVC and its driving factors in the Shengli Coalfield. Based on the idea of differentiation, we proposed the geographically weighted differential factors-artificial neural network (GWDF-ANN) to quantify the contributions of different driving factors on FVC changes in mining areas. The highlights of the study are as follows: (1) For the 50 models, the average RMSE was 0.052. The lowest RMSE was 0.007, and the highest was 0.112. For the MRE, the average value was 0.007, the lowest was 0.001, and the highest was 0.023. The GWDF-ANN model is suitable for quantifying FVC changes in mining areas. (2) Precipitation and temperature were the main driving factors for FVC change. The contributions were 32.45% for precipitation, 24.80% for temperature, 22.44% for mining, 14.44% for urban expansion, and 5.87% for topography. (3) Over time, the contributions of precipitation and temperature exhibited downward trends, while mining and urban expansion showed positive trajectories. For topography, its contribution remains generally unchanged. (4) As the distance from the mining area increases, the contribution of mining gradually decreases. At 200 m away, the contribution of mining was 26.69%; at 2000 m away, the value drops to 17.8%. (5) Mining has a cumulative effect on vegetation coverage both interannually and spatially. This study provides important support for understanding the mechanism of vegetation coverage change in mining areas.

Keywords: quantify; GWDF-ANN; FVC; vegetation cover; mining area

1. Introduction

Coal is an important source of energy in the world. For example, China’s primary energy source is mainly coal [1]. Mining has caused land subsidence, occupation, excavation, and pollution [2,3], considerably reducing vegetation areas. Mining also affects the inherent growth of vegetation. For example, when dust from open-pit mines falls on the plant leaf surface, the stomatal conductance of the leaf decreases, and the amount of carbon dioxide exchange is also reduced. Large-scale mining has severely damaged the grassland ecosystem [4] and accelerated changes in vegetation coverage [5]. In addition to mining, many researchers have shown that precipitation, temperature [6–9], topography [10], and urban expansion [11] are important driving factors altering vegetation coverage in mining areas. However, the contribution of each factor to the change of vegetation coverage in mining areas remains unclear. Therefore, it is crucial to accurately quantify the contributions of the different driving factors in altering the fractional vegetation coverage (FVC) in
mining areas. The quantitative analysis of these factors may also help reveal the internal mechanisms of vegetation coverage change in mining areas.

There are two main approaches based on remote sensing used to investigate the impact of mining on vegetation. One is to use vegetation parameters obtained from satellite monitoring as proxy indicators. Direct changes in proxy indicators are then used to analyze the impact of mining on vegetation. Li and Wang used the temporal and spatial changes in the FVC to reveal the changes in vegetation coverage in the Baorixile mining area, the Shengli mining area, and the Yellow River basin mining area [12–14]. Li used NDVI to analyze the impact of mining on vegetation change in Jungar Banner, China, from 2000 to 2017 [15]. Yang used NDVI and LandTrendr algorithms to monitor vegetation disturbance and restoration in the Kurah coal mining area [16]. However, the proxy parameter for vegetation is direct change, which is affected by the complex coupling of multiple factors and cannot be attributed solely to mining.

The other uses statistical methods to construct a regression model based on ordinary linearity for vegetation parameters and their driving factors. For example, Li used multiple linear regression, spatial correlation, and partial least square regression to analyze the impact of mining on the NDVI changes in the grasslands of Chenbarhu Banner, Inner Mongolia, China [17]. Fu Xiao used the rate of change in greenness (RCV) and the coefficient of variation (CV) to conduct correlation and stepwise regression analysis in studying the driving factors for the vegetation change in Xilinhot, particularly mining activities [18]. These studies mainly use a global regression model, in which the relationship between vegetation parameters and driving factors is assumed to be spatially constant. However, such a relationship often varies significantly in space because the vegetation coverage in mining areas may be affected by spatial factors such as climate, topography, and mining. Each driving factor may also have different effects on FVC at various locations. The phenomenon of how spatial relationships vary at different geographic locations is called spatial nonstationarity [19]. Simple linear regression modeling cannot accurately describe the complex driving processes of these different factors in changing the vegetation coverage in mining areas.

Therefore, the local modeling method, geographically weighted regression (GWR), was developed to deal with spatial nonstationarity. GWR has been widely applied in various fields, such as real estate economics [20,21], land-use science [22,23], ecology [24], and criminology [25,26]. However, GWR is essentially linear modeling. The geographically weighted artificial neural network (GWANN) was proposed to address the limitations of the GWR in describing complex nonlinear processes. GWANN uses the distance attenuation kernel function and bandwidth to calculate the geographic weight of the observed value results based on the GWR, solving the problem in spatial nonstationarity modeling. The model uses ANN to construct a nonlinear function model without making any assumption in completing complex spatial prediction tasks. It has been used to solve problems such as real estate valuation [27], agricultural output estimation [28], and wildfire analysis [29]. GWANN has been shown to be significantly better than the GWR when modeling synthetic data and actual data with nonlinearity and high spatial variance, respectively [27].

Vegetation coverage can be evaluated using the FVC determined using satellite imagery [30]. FVC is the ratio of the vertical projection of vegetation (e.g., leaves, stems, branches) on the ground in each pixel, which directly reflects the amount of vegetation in the target area [31,32]. FVC describes and measures vegetation growth and is the most important and sensitive indicator for vegetation coverage change [33,34].

In this study, we developed the geographically weighted differential factors-artificial neural network (GWDF-ANN). It is a mixed methodology based on the GWR and ANN, which constructs a nonlinear model while considering the nonstationarity of variables. Most importantly, this approach can quantify the contributions of different driving factors to the FVC. We applied the above methods to the various driving factors (i.e., temperature, precipitation, topography, mining, and urban expansion) and quantified each driving factor’s contribution on FVC in the Shengli Coalfield of Inner Mongolia for 2004–2020.
2. Methods

2.1. Geographically Weighted Artificial Neural Network

The artificial neural network is a computing model consisting of many neurons and one-way connections [35]. Each neuron $i$ represents a specific output function called an activation function. The connection between the two neurons $i$ and $j$ are controlled by the weight $w_{ij}$, which is the weighted value of the signal through the connection. These neurons are usually organized in layers, and each neuron is directionally connected to the neuron in the next layer. GWANN is different from the traditional ANN. It consists of an input layer, a hidden layer, and an output layer. A network with a single hidden layer can better approximate any continuous function on the closed subset and bounded subset of $n$-dimensional Euclidean space, given enough hidden neurons.

