Energy Engineering

Tech Science Press

DOI: 10.32604/ee.2022.018750

ARTICLE

Research on Spatial Statistical Downscaling Method of Meteorological Data Applied to Photovoltaic Prediction

Yan Jin[1,*], Dingmei Wang[2], Ruiping Zhang[1] and Haiying Dong[1]

[1]School of Automation and Electrical Engineering, Lanzhou Jiaotong University, Lanzhou, 730000, China

[2]Electric Power Research Institute of State Grid Gansu Electric Power Company, Lanzhou, 730000, China

[*]Corresponding Author: Yan Jin. Email: jinyan2206@163.com

Received: 14 August 2021 Accepted: 29 October 2021

ABSTRACT

Aiming at the low spatial resolution of meteorological data output from a numerical model in photovoltaic power prediction, a geographically weighted statistical downscaling method considers the influence factors such as normalized vegetation index (NDVI), digital elevation model (DEM), slope direction, longitude and latitude is proposed. This method is based on the correlation between meteorological data and NDVI, DEM, slope direction, latitude and longitude, and introduces DEM and local Moran index to improve the regression model, and obtains 100 ∗ 100 m high-resolution meteorological spatial distribution data. Finally, combining the measured data of the study area and the established EOF iterative downscaling method to verify and compare the downscaling results. The results show that the error between the downscaled meteorological data and the measured value is smaller, and the comprehensive downscaling accuracy of the geographically weighted regression method is higher, and the model fitting effect is better. Therefore, this method can effectively improve the influence of errors caused by lower resolution, and provide a more reliable meteorological basis for the prediction of photovoltaic power.

KEYWORDS

Numerical model; meteorological data; resolving power

1 Introduction

As of the end of 2020, the cumulative installed capacity of photovoltaic power stations in China had exceeded 253 GW, a year-on-year decrease of 31.6%. In 2020, photovoltaic power generation will be about 224.3 billion kWh, with a year-on-year increase of 26.3% [1]. With the continuous expansion and improvement of the scale of photovoltaic power stations in China, improving the prediction accuracy of photovoltaic power generation can provide strong technical support for power system dispatching management and multi-energy complementary coordinated control, which also helps to improve the operation stability of power system.

In recent years, a lot of research on improving the prediction accuracy of photovoltaic power has been carried out on the prediction method, and there are relatively few studies on meteorological data that have a great impact on the prediction of photovoltaic power. Literature [2] shows that high-precision numerical weather forecast data plays a vital role in photovoltaic power prediction.

The meteorological data currently used for photovoltaic power prediction has different temporal and spatial resolutions. The Global forecasting system (GFS) model of the US Meteorological Center provides two products of 1° × 1° (latitude and longitude) and 0.5° × 0.5°, while the product resolution of the GRAPES regional forecast model of the China Meteorological Administration has reached 10 km.

The downscaling method uses large-scale weather as a condition to directly convert low-resolution numerical forecast products into regional-scale ground weather change information, so as to achieve downscaling of meteorological data. For example, After comparing the accuracy of monthly precipitation data of Tropical rainfall measurement task (TRMM) and Global precision measurement (GPM), Sheng [3] selected a better random forest algorithm, established a downscaling model according to environmental variables, and obtained annual precipitation data with a resolution of 1 km. Through comparison, it is found that the data after downscaling is very close to the original TRMM Precipitation Data, and the correlation coefficient reaches 0.91; Zhang [4] constructed a statistical downscaling model by using TRMM Precipitation statistics in the study area, NDVI products of medium resolution imaging spectrometer (MODIS) and different terrain influence factors, and obtained the spatial data of 250 m high-resolution radar precipitation in the study area. The results show that the accuracy of the downscaling results is higher than that of the traditional data, and the annual average error is less than that of the original precipitation data; According to the temporal and spatial variation distribution characteristics of meteorology, Fan et al. [5,6] selected predictive factors such as enhanced vegetation index to build a spatial downscaling model of TRMM. The spatial resolution of TRMM data is improved from 0.25° to 1 km and verified by the measured values of meteorological stations in the study area. The results show that the precipitation data with 1 km resolution has high accuracy; Based on statistical and physical models such as random forest and polynomial fitting, Sheng et al. [7] integrated visible light, thermal infrared and surface high-frequency parameters, reduced the microwave soil moisture data of FY-3 satellite (FY3b), and improved its spatial resolution from 25 to 1 km. The verification results show that the overall scale accuracy of the random forest method is the highest and the model fitting effect is the best; In 2014, Ji et al. [8] took Sichuan and Chongqing as the research areas, used regression statistical rules to downscale their TRMM satellite precipitation data, and obtained satellite annual and monthly precipitation grid data with a high spatial resolution (1 km), so as to better monitor and study the drought in Sichuan and Chongqing; in addition, Merlin et al. proposed a downscaling method based on empirical polynomial fitting [9,10], which used low spatial resolution surface temperature, vegetation index, and surface albedo data to construct polynomial functions. The downscaled soil moisture, vegetation index, and surface albedo data were then derived using high-resolution surface temperature, and the method was previously applied to soil moisture products to obtain better-downscaled results; Weng et al. [11–14] proposed a STARFM (Spatio-temporal adaptive reflectance fusion model) data fusion model to reduce the resolution of thermal infrared data (1 km) by fusing MODIS and Landsat data down to 120 m, and then performed surface temperature inversion to finally obtain high spatial resolution surface temperature with good results; Nerini et al. [15,16] compared and analyzed the effect of precipitation data fusion with mean deviation correction, dual kernel smoothing, external drift kriging and Bayesian fusion in Peru. The results show that the application effectiveness of different precipitation fusion methods is affected by the density of ground rainfall stations in the study area, and the dual kernel smoothing method has good applicability in the case of sparse ground rainfall stations.

The above literature provides a good reference for spatial downscaling methods for meteorological data, but most of the above studies focus on plains or basins, where the factors affecting downscaling are single and the instantaneous variability is low. Because the underlying surface is relatively simple,

it can not meet the requirements of spatial downscaling under complex terrain. In order to better explore the downscaling characteristics of meteorological data and the impact of complex terrain on meteorological data, and improve the prediction accuracy of photovoltaic power generation. In this paper, a geographically weighted regression downscaling method considering the influencing factors such as NDVI, DEM and slope direction is proposed. The introduction of DEM and local Moran index can better reduce the error of meteorological data under complex terrain, make the downscaling of meteorological data closer to measured data in this area, and provide a guarantee for the application of high-precision meteorological data in photovoltaic power generation prediction. The effectiveness of this method is verified by experimental comparison and analysis.

2 Principle of Statistical Downscaling

Statistical downscaling is mainly based on historical observation data to establish the statistical relationship between large-scale meteorological elements and regional meteorological elements (Fig. 1), and then use the test data to verify this relationship. Finally, this relationship is applied to the large-scale numerical prediction products output by the local weather prediction numerical model. In other words, the statistical function relationship between the influencing factors of large-scale forecast and the regional weather forecast is established:

$$Y = F(X) \tag{1}$$

where X represents the large-scale weather forecast factor, Y represents the regional weather forecast variable, and F is the established statistical relationship between the large-scale weather forecast factor and the regional weather forecast variable. Generally speaking, F is unknown and needs to be obtained through statistical methods (determined by observational data).

