Research on the Assessment Technology of Land Available for PV

Jiaheng Li 2, Mengjin Hu 1, Hongkui Li 2, Mengju Wei 1, Zhao Liu 1, Xuefei Liu 1

1 State Grid Hebei Economic Research Institute, HeBei, 050000, China
2 State Grid Shijiazhuang Electric Power Supply Company, HeBei, 050000, China

Abstract. Based on the high-resolution satellite image data, the information mining technology of the available surface elements of PV is studied, and the investigation of the available surface elements of PV in 98 counties and cities of South Hebei grid is realized. Based on the large-scale and high-resolution remote sensing data obtained by multi-source remote sensing data fusion technology, the depth-learning-based surface feature recognition technology for photovoltaic development is studied. Based on the method of automatic identification and artificial combination of depth-learning, it can identify the available ground elements (roof, water surface, road surface, dry beach, etc.), the available surface elements of PV in 98 counties and cities of Hebei South Grid were obtained. From the overall point of view, the photovoltaic land, the building occupies the main position, in the four cities are relatively high, are in the 6% ~ 15%

Keywords: Photovoltaic, Land type, Assessment of land resources

1. Introduction

By the end of 2017, China had installed 130 gigawatts of photovoltaic (PV) power, but the country's solar resources and electricity consumption are extremely uneven, leading to a persistent problem of absorption, to this end, the state issued a number of targeted policies and measures to encourage the development of distributed power plants in the central and eastern regions. In the Central and eastern regions of China, where land for photovoltaic construction is relatively scarce, how to comprehensively evaluate the potential of photovoltaic power generation in light of solar energy resources, considering the constraints of topography, landforms and the nature of land use, and regarding the government's industrial policy designation, and market investment behavior will have great practical significance. There have been a number of domestic studies on the potential of photovoltaic power generation. Zhou Yang et al have analysed the spatial and temporal distribution of solar energy resources in Jiangsu province and estimated the amount of rooftop installed capacity in the province, the potential of solar energy resources in Jiangsu province is analyzed in terms of the abundance and stability of solar energy resources, and the spatial and temporal distribution characteristics of solar energy resources in Shanxi province are analyzed by Yan Jiahai, the potential of solar energy resources in Shanxi province is analyzed in terms of solar energy available days, stability and optimal utilization period Liang Yulian analyzed the spatial and temporal distribution characteristics, abundance and stability of solar energy resources in South China, and discussed the suitability of solar energy development under different landforms and underlying surfaces Li Ke et al selected 3 factors of total solar radiation, sunshine duration and effective sunshine days to analyze the land solar energy resource potential in China by using multi-index evaluation method By analyzing the conversion efficiency of monocrystalline silicon modules under different irradiance in Shanghai area, Wang calculated the spatial distribution of PV theoretical annual power generation in China Xie Guohui and others introduced the development potential assessment system of PV power station bases, and Marcel and others used PVGIS data sets to estimate the annual power generation of PV systems in EU member states, the relationship between PV potential and power load in different countries is analyzed. For the potential of photovoltaic power generation, the above research work is involved in different fields, but there are still some problems in the content, such as the lack of pertinence or the lack of comprehensive consideration of factors, so it is very meaningful to explore a comprehensive quantitative evaluation method.Principle of Power Line Communication and Its Application in Energy Internet
2. Investigation of available surface elements of PV

2.1 Calibration and stitching of remote sensing images

This project will use ERDAS remote sensing image processing software, through the geometric correction module of the software, take the State Bureau of Surveying and Mapping data as the control image, and carry out the geometric precision correction of the satellite image, the correction accuracy is less than 1 pixel. After geometric fine correction, the image is a control image, and the registration accuracy is less than 0.5 pixels. In order to ensure the quality of the corrected image and the convenience of the follow-up processing, the following requirements should be followed in the calibration:

The requirements of the control points: the three-dimensional coordinates of the control points should be measured in the field, the points should be set at the obvious corners, the control points should be distributed evenly, and they should be selected in a stable and unchangeable place. The control point used for orthophoto correction of the reference image and the monitoring image should be the same as far as possible, so as to reduce the deviation of two-phase image collocation;

1) The requirements of the correction model should be consistent with the mathematical model used for two-phase image correction;
2) The selection of resampling mode mainly includes bilinear interpolation and bicubic convolution. The information fidelity in the resampling process is higher by using bicubic convolution than by using bilinear interpolation. With the replacement of the computer, the computing speed is no longer a problem. The practice shows that the operation time of the two resampling modes is almost the same, but the image quality varies greatly, so the bicubic convolution method is usually required in the process of image resampling.