An input neuron is a grid cell that contains location information, $y$-value, and $x$-values. In this paper, $y$-value refers to FVC, and $x$-values refer to temperature, precipitation, topography, mining, and urban expansion. The modeling process of GWANN is shown in Figure 1. Many input neurons from the input layer pass into the hidden layer and are then aggregated and converted in the hidden layer using Equation (1). The output of neuron $i$ is calculated using Equation (2). In neurons, people often use a nonlinear hyperbolic tangent activation function, which has the necessary conditions for calculating the network’s continuous and differentiable error gradient [36], as shown in Equation (3). The output for each neuron is then passed to the output layer. Each output neuron of the output layer is assigned to a location in the geographic space and predicts a global observed value. However, the observed value is still not accurate enough. To simulate an exact nonlinear relationship model, the connection weight of the ANN must be adjusted.

$$\text{Layer}_j = \sum_{i \in S_j} w_{ij} o_i$$  \hspace{1cm} (1)

$$o_i = \Phi(\text{Layer}_i)$$  \hspace{1cm} (2)

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$  \hspace{1cm} (3)

where $\text{Layer}_j$ is the network input of neuron $j$, $w_{ij}$ is the connection weight between neuron $i$ and $j$, $o_i$ is the output of neuron $i$, $S_j$ is a group of neurons that have output connection with neuron $j$, $\Phi$ is the activation function, $f(x)$ is the transfer value after activation of the neuron, and $x$ is the parameter value before neuron activation.

Two main steps are usually required to adjust the ANN. First, backpropagation is used to calculate the error signal for the observed value of each neuron [37]. Unlike traditional ANN, GWANN uses a geographically weighted error function instead of a quadratic error function to calculate the error signal. GWANN weighs the difference between the output neuron and the target value ($y$-value) according to the spatial distance between the location of the output neuron and the observed value. These weights can be interpreted as the GWR model. When the output neuron’s position is closer to the observed value, it is given a higher weight than those farther from the observed value. The geographically weighted error function is defined as Equation (4), and the backpropagation error signal is determined using Equation (5).
Figure 1. The modeling process of GWANN. N1, N2, N3, and N4 are the multiple input neurons composing the input layer. The rectangles on the left represent the attributes of each input neuron (y-value and x-values), and Y0 is the prediction.

Gradient descent is used to adjust the connection weight in the second step, as in Equation (6). We use Nesterov’s acceleration gradient [38], which significantly improves training performance [39] and makes training more robust. These two steps are repeated until the termination condition is reached (e.g., the error rate is lower than a predetermined threshold) and the GWANN model construction is completed. Finally, the modeling is completed, and we obtain the prediction(Y0). The GWANN can be implemented on RStudio software. An R package providing an implementation of GWANN can be obtained from https://github.com/jhagenauer/gwann (accessed on 1 July 2021) [27], the purpose of which was to predict Y0 in this study.

\[
E = \frac{1}{2} \sum_{i=1}^{n} v_i (t_i - o_i)^2 \tag{4}
\]

\[
\delta_j = \begin{cases} 
\Phi'(Layer_j) v_j (o_j - t_j) & \text{if } j \text{ is an output neuron} \\
\Phi'(Layer_j) \sum_k \delta_k w_{jk} & \text{otherwise}
\end{cases} \tag{5}
\]

\[
\Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}} = -\eta \delta_j o_i \tag{6}
\]

where \(o_i\) is the output of neuron \(i\), \(o_j\) is the output of neuron \(j\), \(t_i\) is the target value of neuron \(i\), \(t_j\) is the target value of neuron \(j\), \(v_i\) is the geographically weighted distance between the observed value and the position of output neuron \(i\), \(v_j\) is the geographically weighted distance between the observed value and the location of output neuron \(j\), \(n\) is the number of output neurons, \(Layer_j\) is the network input of neuron \(j\), \(\Phi'\) is the derivative of the activation function, \(\delta\) is the value of neuron \(j\) error signal, \(w_{jk}\) is the connection weight between neuron \(j\) and \(k\), \(w_{ij}\) is the connection weight between neuron \(i\) and \(j\), \(E\) is the error function, and \(\eta\) is the learning rate.
2.2. Geographically Weighted Differential Factors-Artificial Neural Network

The GWANN can be used for value prediction but not in exploring the contributions of driving factors. Hence, we developed the geographically weighted differential factors-artificial neural network (GWDF-ANN), which can be obtained from https://figshare.com/s/3d06bb3dc4660396539 (accessed on 15 November 2021). In GWDF-ANN, we adapted the existing architecture (GWANN) and we added a bias for each driving factor before prediction. The idea of differentiation is then utilized to quantify the contribution of each factor. The steps are as follows.

First, we use the FVC and five driving factors (temperature, precipitation, topography, mining, and urban expansion) \( (x_i) \) to build the GWANN model and predict \( Y_0 \). Second, a bias \( (\Delta x_i) \) is added to the \( x_i \), which is calculated using Equation (7). The bias is added to a specific driving factor for all input neurons, while the other driving factors for each input neuron are kept constant. The bias does not affect the learning in the hidden layer of other inputs in computing the contribution of one factor. The setting of the bias is carried out in a series of experiments. If the bias is greater than 0.001, the contribution of the driving factor always changes. But if the bias is less than 0.001, the contribution of the driving factor hardly varies. From experiments, we found that 0.001 is the threshold value at which the contribution of the driving factor tends to stabilize. The bias is added separately for each driving factor and then input into the model to predict \( Y (Y_{xi}) \). The partial derivative for each driving factor is then calculated separately, using Equation (8). Finally, the partial derivative results are normalized using Equation (9), and the contribution of each driving factor to \( Y \) is obtained. The modeling process of GWDF-ANN is shown in Figure 2.

\[
\Delta x_i = x_i \ast 0.001 \tag{7}
\]

\[
G_{x_i} = \frac{Y_{xi} - Y_0}{\Delta x_i} \tag{8}
\]

\[
W_{x_i} = \frac{G_{x_i}}{\sum_{i=1}^{n} G_{x_i}} \tag{9}
\]

where \( x_i \) is the value of the driving factor \( i \), \( \Delta x_i \) is the bias of \( x_i \), \( G_{x_i} \) is the partial derivative of \( x_i \), \( Y_{xi} \) is the prediction after adding a bias to \( x_i \), \( Y_0 \) is the prediction of original data, \( n \) is the number of driving factors, and \( W_{x_i} \) is the contribution of the \( x_i \) factor.