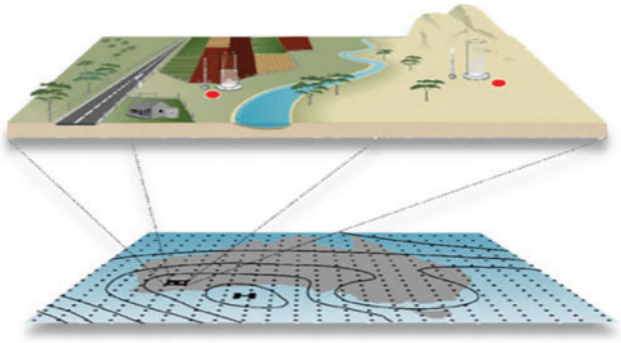


Figure 1: Downscaling diagram

The advantage of statistical downscaling method is that it can apply numerical prediction products with greater physical significance and more accurate simulation to the statistical model, so as to correct the systematic error of the numerical model without considering the influence of boundary conditions on the prediction results. The disadvantage is that the establishment of statistical models requires a large amount of historical observation data.

3 Downscaling Method

Firstly, the historical numerical model products are collected at the same time and the training data set is established. Then, a geographic weighted regression statistical downscaling method combining

DEM, NDVI, slope direction, longitude and latitude is adopted. Extract the relevant information between high-resolution and low-resolution data, take the low-resolution data as the input of the model, and obtain the high-resolution data corresponding to time as the output. Finally, it is compared with the established EOF iterative statistical downscaling method. The results were analyzed and verified by correlation coefficient (R^2), absolute error (MAE) and root mean square error (RMSE). The method and technology roadmap is as follows (Fig. 2):

$$R^2 = \left(\frac{\sum_{i=1}^{n} (P_{Si} - \bar{P}_{Si}) \sum_{i=1}^{n} (P_{Hi} - \bar{P}_{Hi})}{\sqrt{\sum_{i=1}^{n} (P_{Si} - \bar{P}_{Si})^2} \sqrt{\sum_{i=1}^{n} (P_{Hi} - \bar{P}_{Hi})^2}} \right)^2 \tag{2}$$

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |P_{Si} - P_{Hi}| \tag{3}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (P_{Si} - P_{Hi})^2} \tag{4}$$

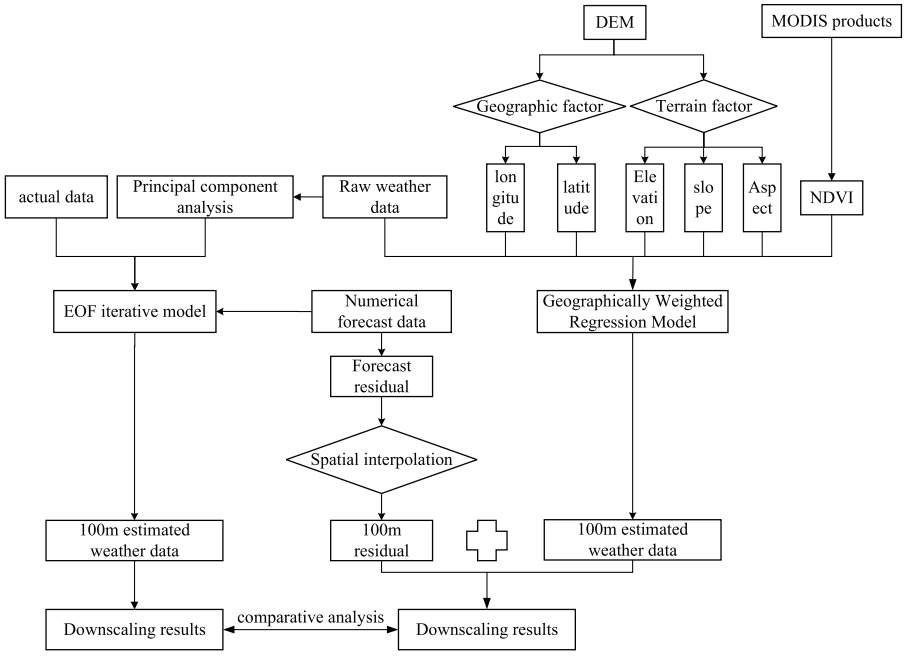


Figure 2: Technical route of statistical downscaling method

In the formula, n represents the number of weather data of the day, I represent the data number, P_{Hi} represents the i-th data value after downscaling, and P_{Si} represents the i-th data measured value.

3.1 Selection of Predictors

In the downscaling process, the selection of predictors has a large impact on the accuracy of the downscaled results. Predictors usually need to have the following characteristics [17,18]:

(1) Sensitive to large-scale climate change, maintaining a significant correlation with downscaling parameters;
(2) The acquisition process is relatively simple and can be simulated continuously and accurately.

In the study of meteorological data downscaling, due to the different leading factors affecting meteorology among regions, its selection is not unified. It is mainly divided into the following meteorological factors, ground factors and location factors (see Table 1).

Table 1: Common predictors

Meteorological factors	Terrain factor	Location factor
Wind field	Elevation	Latitude and longitude
Air pressure	Slope	Sea and land location
Cloud cover	Slope direction	
Evaporation	Fragmentation	
Humidity	Elevation	

Among them, in this research, we strive to collect predictive factors with a resolution of 100 m, taking into account that the climate is affected by geographical location and other factors, resulting in not only complex and diverse spatial distribution but also instantaneous change [19]. Even if we ensure that the underlying surface conditions are the same, the weather data will vary due to changes in geographic location.

3.2 Geographically Weighted Regression Method

The geographically weighted regression (GWR) model is a development on the linear regression model, and its greatest advantage over the general model is the embedding of geographic information into the regression model. The local characteristics of the variables are strengthened by introducing geographic factors, which increases the credibility of the downscaling results. Therefore, DEM and NDVI local Moran index are introduced to investigate the downscaling of meteorological data. Assuming that DEM, NDVI, Slope direction, longitude and latitude are consistent with meteorological data of different spatial scales, the regression model between meteorological data and annual maximum synthetic NDVI, DEM, slope direction, longitude and latitude is established by analyzing the correlation between meteorological elements. As shown in formula (5):

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \varepsilon \tag{5}$$

In the formula, y is the weather data estimated by the model, x_1 is the longitude, x_2 is the latitude, x_3 is the elevation, x_4 is the annual maximum composite NDVI, and x_5 is the slope direction. β represents the regression coefficient, and ε represents the random error of the model.

GWR is a local regression method proposed by Brunsdon et al. [20], which can be used to investigate the non-stationarity of spatial correlation. The GWR model assumes that the coefficients vary with spatial location, and it detects spatial variations in the correlations by adding spatial location

information to the ordinary linear regression process [21], which is written in its general form as:

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{p} \beta_k(u_i, v_i)x_{ik} + \varepsilon_i \tag{6}$$

In the formula, y_i represents the observation value of the i sample point; x_{ik} represents the i independent variable of the k sample point; (u_i, v_i) represents the coordinates of the i regression point; $\beta_k(u_i, v_i)$ is a function of geographic location, representing the i sample The k regression parameter of the point. ε_i represents the random error of the model at the i sample point.

The formula is simplified to:

$$y_i = \beta_{i0} + \sum_{k=1}^{p} \beta_{ik}x_{ik} + \varepsilon_i \tag{7}$$

The following formula is the expression of the parameter estimation vector using the weighted least square method:

$$\hat{\beta}_i = (x^T W_i x)^{-1}(x^T W_i y) \tag{8}$$

where $\hat{\beta}_i$ is the local coefficient estimated at the sampling point i; x and y are independent vectors and variable correlation vectors, respectively; W_i represents the weighting matrix, and this effect gives a greater weight at the point closer to the i point.