2.2 Research on multi-scale segmentation algorithm of remote sensing image

In this paper, height information and vegetation index are added to the existing multi-scale segmentation methods. In this study, DSM was used to represent height information, while NDVI was the most representative vegetation index. As well as spectral and shape information, height information, as a factor of heterogeneity, is called DSM heterogeneity. The image of NDVI is represented by the weight of NDVI, which is set according to different NDVI, and reflects the image of NDVI in the process of image segmentation. The formula for calculating the algorithm is as follows:

\[ f = w_{NDVI} \cdot h_{color} + w_{shape} \cdot h_{shape} + w_{DSM} \cdot h_{DSM} \]

Where \( w_{NDVI} \) describes the effect of NDVI; \( h_{color} \), \( h_{shape} \) and \( h_{DSM} \) for spectral, shape, and DSM heterogeneity, respectively. \( w_{color} \cdot w_{shape} \cdot w_{DSM} \) for the weight of spectral, shape, and DSM heterogeneity, respectively. DSM can effectively describe the complex building environment, the building is designed and built, not the natural formation, so it is unexpected and unpredictable. Compared with other types of ground features, the DSM of buildings is obviously different and has clear boundaries. The multi-scale segmentation method not only uses DSM as a layer to calculate the spectral heterogeneity, but also as a factor to determine whether the adjacent objects merge.

\[ h_{DSM} = D_{MDiff} + \left( P_{d1} + P_{d2} / E_{diff} \right) \]

\[ D_{MDiff} = \left| D_{mean1} - D_{mean2} \right| \]

\( D_{MDiff} \) is the difference between the average DSM of adjacent objects before merging, and \( E_{diff} \) is the number of adjacent pixels of adjacent objects. \( P_{d1} \) is the number of pairs of adjacent pixels in which the DSM difference is greater than 3, and \( P_{d2} \) is the total value of the DSM difference of adjacent pairs of adjacent pixels in which the DSM difference is greater. \( D_{mean1} \) and \( D_{mean2} \) are the average DSM values of adjacent objects, respectively. \( h_{DSM} \) is calculated specifically as shown in Figure 10, where two objects 1 and 2 are represented in green and yellow, respectively, and the numbers are the DSM values of the pixels, respectively, and the DSM averages of objects 1 and 2 are calculated, 1.07 and 4.27, respectively, and calculated \( D_{MDiff} \) as 3.2. The logarithm of the adjacent pixels of the two objects is then \( P_{d1} \) calculated to be 5 pairs. The number of pairs of adjacent pixels larger than 3 is 5, and the sum of DSM difference of these 5 pairs is 20, i.e. \( P_{d2} = 20 \).

\[ DSM_{MDiff} = 3.22 \quad E_{diff} = 5 \quad P_{d1} = 5 \]

\( DSM_{MDiff} \) is determined by the average NDVI of the two adjacent objects before the merge. The formula is as follows:

\[ NDVI_{mean} = (NIR_{mean} - R_{mean}) / (NIR_{mean} + R_{mean}) \]  \( (1) \)

\[ w_{NDVI} = 1, \quad NDVI_{mean1} \cdot NDVI_{mean2} > 0 \]

\[ w_{NDVI} = 0, \quad NDVI_{mean1} \cdot NDVI_{mean2} < 0 \]  \( (2) \)

Where, \( NDVI_{mean1} \) and \( NDVI_{mean2} \) are the average NDVI of the first two adjacent objects to be merged, respectively.

2.3 Target region extraction

BP neural network is used to extract the target area. BP neural network is a multi-layer feed-forward network. It uses error Backpropagation as learning rules to supervise learning. It trains a large number of condition sets of objects to be classified as known samples, find out the relationship between the feature attributes (patterns) of the input samples and the classification categories. BP
neural network has a simple structure and strong Operability. It has been widely used in image classification because of its mature technology and clear processing flow. Based on the selected samples, the BP neural network model is constructed by taking the extracted optimal feature set as the input of BP neural network to realize the recognition of surface elements. The BP (back propagation) neural network, a multi-layered Feedforward neural network trained by an error back propagation algorithm developed in 1986 by scientists led by Rumelhart and McClelland, is now the most widely used. The aim of BP neural network is to minimize the error mean square deviation between the actual output value and the expected output value of BP neural network. The calculation process of the basic BP algorithm includes forward calculation (i.e. the forward propagation of the signal) and backward calculation (i.e. the backward propagation of the error). In the forward propagation, the input signal acts on the output node through the hidden layer, and generates the output signal through the nonlinear transformation to judge whether the actual output matches the expected output or not, the error is propagated backward. Error back-propagation is to transfer the output error from the hidden layer to the input layer, and distribute the error to all units in each layer, so as to adjust the weight of each unit. The error decreases along the node gradient when the connection strength between the input node and the hidden node, the connection strength and the threshold value between the hidden node and the output node are adjusted. Then, through multiple training, to obtain the corresponding with the minimum error of the weight and threshold), stop the training.