![Figure 2. Cont.](image-url)
Figure 2. The modeling process of GWDF-ANN. N1, N2, N3, and N4 are the multiple input neurons composing the input layer. The first rectangular box shows the model after the first driving factor has been added with a bias to $x_1 (x_1 + \Delta x_1)$. The $Y_{x1}$ represents the predicted result after adding a bias. The second rectangular box shows the model after the first driving factor has been added with a bias to $x_2 (x_2 + \Delta x_2)$. The $Y_{x2}$ represents the predicted result after adding a bias. We do not show all the driving factors. The last rectangular box shows the model after the first driving factor has been added with a bias to $x_5 (x_5 + \Delta x_5)$. The $Y_{x5}$ represents the predicted result after adding a bias.

3. A Case Study

3.1. Study Area

Shengli Coalfield is located in Xilinhot, Inner Mongolia, China (115°18′~117°06′ E, 43°02′~44°52′ N), as shown in Figure 3. It belongs to the mid-temperate semiarid continental monsoon climate. From 2004 to 2020, the monthly 24 h average temperature ranges from −18.4 °C in January to 22.4 °C in July, with an annual mean temperature of 2.76 °C. Most of the 250 mm (9.84 inches) annual rainfall occur in July and August. The average precipitation and temperature from January to December 2004–2020 are shown in Figure 4a. In winter, northwest winds predominate, while south and southeast winds are more prevalent in summer. Xilinhot’s usable pasture area is 1,369,000 hectares and is a significant processing and production base for green livestock products in China. The grasslands are composed of meadow grassland, typical grassland, and dune grassland. Xilinhot is also rich in mineral resources, with more than 30 billion tons of proven coal reserves.
The Shengli Coalfield is a lignite coalfield with the thickest coal seam, and it holds the largest reserves (estimated at 22.7 billion tons) in China. Its thick coal seam, shallow burial, and simple geological structure are suitable for centralized development. The coal quality is well suited for power generation, liquefaction, and chemical industries. The coalfield is located three kilometers north of Xilinhot. In 2004, Shengli Coalfield started the mining project and entered the large-scale construction phase in 2006.

A field survey was carried out in April 2020, in which drone images of the study area were taken and discussed with the residents. We found that the impact of mining activities on the vegetation two kilometers away from the coalfield boundary was minimal. While a larger area would have been more favorable, we were hampered by data acquisition constraints. In this paper, we selected a portion of the Shengli Coalfield to investigate the effects of mining and other factors on the vegetation using our proposed approach. We set a two-kilometer buffer and selected five directions (directions 1–5, Figure 3c) with uniform and random distribution to acquire more representative results. Given that the south of the Shengli Coalfield is mainly urban communities with little vegetation coverage, we did not
include the south direction in this study. The placement of direction 5 was determined by considering the land cover, which was incompatible with the other four. Each direction was set up with ten analysis units placed at 200 m intervals. GWDF-ANN models were constructed separately for each analysis unit.

3.2. Fractional Vegetation Coverage

In this study, FVC was used as a proxy for vegetation coverage change. FVC can be calculated by the Normalized Difference Vegetation Index (NDVI) using the pixel dichotomy model. First, Landsat Collection 2 surface reflectance products were accessed on Google Earth Engine (GEE) from 2004 to 2020. All Collection 2 surface reflectance products were created with a single-channel algorithm jointly produced by the Rochester Institute of Technology (RIT) and National Aeronautics and Space Administration (NASA) Jet Propulsion Laboratory (JPL). This 30 m spatial resolution product is derived from Landsat 5 TM and Landsat 8 OLI sensors and has undergone radiation calibration and atmospheric correction.

We selected satellite imageries from July 1 to September 30, when vegetation grows vigorously. The masking and the removal of clouds on the images were implemented on the GEE. Cloud removal was achieved using the QA quality band of Landsat and filtering the pixel values by bit manipulation. Masking was then used to remove clouds, cloud shadows, and snow pixels. The processed satellite images were synthesized using the maximum function in GEE. The maximum algorithm reduces image collection by calculating the maximum values at each pixel across the stack of all matching bands from July to September for each year. In other words, we used this reduction function for the images of every year. Years 2008, 2012, and 2018 were excluded due to excessive cloudiness. The information of satellite imageries is shown in Table 1.

We then calculated the NDVI for each pixel in each year using Equation (10). We corrected the NDVI for Landsat 5 TM to Landsat 8 OLI to resolve sensor differences between satellites. Previous studies have proposed methods to calibrate Landsat 5 TM, Landsat 7 ETM+, and Landsat 8 OLI [40–42]. We followed their procedure and selected some sample plots in Xilinhot to calibrate these sensors. We extracted the NDVI from TM and ETM+ for the same dates and performed regression analysis on the two NDVI groups to obtain the correction equation for TM to ETM+. The same method was used to calibrate ETM+ to OLI. Finally, the TM was corrected to OLI.