The common spatial weight functions are inverse distance function, distance threshold function method, double square function and Gaussian function [22]. The Gaussian function uses a continuous monotonic function to represent the relationship between the weights W_{ij} and d_{ij} distances, and the functional expressions of the two are:

$$W_{ij} = \exp(-(\varphi d_{ij})^k) \tag{9}$$

In the formula, φ represents the attenuation parameter and gives the parameter k of the smoothing factor. The form of the Gaussian function used in the GWR model function is as follows:

$$W_{ij} = \exp(-(d_{ij}/c)^2) \tag{10}$$

where c denotes bandwidth, the larger the bandwidth, the slower the spatial weights decay with increasing distance, too large or too small will lead to inaccurate estimation of regression parameters. The common methods to determine the c values are AIC (information criterion) method and the CV (cross-validation) method [23].

Through the GWR model established above, the functional relationship between meteorological elements and the elevation, slope direction, NDVI, etc., of the reporting area is established, and the spatial scale conversion of low-resolution meteorological data is realized. The scale conversion process is divided into the following steps:

(1) First, the grid is divided (Fig. 3), and the pre-processed original NDVI, elevation, slope and slope direction data are respectively upgraded to 100 ∗ 100 m spatial resolution, and the corresponding information is extracted from the network;
(2) Establish a correlation regression model of meteorological data and predictive factors on 100 m;
(3) Based on the original data, the regression model obtained in Step (2) is used to simulate the meteorological data in each period;

(4) Calculate the residuals of the original meteorological data and the data during the forecast period, and interpolate the residuals of the data to a spatial resolution of 100 m;
(5) Input the prediction factor of 100 m spatial resolution into the model in Step (2), and get its 100 m meteorological data;
(6) Add the 100 m meteorological data in Step (5) to the 100 m meteorological data residual in Step (4) to obtain downscaled meteorological data with a spatial resolution of 100 m.

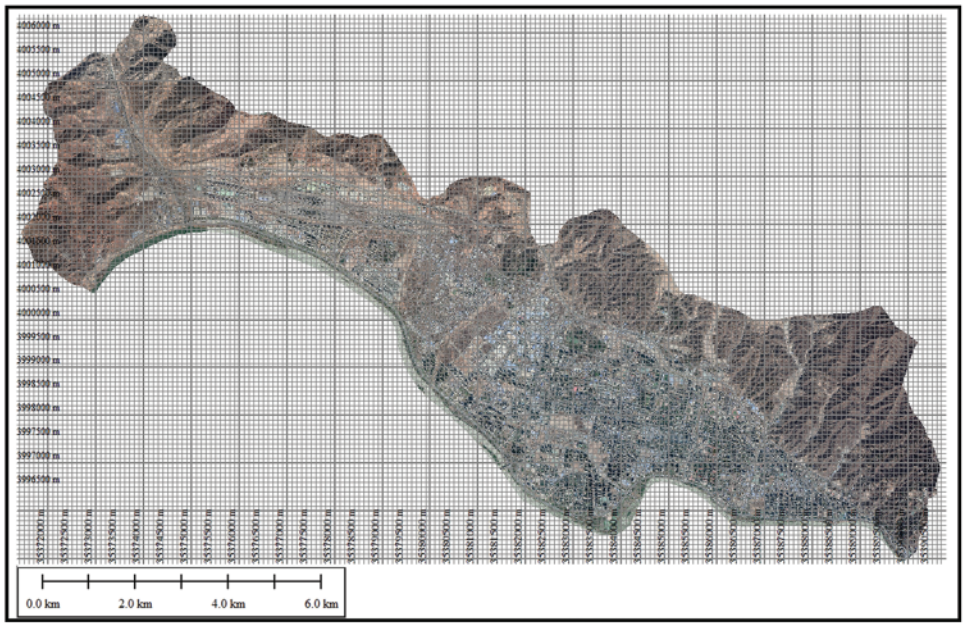


Figure 3: Meshing of anning district

3.3 EOF Iterative Method

The iterative method constructs a matrix containing the forecast field and the forecast object and performs multiple expansions and reconstructions to obtain the forecast results. The idea of the model design method is to construct a matrix composed of two elements, including the historical and real-time prediction field and the location of the prediction field of the object, and decompose them into two spatial function fields and real-time prediction function fields. The spatial function field can reflect the relationship of various weather and physical coordination between the prediction field and the predicted object field, and the time function field can reflect the international and geographical interannual changes of time. By integrating a variety of information between the prediction space-time field and the prediction object space-time field, the prediction with complex relationships can be obtained. The advantage of this method is that the amount of calculation is small and easy to implement. Compared with other methods, it directly applies the corresponding relationship between the prediction factor field and the actual data field. The method flow chart is as follows:

This article first uses principal component analysis (PCA) to screen it before the EOF iteration. PCA is a statistical analysis method to grasp the main contradictions of things, which can resolve the main influencing factors from multiple things, reveal the essence and simplify complex problems. PCA has been widely used for downscaling and compression of downscaled predictors, which can simplify a large number of predictors into a few new comprehensive index factors. The steps are shown in Fig. 4.

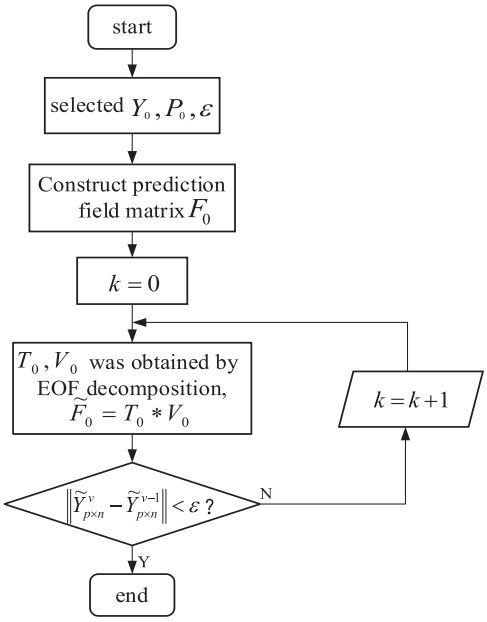


Figure 4: EOF iteration flow chart

Calculate the covariance matrix S of the matrix X.

$$S = \frac{1}{m-1}\sum_{j=1}^{m}(x_j - \bar{x})(x_j - \bar{x})^T \tag{11}$$

$$\bar{x} = \frac{1}{m}\sum_{j=1}^{m} x_j \tag{12}$$

Calculate the eigenvalue E and eigenvector $\lambda_i (i = 1, 2, \cdots, n)$ of S, and sort the eigenvalues in order $\lambda_1 > \lambda_2 > \cdots > \lambda_n$.

Call $\lambda_1 / \sum_{i=1}^{n} \lambda_i$ the contribution rate of the first principal component, and call $\sum_{j=1}^{k} \lambda_i / \sum_{i=1}^{n} \lambda_i$ the cumulative contribution rate of the first k principal components. If the cumulative contribution rate of the first k principal components exceeds 85%, it is considered that the first k principal components basically contain the original indicator information.

The main component field of meteorological forecast factors and the forecast field matrix of actual meteorological elements are constructed as follows:

$$F^0_{(m+p)\times(n+1)} = \begin{pmatrix} P^0_{m\times n} & P^0_{m\times 1} \\ Y^0_{p\times n} & Y^0_{p\times 1} \end{pmatrix} \tag{13}$$

where $P^0_{m\times n}$ and $P^0_{m\times 1}$ are respectively a block matrix composed of several (a total of m spatial points) model meteorological factor field n sample years and forecast years, $Y^0_{p\times n}$ and $Y^0_{p\times 1}$ are (a total of p spatial points) live meteorological element field n samples A block matrix composed of the data of the year and the initial value of the meteorological element field for a given forecast year.