The BP neural network classification model is trained according to the input of the sample of the image object selected artificially, the characteristic value of the image object is taken as the input, analyzes and judges whether it is a building, and outputs the result, if the error is large, then adjust the weight value to calculate the output again until the classification precision is satisfied. If the training results meet the requirements, the network model is used to extract the building. The specific training process is as follows:

1) Select a group of training samples, the image object's characteristic is the input information, whether it is the building as the expected output. The input data has M eigenvalues \( X = (x, x, ... X)^t \), the input layer has M neurons. The expected output is. Take each feature as each node of the input layer.
2) The input layer passes through each hidden layer in turn, and the output \( x = \gamma^L (L) , 1 = 1,2, ... L, L \) is the number of implied layers.
3) Reach the output layer, calculate the actual output of the network, and the expected output, the error. At this point, complete the entire network forward propagation.
4) The error of forward propagation calculation is propagated from the output layer, through the middle layer, to the first hidden layer, and distributed to the neurons of each layer according to different weights.
5) Combining with the error of each neuron, the connection weight of each neuron is revised repeatedly.
6) Repeat steps 2 and 3, terminate the training if the error meets the requirement, otherwise continue the iterative training until a satisfactory neural network model is obtained. In order to solve the problem that BP neural network is not easy to converge, the initial weights and thresholds of BP neural network are often initialized randomly. Setting appropriate learning rate and momentum term is also an effective method to improve the convergence rate of network training.

![Fig. 2 Algorithm flow](image)

3. Case Study

By analyzing the land-use situation of each city and counting the land resources available for solar energy, it can be seen that the main types of land-use in the cities in southern Hebei are cultivated land, grassland and urban buildings, however, the development and utilization value of cultivated land and grassland is low, and it does not have the condition of photovoltaic development and utilization. On the whole, the area of land available for PV in south Hebei network is 5,350 km², including 1,426.27 km² in Baoding, 1,220.55 km² in Cangzhou, 1,941.84 km² in Handan, 559.75 km² in Hengshui, 1,866.44 km² in Shijiazhuang and 1,514.81 km² in Xingtai. Among the four types of photovoltaic usable land, buildings occupy the main position, and are relatively high in the four cities, all in the 6% ~ 15%. The other three types of area, relative to the building, accounting for less, all three of the total area of the city to 1% of the total area. In terms of the area of land available for PV, Shijiazhuang and Handan have the largest area, at about 1,900 km², followed by Xingtai and Baoding at about 1,500 km², Cangzhou at about 1,220 km² and Hengshui at least, only about 560 km² of photovoltaic land is available.
### Table 1 Land use type in southern Hebei

<table>
<thead>
<tr>
<th>Cities</th>
<th>Buildings</th>
<th>Bare land</th>
<th>Roads</th>
<th>Dry beaches</th>
<th>Water bodies</th>
<th>Others</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baoding</td>
<td>1295.24</td>
<td>65.20</td>
<td>39.43</td>
<td>26.40</td>
<td>108.68</td>
<td>20701.05</td>
<td>22236</td>
</tr>
<tr>
<td>Cangzhou</td>
<td>1143.68</td>
<td>4.31</td>
<td>33.69</td>
<td>38.87</td>
<td>499.38</td>
<td>12394.07</td>
<td>14114</td>
</tr>
<tr>
<td>Handan</td>
<td>1856.21</td>
<td>32.93</td>
<td>40.62</td>
<td>12.08</td>
<td>26.19</td>
<td>10087.97</td>
<td>12056</td>
</tr>
<tr>
<td>Hengshui</td>
<td>517.96</td>
<td>2.68</td>
<td>25.87</td>
<td>13.24</td>
<td>37.04</td>
<td>8254.21</td>
<td>8851</td>
</tr>
<tr>
<td>Shijiazhang</td>
<td>1722.26</td>
<td>72.85</td>
<td>46.26</td>
<td>25.07</td>
<td>94.81</td>
<td>12072.75</td>
<td>14034</td>
</tr>
<tr>
<td>Xingtai</td>
<td>1421.17</td>
<td>33.96</td>
<td>41.75</td>
<td>17.93</td>
<td>22.47</td>
<td>10900.72</td>
<td>12438</td>
</tr>
</tbody>
</table>

### 4. Summary

In order to reduce the cost of PV Land Use Calculation, the project combines object-oriented and deep learning methods. Considering the different types of PV land use, the multi-scale segmentation method is adopted firstly, to segment the remote sensing image and lay a good foundation for the following land use extraction, then in order to give full play to the performance of the classifier, a feature selection method based on feature clustering is proposed, by selecting a small number of samples, the optimal feature set of land use extraction is selected, which contains enough category information. The precision of the buildings extracted with optimal features is similar to that of the buildings extracted with all features, but the buildings extracted with optimal features will reduce the time of calculating a large number of features and be more efficient. BP neural network is used to extract the available land resources of PV by using the optimal feature set, which provides a good data support for PV potential assessment.

### Acknowledgments

This research was supported by State Grid Hebei Economic Power Co, Ltd. Investigation on the potential of solar energy and solar energy in provinces based on high-resolution satellite images, Grant B704JY210065.

### References