We then calculated the FVC according to the pixel dichotomy model using Equation (11). Pixel dichotomy assumes that ground pixels exist in a linear mixture of vegetation and soil. The pixel’s NDVI value is only affected by vegetation coverage. We selected plots that were completely bare. The NDVI pixel values from 2013 to 2020 were sorted in descending order, and the 95% percentile was used as NDVIsoil. The NDVIveg and NDVIsoil are stable from year to year with the long time series. We then selected some with complete vegetation cover during the peak growing season. The NDVI values in these plots were sorted in ascending order, and the 95% percentile was selected as NDVIveg. The long time series of pixel values can ensure the interannual stability of NDVIveg and NDVIsoil. The NDVIveg and NDVIsoil values are suitable since the different satellite sensors were corrected for consistency. The use of the 95% percentile pixel values reduces anomalous noisy pixels. The NDVIsoil was calculated to be 0.08, while NDVIveg was 0.7.

\[
\text{NDVI} = \frac{\rho_{\text{nir}} - \rho_{\text{red}}}{\rho_{\text{nir}} + \rho_{\text{red}}} \quad (10)
\]

\[
\text{FVC} = \frac{\text{NDVI} - \text{NDVI}_{\text{soil}}}{\text{NDVI}_{\text{veg}} - \text{NDVI}_{\text{soil}}} \quad (11)
\]

where \(\rho_{\text{nir}}\) is the surface reflectance in the near-infrared band, \(\rho_{\text{red}}\) is the surface reflectance in the red band, \(\text{NDVI}_{\text{soil}}\) is the NDVI for pure bare soil pixel, and \(\text{NDVI}_{\text{veg}}\) is the NDVI for pure vegetation pixel.
Table 1. Information of satellite imageries.

| Sensor       | Year | Path/Row  | Date         | Number |
|--------------|------|-----------|--------------|--------|
| Landsat 8 OLI| 2020 | 124/29    | 9 July       | 16     |
|              |      | 124/30    | 9 July       |        |
|              |      | 125/29    | 16 July      |        |
|              | 2019 | 124/29    | 7 July       |        |
|              |      | 124/30    | 7 July       |        |
|              |      | 125/29    | 14 July      |        |
|              | 2017 | 124/29    | 1 July       |        |
|              |      | 124/30    | 1 July       |        |
|              |      | 125/29    | 8 July       |        |
|              | 2016 | 124/29    | 14 July      |        |
|              |      | 124/30    | 14 July      |        |
|              |      | 125/29    | 5 July       |        |
|              | 2015 | 124/29    | 12 July      |        |
|              |      | 124/30    | 12 July      |        |
|              |      | 125/29    | 3 July       |        |
|              | 2014 | 124/29    | 9 July       |        |
|              |      | 124/30    | 9 July       |        |
|              |      | 125/29    | 16 July      |        |
|              | 2013 | 124/29    | 6 July       |        |
|              |      | 124/30    | 6 July       |        |
|              |      | 125/29    | 13 July      |        |
|              | 2011 | 124/29    | 1 July       |        |
|              |      | 124/30    | 1 July       |        |
|              |      | 125/29    | 8 July       |        |
| Landsat 5 TM | 2010 | 124/29    | 14 July      |        |
|              |      | 124/30    | 14 July      |        |
|              |      | 125/29    | 5 July       |        |
|              | 2009 | 124/29    | 11 July      |        |
|              |      | 124/30    | 11 July      |        |
|              |      | 125/29    | 2 July       |        |
Table 1. Cont.

| Sensor  | Year  | Path/Row | Date       | Date          | Number |
|---------|-------|----------|------------|---------------|--------|
| 2007    | 124/29| 6 July   | 22 July    | 7 August      | 24 September |
|         | 124/30| 6 July   | 22 July    | 7 August      | 24 September |
|         | 125/29| 13 July  | 30 July    | 30 August     | 15 September |
|         | 124/29| 3 July   | 29 July    | 4 August      | 20 August    |
|         | 124/30| 3 July   | 29 July    | 4 August      | 20 August    |
|         | 125/29| 10 July  | 27 July    | 27 August     | 12 September |
|         | 124/29| 16 July  | 17 August  | 2 September   | 18 September |
|         | 124/30| 16 July  | 17 August  | 2 September   | 18 September |
|         | 125/29| 7 July   | 23 August  | 7 August      | 24 August    |
|         | 124/29| 13 July  | 29 July    | 14 August     | 30 August    |
|         | 124/30| 13 July  | 29 July    | 14 August     | 30 August    |
|         | 125/29| 4 July   | 20 July    | 5 August      | 21 August    |
3.3. Driving Factors

For this study, temperature, precipitation, topography, mining, and urban expansion were selected as driving factors. Precipitation and temperature data were obtained from the China Meteorological Data Network (http://data.cma.cn (accessed on 1 January 2021)) (station number 54,102). We collected the monthly cumulative precipitation (unit: mm) and monthly average temperature (unit: °C) for Xilinhot in 2004–2020 (except 2008, 2012, and 2018) from station number 54,102. Numerous studies have shown that vegetation response to climate and precipitation may exhibit varying time lag effects for different regions [43,44]. It remains unclear how precipitation and temperature correlate with the FVC in the study area at different months of the year. Therefore, we performed a Pearson correlation analysis on temperature, precipitation, and FVC for each month. The correlation coefficients between FVC and precipitation and temperature are presented in Figure 4b. The results show that the cumulative precipitation from July to September and the average temperature from July to September have the highest correlations with the FVC at 0.660 and −0.616, respectively. Hence, they were selected as driving factors.

The topography data were derived from the digital elevation model (DEM) of the Shuttle Radar Topography Mission (SRTM), released by NASA in 2014. The spatial resolution is 30 m. Since topography does not significantly change year to year and annual topography data is unavailable, the same topography data was used for each year (see Figure 5). The mining factor was determined using the annual coal production, and the shortest Euclidean distance between the grid cell and the mining boundary and is calculated by Equation (12). Aside from the Shengli Coalfield, we also accounted for the effect of the West No. 2 Mine. This Euclidean distance refers to the shortest distance from each grid cell to each mine boundary. Coal production refers to the total output of the Shengli Coalfield and the West No. 2 Mine. The urban expansion was quantified using the annual urban population of Xilinhot and the shortest Euclidean distance between the grid cell and the mining boundary, as shown in Equation (13).