Perform EOF decomposition on matrix $F^0_{(m+p)\times(n+1)}$ to obtain eigenvector matrix $V^0_{(m+p)\times(n+1)}$ and time coefficient matrix $T^0_{(m+p)\times(n+1)}$. The fitting field is obtained by using the eigenvectors corresponding to

the first k largest eigenvalues and their corresponding time coefficients (main components), namely:

$$\tilde{F}^0 = V^0_{(m+p)\times k} \times T^0_{k\times(n+1)} \tag{14}$$

$$\tilde{F}^0 = \begin{pmatrix} \tilde{P}^0_{m\times n} & \tilde{P}^0_{m\times 1} \\ \tilde{Y}^0_{p\times n} & \tilde{Y}^0_{p\times 1} \end{pmatrix} \tag{15}$$

where $\tilde{Y}^0_{p\times n}$ is the value of iteration 0. The percentage of the sum of the first k largest eigenvalues and the sum of all eigenvalues (the number is the rank of the normal matrix $F^0_{(m+p)\times(n+1)} \times F^T_{(m+p)\times(n+1)^0}$) is called the truncated explanatory variance of the EOF fitting field. In this paper, it is called EOF iterative truncated explanatory variance or EOF iterative truncated principal component. It cannot take $\min(m+P, n+1)$, otherwise, it cannot be iterated and can only recover the original field.

Reconstructing the matrix above $\tilde{F}^0$, we get:

$$F^1_{(m+p)\times(n+1)} = \begin{pmatrix} P^0_{m\times n} & P^0_{m\times 1} \\ Y^0_{p\times n} & \tilde{Y}^0_{p\times 1} \end{pmatrix} \tag{16}$$

Perform EOF decomposition on the matrix $F^1_{(m+p)\times(n+1)}$ to obtain the eigenvector matrix $V^1_{(m+p)\times(n+1)}$ and the time coefficient matrix $T^1_{(m+p)\times(n+1)}$, and still use the first k eigenvectors and their corresponding time coefficients to obtain the fitting field, namely:

$$\tilde{F}^1 = V^1_{(m+p)\times k} \times T^1_{k\times(n+1)} \tag{17}$$

$$\tilde{F}^1 = \begin{pmatrix} \tilde{P}^1_{m\times n} & \tilde{P}^1_{m\times 1} \\ \tilde{Y}^1_{p\times n} & \tilde{Y}^1_{p\times 1} \end{pmatrix} \tag{18}$$

where $\tilde{Y}^1_{p\times n}$ is the value of the first iteration, to construct the matrix, we get:

$$F^2_{(m+p)\times(n+1)} = \begin{pmatrix} P^0_{m\times n} & P^0_{m\times 1} \\ Y^0_{p\times n} & \tilde{Y}^1_{p\times 1} \end{pmatrix} \tag{19}$$

Repeat the above operation. When the two iterative values meet $||\tilde{Y}^v_{p\times n} - \tilde{Y}^{v-1}_{p\times n}|| < \varepsilon$ ($\varepsilon = 0.01$), $\tilde{Y}^v_{p\times n}$ is the predicted value of the prediction field.

4 Case Analysis

4.1 Overview of the Study Area

Anning District is located on the North Bank of the Yellow River in Lanzhou. It is in the northwest suburb of Lanzhou. It is located at 103°34–103°47 east longitude and 36°5–36°10 north latitude, with a total area of 82.33 square kilometers (Fig. 5). The river valley in the Lanzhou section of the Yellow River is high in the north and low in the south. The north is a rocky mountain covered with loess, and the south is a wide river valley. According to the climate division of Lanzhou City, the Southern Sichuan beach belongs to the warm and semi-dry climate area, and the northern hilly land belongs to the shallow mountain mild, and semi-dry climate area.

In this section, two different areas in Anning District (as shown in Fig. 6) are selected to obtain their meteorological data change values through actual measurement. It can be seen from Fig. 7 that there will be differences in irradiance, temperature, relative humidity, and wind speed in different adjacent areas. Therefore, the study of topographic factors such as terrain, vegetation, and the slope is of great significance for the downscaling of meteorological data.

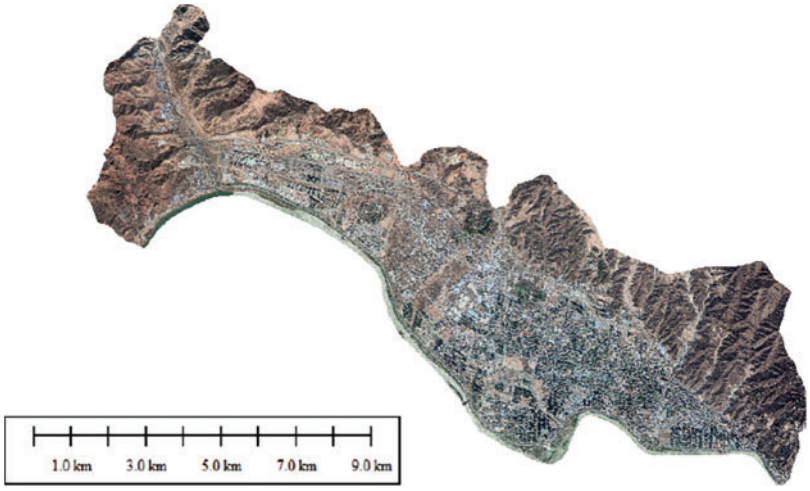


Figure 5: Map of anning district

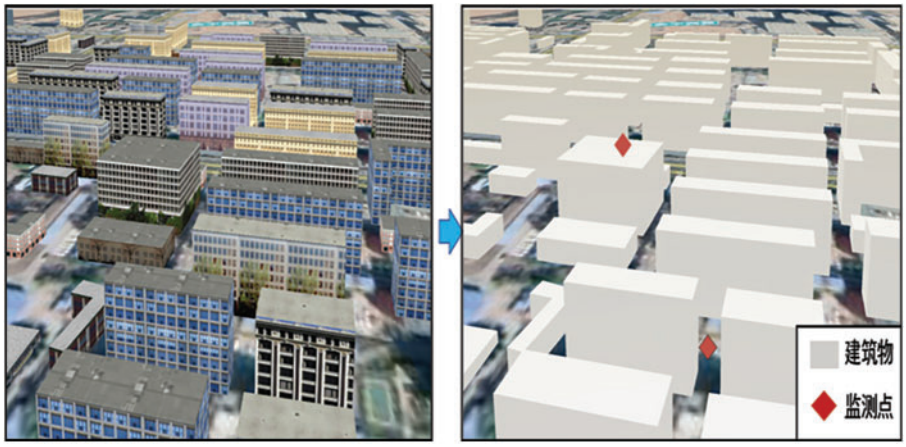
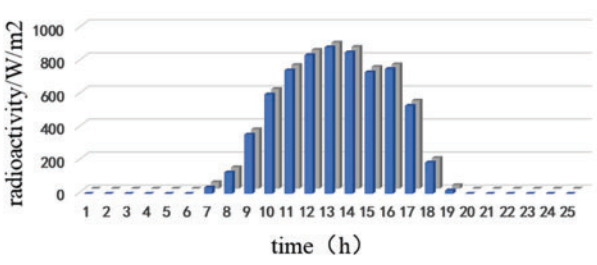
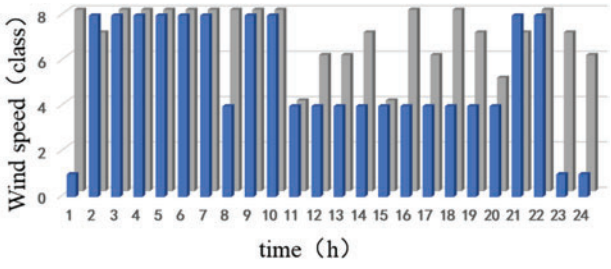