![Figure 5. The spatial distribution of topography.](image)

Annual coal production was provided by the National Energy Investment Group Co., Ltd. (Beijing, China). Coal production did not increase annually. The annual coal production from 2004 to 2020 is shown in Figure 6a. The urban population was obtained from the Xilinhot Statistical Yearbook, while the mining and urban boundaries were generated using Landsat imageries. In this study, we quantified mining and urban expansion from 2004 to 2020 (except 2008, 2012, and 2018). The quantitative results for the mining factor in 2020...
are presented in Figure 6b. The results for the urban expansion factor in 2020 are shown in Figure 7.

\[ X_{\text{mine}} = \frac{MA}{ED_{\text{mine}} + 1} \]  
\[ X_{\text{urban}} = \frac{NP}{ED_{\text{urban}} + 1} \]

where \( MA \) is the mined amount (unit: \( 10^4 \) m\(^3\)), \( ED_{\text{mine}} \) is the shortest Euclidean distance between the grid cell and the boundary of the mining area (unit: km), \( NP \) is the urban population of Xilinhot, and \( ED_{\text{urban}} \) is the shortest Euclidean distance between the grid cell and the urban boundary (unit: km).

**Figure 6.** (a) Coal production from 2004 to 2020; (b) quantitative results of the mining factor in 2020 (normalized).

**Figure 7.** Quantitative results of the urban expansion factor in 2020 (normalized).
3.4. Model Building

We set a two-kilometer buffer in the study area and delineated the buffer zone uniformly and randomly in five directions (directions 1–5, Figure 3c). Each direction had ten analysis units placed at 200 m intervals. The analysis unit was composed of 81 grid cells, separated by 30 m. Each unit was a circular area with a 90 m radius and was used to train a model; 50 driving models were established.

Before model training, we normalized the FVC, temperature, precipitation, topography, mining, and urban expansion factors for 2004–2020 (except 2008, 2012, and 2018) using Equation (14). The normalized data were then extracted to the corresponding grid cells (y-value and x-values). In this case, the input values were mapped on a 0–1 range with no dimension.

\[
Z_i = \frac{x_i - \max_{1 \leq i \leq n} (x_i)}{\max_{1 \leq i \leq n} (x_i) - \min_{1 \leq i \leq n} (x_i)}
\] (14)

where \(Z_i\) is the normalized value of the \(Z\) factor in grid cell \(i\), \(x_i\) is the original value of the grid cell \(i\), and \(n\) is the number of grid cells.

The 15-year grid cells (y-value and x-values) for each analysis unit were input into the model as the input layer. The FVC value (Y_0) was predicted for each analysis unit. The input data were then modified. Biases were added to the five driving factors (\(x_i + \Delta x_i\)) and were input back into the model to predict the FVC value (Y_{xi}) for each driving factor. Finally, the GWDF-ANN method was used to obtain the contribution of each grid cell. The average value of all the grid cells in each analysis unit was calculated to determine the contribution of the analysis unit. The procedure was repeated for each analysis unit to obtain the contribution (W) of the different analysis units. Each analysis unit uses an independent GWDF-ANN model.

4. Results

4.1. Modeling Results and Accuracy

Models were generated for the different analysis units. We selected the predicted FVC results (Y_0) of the first analysis unit in each direction and compared them with the actual FVC results (y-value), as shown in Figure 8. The predicted FVC values are similar to the actual results. RMSE and MRE were used as the model evaluation indicators, and the calculation methods are shown in Equations (15) and (16). The RMSE and MRE are typical accuracy evaluation indicators widely used in evaluating modeling accuracy in geospatial modeling [45,46]. The results of the RMSE and MRE are shown in Table 2. The average RMSE for the 50 groups of models is 0.052. The minimum RMSE for analysis unit 16 is 0.007, and the highest RMSE for analysis unit 10 is 0.112. The average MRE value is 0.007. The lowest MRE for analysis unit 6 is 0.001, while the highest MRE for analysis unit 46 is 0.023. The results suggest that the GWDF-ANN model is suitable for quantifying the change of FVC in mining areas.
4.2. FVC Spatial Changes and Quantitative Results of Driving Factors

Figure 9 shows the spatial distribution of FVC values within two kilometers from the Shengli Coalfield. The FVC values are in the 0–1 range; the higher the value, the greater the vegetation coverage. From 2004 to 2010 (except 2008), FVC decreased annually. The average FVC was 0.48, 0.40, 0.37, 0.34, 0.34, and 0.30, decreasing by 37.02% in six years. The FVC had a short-term increase from 2011 to 2013 and then decreased from 2013 to 2019. The lowest FVC value was 0.30, occurring in 2010, and the highest was 0.63 in 2020.

We calculated the average contributions of all the analysis units in the five directions. We found that the most prominent driving factor for FVC is precipitation (32.45%), followed by temperature (24.80%), mining (22.44%), and urban expansion (14.44%). The topography had the lowest contributions, with 5.87%. When combined, precipitation and temperature drive more than half (57.25%) of the FVC changes.
We then computed the average factor contributions of the ten analysis units for a given year at varying distances from the boundary of the mining area, and the results are displayed in Figure 10. The contribution of each driving factor shows similar characteristics. Over the years, the contributions of the different driving factors constantly changed. Precipitation and temperature have a downward trend, while mining and urban expansion have an increasing trajectory. The topography contributions remained unchanged.

Figure 10. The contribution of driving factors from 2004 to 2020. (a–e): Directions 1–5. Topo represents the factor of topography; Pre represents the factor of precipitation; Temp represents the temperature factor; Mine represents the factor of mining, and Urban represents the factor of urban expansion.
To compare the contribution of a particular driving factor in different years, take direction 1 as an example. Precipitation had the highest contribution at 37.28% in 2004 and the lowest contribution in 2019 at 30.68%. Temperature’s contribution was at 25.37% in 2004 and it contributed the least in 2020 at 21.88%. The mining effect increased year by year, closely following temperature. Temperature’s contributions in 2016, 2017, and 2020 were comparable. The trend suggests that temperature has become the second major driving factor. The contribution of mining from 2004 to 2020 increased from 18.38% to 22.23%, in which the fastest growth occurred in 2010–2011, rising by 10.04 percentage points. Urban expansion had a relatively flat growth trend and had declined in specific years. Still, urban expansion’s overall contributions have grown over the years, increasing from 12.73% in 2004 to 15.77% in 2020. For topography, the interannual contributions had very little change, only fluctuating within a 1% variation range.

The interannual contributions of each driving factor in the five directions had similar fluctuation characteristics, with the fifth direction being slightly different. The contribution of mining in direction 5 increased significantly interannually. In 2004, the major contributing factors were precipitation (32.42%), mining (25.47%), and temperature (24.75%). For 2020, mining (31.52%), precipitation (28.75%), and temperature (21.43%) were the main contributors. During the early stages of mining in 2004, its contribution was the second leading driving factor, surpassing that of temperature. However, over the years, temperature’s role gradually diminished while mining’s influence increased. From 2004 to 2020, contributions from mining increased from 25.47% to 31.52%, growing by 23.77 percentage points. After 2016, mining surpassed precipitation and became the leading driving factor.

Interannually, the contributions from precipitation and temperature show a fluctuating trend. The increase in the contribution of the precipitation is accompanied by the decrease in the influence of temperature. Likewise, the reduction in the contribution of precipitation is accompanied by the increase in the contribution of temperature. In direction 4, the contribution of precipitation showed a downward trend in 2004–2006, 2014–2017, and 2019–2020, and an upward trend in other years. Temperature showed an upward trajectory in 2004–2006, 2014–2017, and 2019–2020, and had a downward trend in other years.

We then took the average contribution of each factor in the ten analysis units at varying distances from the boundary of the mining area in the same direction. We calculated the change in the contribution of each factor in different directions, and the results are shown in Figure 11. The contribution of driving factors in direction 5 is noticeably different from other directions. In direction 5, the contribution of mining increased significantly, while the contributions of precipitation, temperature, urban expansion, and topography decreased.

![Figure 11. The contribution of driving factors in different directions.](image-url)
We computed the mean contribution of each factor in the ten analysis units at different distances from the boundary of the mining area in the same direction for 2004–2020 (except 2012 and 2018). We drew pie charts detailing the factor contribution of different directions (see Figure 12). Precipitation had the maximum contribution of 33.50% in direction 1 and a minimum of 30.94% in direction 5. For temperature, the maximum contribution was in direction 2 at 25.61%, and the minimum in direction 5 at 22.76%. The maximum contribution of mining in direction 5 is 28.36%, and the minimum in direction 1 is 20.93%. Urban expansion had the largest share of contribution in direction 4 at 14.89% and the smallest in direction 5 at 12.88%. For topography, the maximum contribution was 6.13% in direction 1 and the minimum at 5.06% in direction 5.

![Pie charts showing factor contribution](image1)

**Figure 12.** The pie chart of driving factor contribution in different directions. (a–e): Directions 1–5. Topo represents the factor of topography; Pre represents the factor of precipitation; Temp represents the factor of temperature; Mine represents the factor of mining, and Urban represents the factor of urban expansion.

The contributions of the various driving factors varied for the different directions. In directions 1–3, the order of driving factors from highest to lowest contribution was precipitation, temperature, mining, urban expansion, and topography. In direction 4, the ranking from highest to lowest was precipitation, temperature, mining, topography, and urban expansion. In this direction, the topography influence exceeded that of urban expansion. In direction 5, the order from highest to lowest contribution was precipitation, mining, temperature, urban expansion, and topography. Mining surpassed temperature to become the number two leading driving factor, second only to precipitation.

### 4.3. Contribution Analysis of the Mining Factor

We took the average value of mining contribution for each analysis unit from 2004 to 2020 (except 2008, 2012, and 2018) and analyzed the mining contributions at varying distances from the Shengli Coalfield boundary. We plotted the mean values for the different directions, and the results are presented in Figure 13. In all five directions, the contribution of mining showed apparent distance attenuation. As the distance from the boundary of the Shengli Coalfield increased, the contribution from mining exhibited a downward
trend. This means that the farther away a point is from the mining area, the lower the contribution of mining. The maximum contribution of mining in the five directions is 26.69% at 200 m from the mining area boundaries, which drops to 17.8% at 2000 m, decreasing by 33.31%. As the distance from the mining area boundary increases from 200 m to 2000 m, the contribution of mining in directions 1–5 decreases by 0.09, 0.10, 0.09, 0.08, and 0.08, respectively.

Figure 13. The contribution of the mining factor at different distances from the boundary of Shengli Coalfield.

5. Discussion

The GWDF-ANN method proposed in this study quantifies the contribution of driving factors to changes in the FVC. In particular, this study explores the impact of mining on FVC. Some results need further explanation and discussion.

The main driving factor of FVC around the mining area is not mining but precipitation, followed by temperature. In this study, the average contributions of precipitation and temperature from 2004 to 2020 in all analysis units were 32.45% and 24.80%, respectively. Since the study area was set as a buffer zone, two kilometers away from the mine boundary, the spatial heterogeneity of temperature and precipitation was almost absent. Therefore, the temperature and precipitation driving factors for each analysis unit were the same station data. The differences in temperature and precipitation were mainly between years. For example, the cumulative precipitation from July to September was high in 2004 and 2020 (274.3 mm and 268.4 mm) and low in 2014 and 2017 (93.9 mm and 93.1 mm). These values are consistent with our results, as in direction 1, the contribution of precipitation factor was high in 2004 and 2020 (37.28% and 33.27%) and low in 2014 and 2017 (32.07% and 32.93%). We found that more than 50% of the vegetation changes had been driven by precipitation and temperature and that the sum of their contributions reached 57.25%. Many studies have shown that vegetation change is very responsive to precipitation and temperature [47–50], consistent with the results of this study. In arid and semi-arid grasslands, the contribution of precipitation on FVC is significantly greater than that of temperature.

In addition to climate factors, this study quantified the impact of mining and urban expansion. Hui and Aman Fang found that mining significantly degrades vegetation coverage [51,52], but its specific contribution to vegetation coverage change has not been thoroughly explored. From 2004 to 2020, the contribution of mining within two kilometers around the boundary of the Shengli Coalfield was 22.44%. Mining causes severe degradation in groundwater levels and noticeable soil desertification, resulting in harsh environments for vegetation growth. Coal mine dust falling on plant leaf surface may
also affect plant growth and cause other serious problems. Around the mining area, FVC degradation caused by mining is apparent.