Figure 6: 3D building model drawing

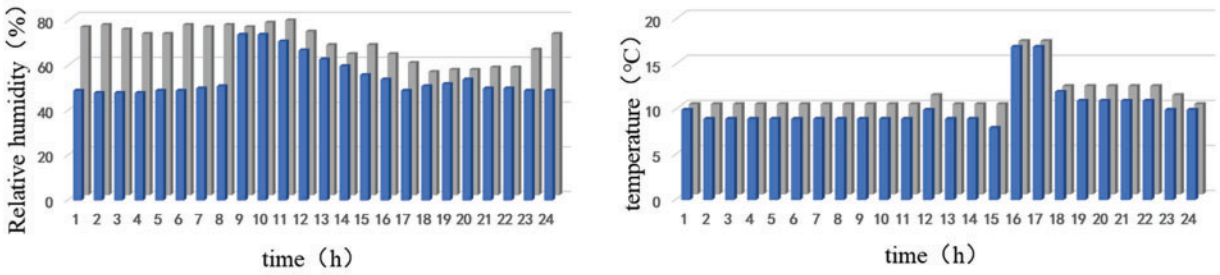


(a)



(b)

Figure 7: (Continued)



(c) (d)

Figure 7: Comparison of meteorological elements in different regions

4.2 Data Source and Preprocessing

This paper selects meteorological change products in the numerical weather forecast products provided by the National Meteorological Administration as the original data. The selected product range is the meteorological area of Anning District, Lanzhou City, Gansu Province, and the actual measurement range is the area of Lanzhou Jiaotong University (Fig. 8). Among them, the data of a total of 2976 h in July 2019 and a total of 96 h on August 01 are selected as the experimental training data and test data, respectively.

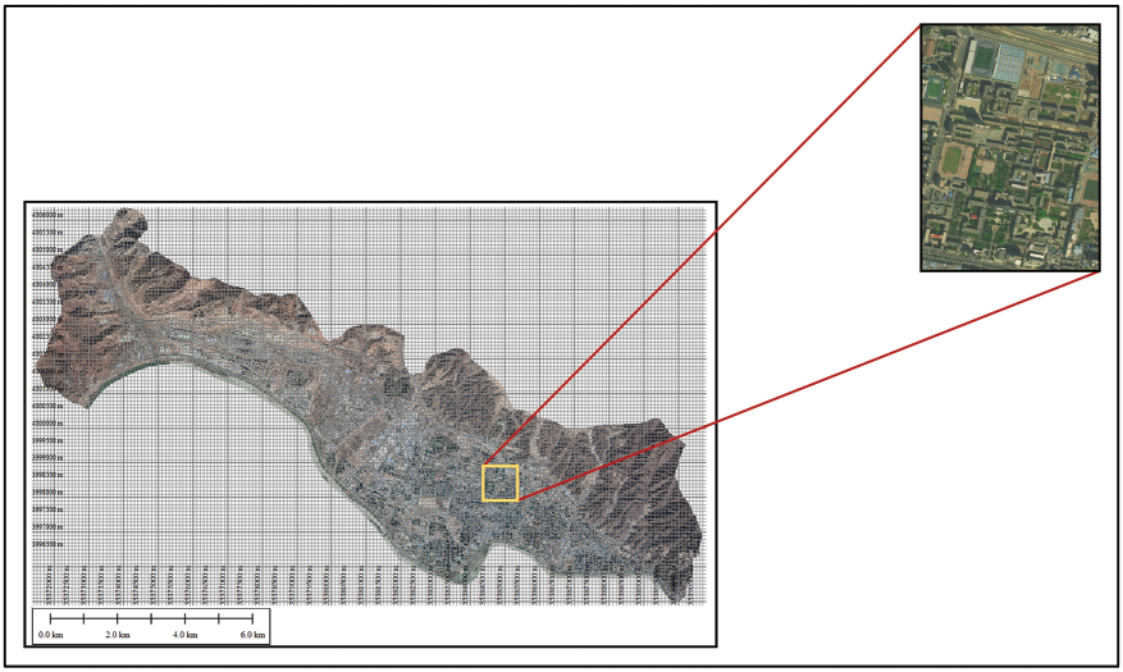


Figure 8: Downscale measured range

Vegetation is one of the most important factors influencing meteorological factors, and among various vegetation indices, NDVI can more accurately describe vegetation quantity, quality, growth and cover [24], so NDVI is used as one of the core variables for downscaling. The NDVI data selected for this study are MOD13A3 products for the period January to December 2019, with a spatial resolution of 100 m. This product uses 16-day 100 m MODIS VI output to create a one-month

composite. By calculating in pixels, whenever 16-day composite data is collected, a weighting factor is applied to the input data according to the degree of time overlap. On this basis, the 12-month average of the year was calculated, and the annual NDVI value was obtained. Rasters smaller than 0 in NDVI mostly represent abnormal pixels, which need to be removed from the data. We take 2019 as an example to process the NDVI data as shown in Fig. 9.

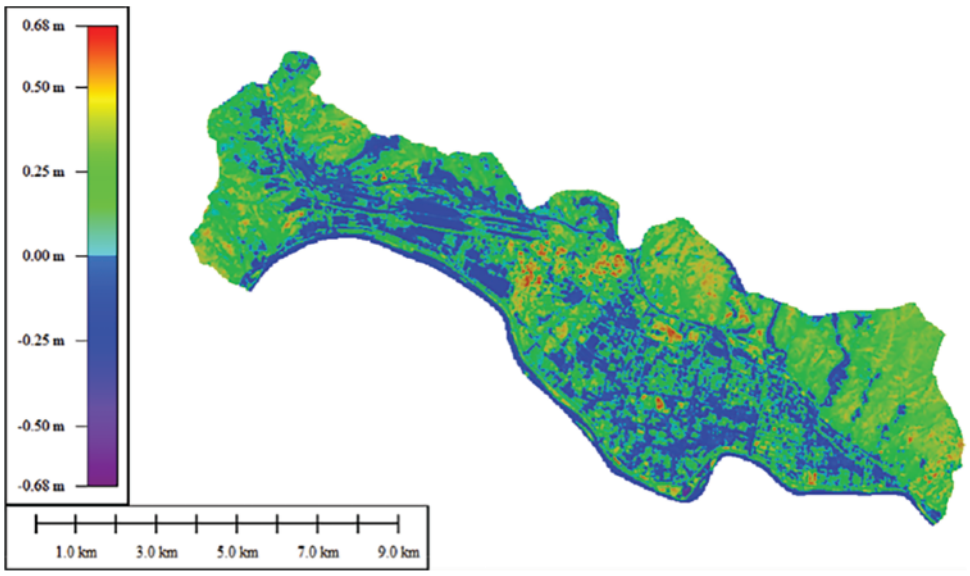


Figure 9: NDVI coverage index

The DEM data comes from the geospatial data cloud, the website is http://www.gscloud.cn. Select the DEM data product of the required area, as shown in Fig. 10. This product includes two spatial resolutions of 30 and 90 m. Each 90 m data point is obtained by the arithmetic average of 9(30 × 30 m) data points, and each grid is a file. This study uses SRTM data with a resolution of 90 m. Use Arcgis software to carry out a series of preprocessing of the data to obtain elevation data with a resolution of 100 m in the survey area, and extract other topographic indices based on this information such as slope and slope direction [25].

Due to the high altitude of natural resources, sufficient sunlight, obvious continental climate, human factors and river erosion in the Anning area. In order to effectively eliminate the interference of the above factors, the local Moran index is introduced to modify it. This index can accurately identify spatial agglomeration, detected spatial variability, etc. The local Moran index of NDVI directly reflects the spatial variability of vegetation, and the calculation formula is as follows:

$$I_j = \frac{n^2}{\sum_j \sum_k w_{jk}} \times \frac{(x_j - \bar{x}) \sum_k w_{jk}(x_k - \bar{x})}{\sum_k (x_k - \bar{x})^2} \tag{20}$$

In the formula, I_j is the local Moran index, n is the sample number of the spatial reference object location unit, x_j is the attribute value of the spatial reference object location unit j, w_{jk} represents the spatial weight matrix, which represents the distance between the reference object location unit k and j influence level.

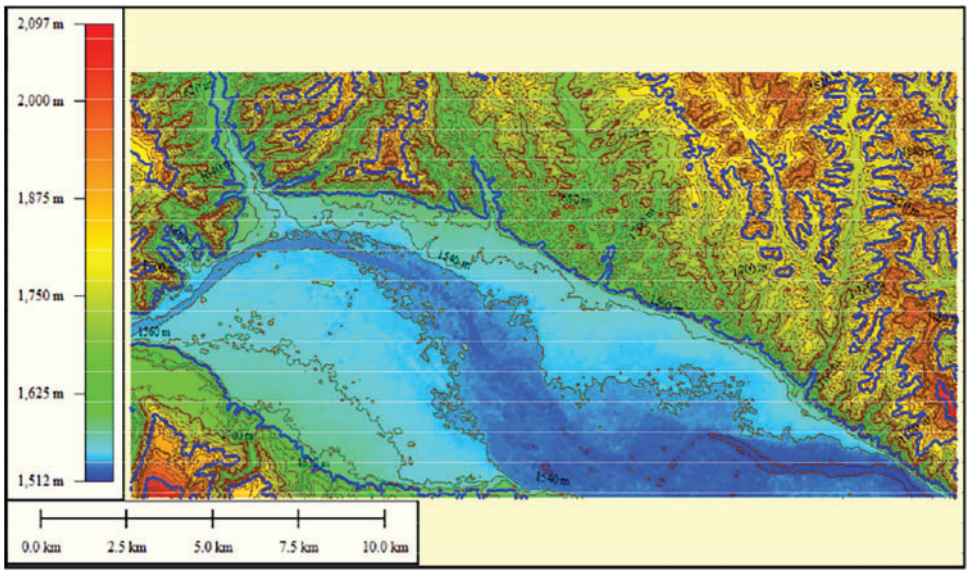


Figure 10: Digital elevation map

4.3 Result Analysis

In order to verify the accuracy of the downscaling results and evaluate the downscaling performance, this paper analyzes the irradiance, wind speed, relative humidity and temperature data and the original data through RMSE, MAE and R^2. The results are shown in Table 2 below.

Table 2: Summary of simulation results of correlation coefficients before and after downscaling of meteorological elements

Meteorological elements	Data	R^2	MAE	RMSE
radioactivity	Original	0.965	2.084	3.081
	Downscaling	0.972	0.693	1.093
Wind speed	Original	0.920	0.268	0.317
	Downscaling	0.985	0.070	0.098
Relative humidity	Original	0.951	1.954	2.299
	Downscaling	0.989	0.630	0.922
Temperature	Original	0.757	1.665	1.907
	Downscaling	0.911	0.819	1.106

First, analyze the differences in the downscaling results of different meteorological data sets based on the regional measured data. Fig. 11 is a scatter plot of the measured weather data and the data before and after downscaling. Among them, Figs. 11a and 11b are the scatter plots of the measured irradiance data and the data before and after downscaling, respectively. The R^2 value before and after downscaling is almost unchanged, from 0.965 to 0.972, an overall increase of 0.07. MAE and RMSE were reduced by 1.391 and 1.988, respectively. Figs. 11c and 11d are the scatter plots of the

measured wind speed data and the data before and after downscaling, respectively. The R^2 value before and after downscaling increased from 0.920 to 0.985, an increase of 0.065. The values of MAE and RMSE decreased from 0.268 and 0.317 to 0.070 and 0.098 respectively before and after downscaling. Figs. 11g and 11h are the scatter plots of the measured temperature data and the data before and after downscaling, respectively. The R^2 value increased significantly before and after downscaling, from 0.757 to 0.911, an overall increase of 0.154. The values of MAE and RMSE decreased slightly, from 1.665 and 1.907 to 0.819 and 1.106, respectively. Finally, the R^2 value of the relative humidity data before and after downscaling increased from 0.951 to 0.9989, and the overall R^2 value changed by 0.03. The MAE and RMSE values decreased significantly before and after downscaling, from 1.954 and 2.299 to 0.630 and 0.922, respectively. The R^2 values after downscaling of the four sets of data all show an upward trend, while the MAE and RMSE values tend to decrease. Although Figs. 11a and 11b show that the data changes before and after downscaling are not obvious. In addition, in the verification process, the slope and intercept of the regression line were obtained by fitting a scatter plot, and it was found that the deviation of the slope and intercept of the regression line was significantly reduced after downscaling. This is mainly due to the combination of downscaling methods that can better express the relationship between meteorological data and regional surface variables.