The city’s rapid expansion and industrialization have resulted in decreased FVC. From 2004 to 2020, the contribution rate of urban expansion within two kilometers from the Shengli Coalfield is 14.81%. From 2004 to 2020, the urban area of Xilinhot expanded by 129.184 km², and the urban population increased by 42,697 people. Urban expansion reduces the dependence of plant growth on climate factors and has a particular impact on FVC, consistent with Song’s and Noa’s findings [53,54].

As for the topography, its influence on FVC is relatively small, and its interannual contribution remains unchanged. The topography showed little change from year to year, and its differentiation is mainly spatial. There is a 45 m elevation difference in the study area, high in the west and low in the east. However, the impact of topographic factors does not significantly change in different directions. This may be because the elevation of the mining area does not have a particular contribution to the vegetation.

The effects of mining have noticeable spatial and temporal variations. Temporally, the value of the mining factor varies with the coal production. The findings also show that the contribution of mining to FVC continued to increase annually, even as coal production decreased in certain years. In 2009, 2011–2015, and 2019, while coal production in the Shengli Coalfield declined from the previous year, the contribution of mining did not decrease. Hence, the results suggest that the impact of mining on FVC may have a cumulative effect in time. Years of continuous mining have exacerbated the adverse effects of mining operations on FVC.

Spatially, the further away from the mine boundary, the lower the value of the mining driving factor. The contribution of mining shows distance attenuation around the mining area. Within 200 m from the periphery of the mining area, mining’s contribution was 26.69%; at 2000 m, it dropped to 17.84%. As the distance from the mining area increases, the contribution of mining decreases until it reaches some point where the value is zero. This suggests that the impact of mining on FVC may be limited. For example, Aman Fang studied the influence area of Bao’s coal in eastern Inner Mongolia [51]. The scope of mining influence on FVC needs to be further explored.

The contribution of mining in direction 5 was significantly higher than in other directions. This may be because the Shengli Coalfield is close to the West No. 2 Mine (about three kilometers away), affecting the FVC. The West No. 2 Mine started operations in 2006, and by 2020, the cumulative coal produced was 825,110,800 m³, about 0.91 times the cumulative coal production of the Shengli Coalfield. The high mining impact in direction 5 was likely caused by the cumulative effect of the Shengli Coalfield and the West No. 2 Mine.

Changes in the FVC are caused by the coupling of multiple factors. We only considered the contribution of precipitation, temperature, mining, urban expansion, and topography. Aside from these factors, overgrazing may have a considerable impact on the FVC [55]. Ecological environment restoration policies such as vegetation reclamation have also been found to significantly affect the FVC in mining areas [56,57]. Subsequent studies can analyze other parameters affecting vegetation coverage in mining areas, including grazing activities, environmental protection policies, evaporation, and soil types. In addition, there are currently few quantitative studies on mining and urban expansion. How to scientifically quantify the impact of mining and urban expansion requires further exploration.

6. Conclusions

This study proposed the GWDF-ANN to quantify the contributions of different driving factors on FVC changes in mining areas. Based on the results, some conclusions were reached, as follows.

1. For the 50 models, the average RMSE was 0.052 and the average MRE was 0.007. The GWDF-ANN model is suitable for quantifying FVC changes in mining areas.
(2) Precipitation and temperature were the main driving factors for FVC change. The contributions were 32.45% for precipitation, 24.80% for temperature, 22.44% for mining, 14.44% for urban expansion, and 5.87% for topography.

(3) The contributions of precipitation and temperature on vegetation cover exhibited downward trends, while mining and urban expansion showed positive trajectories. For topography, its contribution remains generally unchanged.

(4) The contribution of mining showed apparent distance attenuation. At 200 m away, the contribution of mining was 26.69%; at 2000 m away, the value drops to 17.8%.

(5) Mining has a cumulative effect on vegetation coverage both interannually and spatially. In the future, more driving factors, such as grazing and soil quality, can be considered to improve model accuracy.

Author Contributions: Conceptualization, J.L. and C.Z. (Chengye Zhang); methodology, J.L., C.Z. (Chengye Zhang) and T.Q.; validation, T.Q., H.Z. and J.G.; formal analysis, J.L. and T.Q.; investigation, J.G., H.X., C.Z. (Caiyue Zhang) and Y.Z.; writing—original draft preparation, J.L. and T.Q.; writing—review and editing, T.Q. and C.Z. (Chengye Zhang); visualization, T.Q. and H.Z.; supervision, T.Q. and H.Z.; project administration, C.Z. (Chengye Zhang); funding acquisition, C.Z. (Chengye Zhang). All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the State Key Laboratory of Water Resource Protection and Utilization in Coal Mining (Grant number GJNY-20-113-14), National Natural Science Foundation of China (Grant number 41901291), Fundamental Research Funds for the Central Universities (Grant number 2021YQDC02).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The Landsat data and topography data can be downloaded from GEE (https://developers.google.com/earth-engine/datasets/, accessed on 15 November 2021). And the precipitation and temperature data can be downloaded from the China Meteorological Data Network (http://data.cma.cn, accessed on 15 November 2021).

Acknowledgments: The authors express gratitude to the National Energy Investment Group Co., Ltd. (Beijing, China) for providing the necessary information and data.