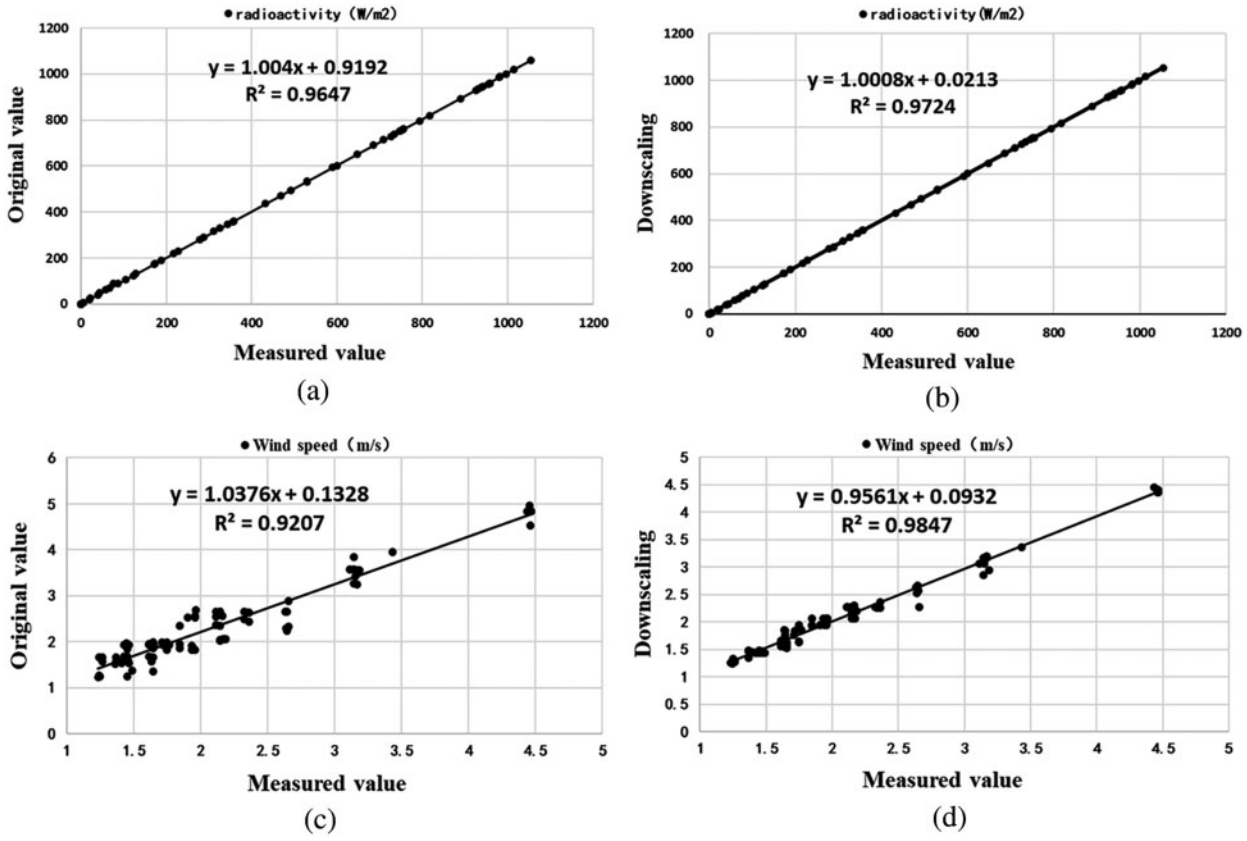


Figure 11: (Continued)

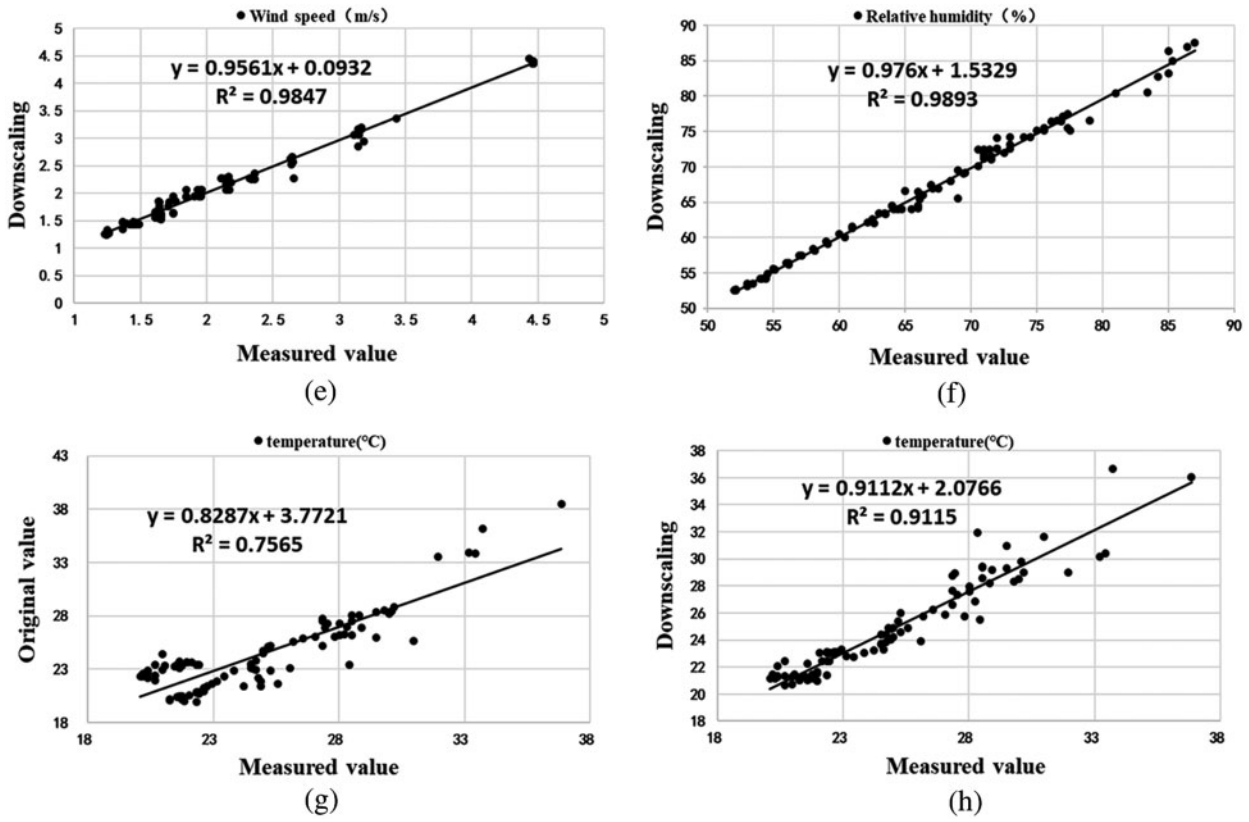


Figure 11: Comparison of scatter before and after downscaling of meteorological data

In order to further illustrate the effectiveness of the method proposed in this paper, this section compares the reduced scale results with the EOF method and evaluates the effect according to four different meteorological elements: irradiance, wind speed, relative humidity and temperature. Fig. 12 shows the downscaling curves of different meteorological elements under the statistical downscaling methods of original values, measured values, EOF iteration and geographically weighted regression.

In general, the geographically weighted regression method can more accurately reflect the time change than the EOF iterative method. Due to the small spatial scale, the change of irradiance before and after downscaling is not very obvious, and the change of wind speed, relative humidity and temperature are more prominent affected by geographical environmental factors. The EOF downscaling results are similar to the measured data over time, which is mainly due to the fact that the EOF downscaling results are very dependent on the measured data and can not better characterize the change details of meteorological data from time and space scales.