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

References
1. Bai, X.; Ding, H.; Lian, J.; Ma, D.; Yang, X.; Sun, N.; Xue, W.; Chang, Y. Coal production in China: Past, present, and future projections. Int. Geol. Rev. 2017, 60, 535–547. [CrossRef]
2. Wang, Y.J. Research Progress and Prospect on Ecological Disturbance Monitoring in Mining Area. Acta Geod. Cartogr. Sin. 2017, 45, 1705–1716.
3. Liu, S.J.; Wang, Z.; Mao, Y.C.; Xu, B.S.; Wu, L.X. Multi-source Remote Sensing Technology for Monitoring Safety and Environment in Mine. Geomat. Spat. Inf. Technol. 2015, 38, 98–100.
4. Wu, Z.; Lei, S.; Lu, Q.; Bian, Z. Impacts of Large-Scale Open-Pit Coal Base on the Landscape Ecological Health of Semi-Arid Grasslands. Remote Sens. 2019, 11, 1820. [CrossRef]
5. Qian, T.; Bagan, H.; Kinoshita, T.; Yamagata, Y. Spatial–Temporal Analyses of Surface Coal Mining Dominated Land Degradation in Holingol, Inner Mongolia. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2014, 7, 1675–1687. [CrossRef]
6. Parmesan, C.; Yohe, G. A globally coherent fingerprint of climate change impacts across natural systems. Nature. 2003, 421, 37–42. [CrossRef] [PubMed]
7. Chen, L.F.; Zhang, H.; Zhang, X.Y.; Liu, P.H.; Zhang, W.C.; Ma, X.Y. Vegetation changes in coal mining areas: Naturally or anthropogenically Driven? Cattota 2022, 208, 105712. [CrossRef]
8. Hao, L.; Sun, G.; Liu, Y.; Gao, Z.; He, J.; Shi, T.; Wu, B. Effects of precipitation on grassland ecosystem restoration under grazing exclusion in Inner Mongolia, China. Landsc. Ecol. 2014, 29, 1657–1673. [CrossRef]
9. Chuai, X.W.; Huang, X.J.; Wang, W.J.; Bao, G. NDVI, temperature and precipitation changes and their relationships with different vegetation types during 1998–2007 in Inner Mongolia, China. Int. J. Climatol. 2013, 33, 1696–1706. [CrossRef]
10. Pang, J.; Du, Z.Q.; Zhang, X.Y. Understanding of the relationship between vegetation change and physical geographic factors based on Geographical Detector. In Proceedings of the International Conference on Intelligent Earth Observing and Applications (IEOAs), Guilin, China, 9 December 2015; pp. 23–24.
40. Flood, N. Continuity of Reflectance Data between Landsat-7 ETM+ and Landsat-8 OLI, for Both Top-of-Atmosphere and Surface Reflectance: A Study in the Australian Landscape. Remote Sens. 2014, 6, 7952–7970. [CrossRef]

41. Li, P.; Jiang, L.; Feng, Z. Cross-Comparison of Vegetation Indices Derived from Landsat-7 Enhanced Thematic Mapper Plus (ETM+) and Landsat-8 Operational Land Imager (OLI) Sensors. Remote Sens. 2013, 6, 310–329. [CrossRef]

42. She, X.; Zhang, L.; Cen, Y.; Wu, T.; Huang, C.; Baig, M. Comparison of the Continuity of Vegetation Indices Derived from Landsat 8 OLI and Landsat 7 ETM+ Data among Different Vegetation Types. Remote Sens. 2015, 7, 13485–13506. [CrossRef]

43. Bai, S.Y.; Wang, L.; Shi, J.Q. Time lag effect of NDVI response to climatic change in Yangtze River basin. Chin. J. Agrometeorol. 2012, 33, 579–586.

44. Wang, J.B.; Yang, X.L.; Zhang, Y.H.; Wang, X.W. Correlation Between NDVI and Meteorological Factors in Zhangye. Chin. Agric. Sci. Bull. 2019, 35, 85–90.

45. Liu, Y.L.; Guo, L.; Jiang, Q.H.; Zhang, H.T.; Chen, Y.Y. Comparing geospatial techniques to predict SOC stocks. Soil Tillage Res. 2015, 148, 46–58. [CrossRef]

46. Arora, A.; Pandey, M.; Mishra, V.N.; Kumar, R.; Rai, P.K.; Costache, R.; Punia, M.; Di, L. Comparative evaluation of geospatial scenario-based land change simulation models using landscape metrics. Ecol. Indic. 2021, 128, 107810. [CrossRef]

47. Chen, Z.F.; Wang, W.G.; Fu, J.Y. Vegetation response to precipitation anomalies under different climatic and biogeographical conditions in China. Sci. Rep. 2020, 10, 1–16. [CrossRef]

50. Zhang, Y.; Liang, W.; Liao, Z.; Han, Z.; Xu, X.; Jiao, R.; Liu, H. Effects of climate change on lake area and vegetation cover over the past 55 years in Northeast Inner Mongolia grassland, China. Theor. Appl. Climatol. 2019, 138, 13–25. [CrossRef]

51. Fang, A.; Yao, M.; Chen, W.; Dong, J. Assessment of Surface Ecological Quality of Grassland Mining Area and Identification of Its Impact Range. Nat. Resour. Res. 2021, 30, 3819–3837. [CrossRef]

52. Hui, J.; Bai, Z.; Ye, B.; Wang, Z. Remote Sensing Monitoring and Evaluation of Vegetation Restoration in Grassland Mining Areas—A Case Study of the Shengli Mining Area in Xilinhot City, China. Land 2021, 10, 743. [CrossRef]

53. Ohana-Levi, N.; Paz-Kagan, T.; Panov, N.; Peeters, A.; Tsoar, A.; Karnieli, A. Time series analysis of vegetation-cover response to environmental factors and residential development in a dryland region. Gisci. Remote Sens. 2019, 56, 362–387. [CrossRef]

54. Song, X.P.; Hansen, M.C.; Stehman, S.V.; Potapov, P.V.; Tyukavina, A.; Vermote, E.F.; Townshend, J.R. Author Correction: Global land change from 1982 to 2016. Nature 2018, 563, E26. [CrossRef]

55. Wang, Z.; Deng, X.; Song, W.; Li, Z.; Chen, J. What is the main cause of grassland degradation? A case study of grassland ecosystem service in the middle-south Inner Mongolia. Catena 2017, 150, 100–107. [CrossRef]

56. Xu, G.; Zhang, J.; Li, P.; Li, Z.; Lu, K.; Wang, X.; Wang, F.; Cheng, Y.; Wang, B. Vegetation restoration projects and their influence on runoff and sediment in China. Ecol. Indic. 2018, 95, 233–241. [CrossRef]

57. Jiang, G.; Han, X.; Wu, J. Restoration and management of the Inner Mongolia grassland require a sustainable strategy. AMBIO A J. Hum. Environ. 2006, 35, 269–270. [CrossRef]