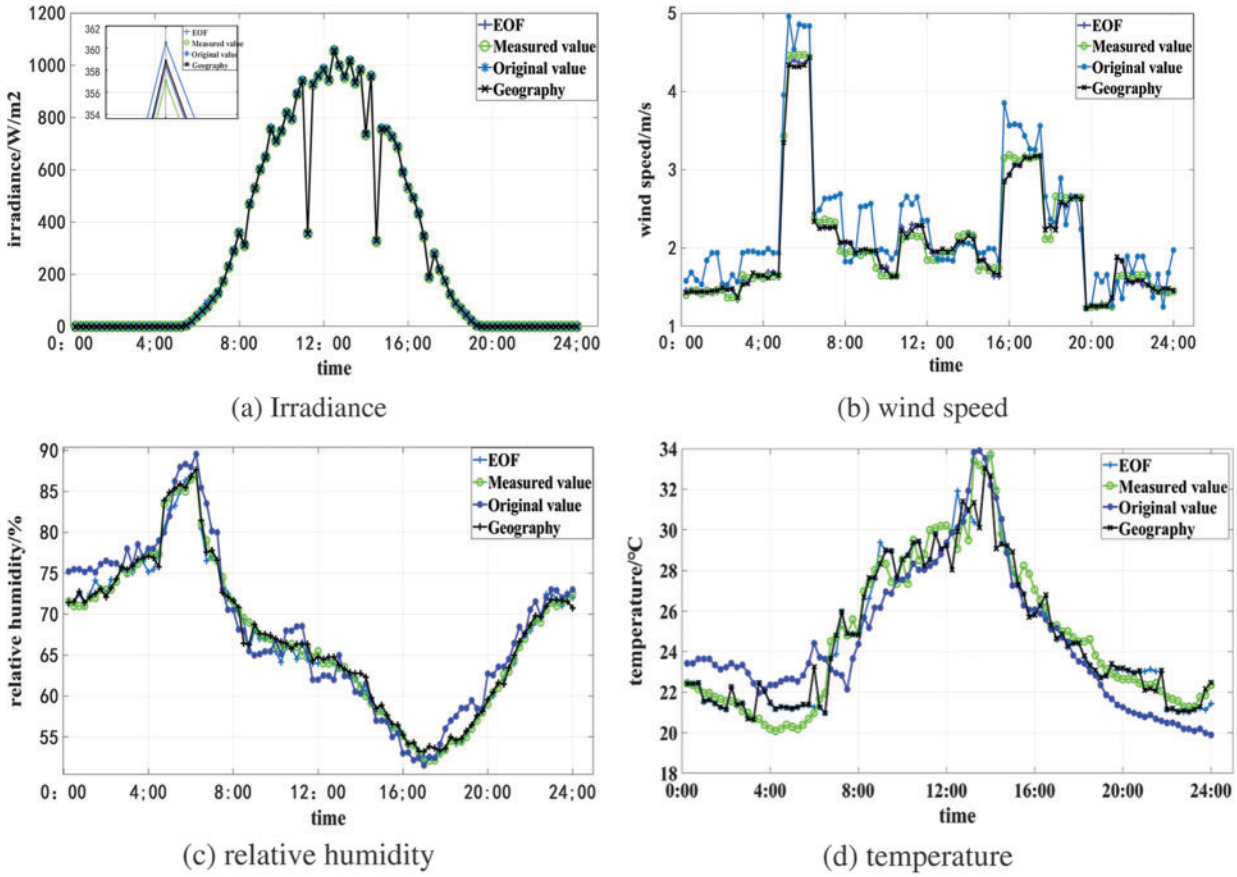


Figure 12: Simulation comparison of meteorological elements by different models

5 Conclusion

Based on the correlation between meteorological data and DEM, NDVI, Slope direction, and latitude and longitude, this paper constructs a geographically weighted spatial downscaling method that takes into account the influence factors of DEM, NDVI, Slope direction, and latitude and longitude. Use the data provided by the National Meteorological Administration to downscale it to increase its spatial resolution to 100 m. Through comparative analysis of the measured data and EOF iteration method in the study area, the following conclusions are drawn:

(1) Compared with the error between the original meteorological data and the measured meteorological data, the error of the meteorological data after downscaling by this method is obviously smaller. For example, the correlation coefficient between the original temperature data and the measured temperature data is 0.757, while the correlation coefficient between the downscaling temperature data and the measured temperature data is 0.911. Obviously, this method can not only obtain high-precision meteorological data, but also provide a more accurate meteorological basis for the prediction of photovoltaic power.

(2) Through the comparative analysis of geographical weighted regression downscaling method and EOF iterative downscaling method for the meteorological data in the study area, it is found that the error of EOF iteration is obviously large and has certain limitations on

reflecting the spatial distribution characteristics of meteorological data, while the statistical downscaling method combined with DEM, NDVI, slope direction, longitude and latitude, and other influencing factors have a better effect, The results can reflect the spatial distribution characteristics of Meteorology in the study area in detail.

Funding Statement: This work is supported by the Science and Technology project of State Grid Gansu Electric Power Company (No. 52272219002).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. State Energy Administration (2020). Grid connected operation of photovoltaic power generation in 2020. http://www.nea.gov.cn.
2. Li, Y., Wang, Z., Chu, H. J., Qin, Z., Bao, D. M. (2014). Photovoltaic power generation prediction system based on extjs. *Power Generation Equipment, 28(1),* 70–74.
3. Sheng, X. (2020). *Spatial downscaling and spatiotemporal variation analysis of satellite precipitation data over the Qinghai Tibet Plateau.* Nanjing University of Information Technology.
4. Zhang, X. (2013). *Downscaling and spatiotemporal analysis of satellite precipitation data in the Middle Tianshan Mountains.* Northwest Normal University, China.
5. Fan, D. (2017). *Downscaling of TRMM satellite precipitation data in the middle reaches of the Yellow River.* Henan University of Technology, China.
6. Fan, L. J., Fu, C. B., Chen, D. L. (2005). Research progress of statistical downscale method on prediction of future regional climate change scenarios. *Geoscience Progress, (3),* 320–329.
7. Sheng, J. H., Rao, P. (2021). Study on downscaling method of soil moisture data based on Fengyun meteorological satellite. *Journal of Infrared and Millimeter Wave, 40(1),* 74–88.
8. Ji, T., Liu, R., Yang, H., He, T. R., Wu, J. F. (2015). Spatial downscaling of precipitation based on multi-source remote sensing data: A case study of Sichuan and Chongqing. *Acta Geoinformatics, 17(1),* 108–117. DOI 10.3724/SP.J.1047.2015.00108.
9. Merlin, O., Walker, J., Chehbouni, A., Kerr, Y. (2008). Towards deterministic downscaling of SMOS soil moisture using MODIS derived soil evaporative efficiency. *Remote Sensing of Environment, 112(10),* 3935–3946. DOI 10.1016/j.rse.2008.06.012.
10. Merlin, O., Rudiger, C., Albitar, A., Richaume, P., Walker, J. et al. (2012). Disaggregation of SMOS soil moisture in Southeastern Australia. *IEEE Transactions on Geoscience and Remote Sensing, 50(5),* 1556–1571. DOI 10.1109/TGRS.36.
11. Chen, B., Xu, B. (2014). A unified spatial-spectral-temporal fusion model using Landsat and MODIS imagery. *2014 Third International Workshop on Earth Observation and Remote Sensing Applications (EORSA)*, pp. 256–260. Changsha. DOI 10.1109/EORSA.2014.6927890.
12. Feng, J., Masek, J., Schwaller, M., Hall, F. (2006). On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance. *IEEE Transactions on Geoscience and Remote Sensing, 44(8),* 2207–2218. DOI 10.1109/TGRS.2006.872081.
13. Houborg, R., McCabe, M. F., Gao, F. (2015). Downscaling of coarse resolution LAI products to achieve both high spatial and temporal resolution for regions of interest. *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 3317–3320, Milan, Italy. DOI 10.1109/IGARSS.2015.7326528.
14. Weng, Q., Fu, P., Gao, F. (2014). Generating daily land surface temperature at landsat resolution by fusing landsat and MODIS data. *Remote Sensing of Environment, 145(8),* 55–67. DOI 10.1016/j.rse.2014.02.003.

15. Nerini, D., Zulkafli, Z., Wang, L., Onof, C., Buytaert, W. et al. (2015). A comparative analysis of TRMM-rain gauge data merging techniques at the daily time scale for distributed rainfall-runoff modeling applications. *Journal of Hydrometeorology, 16(5),* 2153–2168. DOI 10.1175/JHM-D-14-0197.1.
16. Wu, H. C. (2020). *Ground and TRMM precipitation data fusion based on deep learning.* University of Electronic Science and Technology, China.
17. Liu, Y., Fan, K., Zhang, Y. (2013). Statistical downscaling prediction of summer precipitation at stations in China based on CFS model. *Atmospheric Science, (6),* 1287–1296. DOI 10.3878/j.issn.1006-9895.2012.
18. Gao, H. X. (2011). *Study on Statistical Downscaling Method of Regional Climate Change in China*. Nanjing University, China.
19. Liu, J., Xia, J., Zou, L., Wang, Q., Yu, J. Y. (2018). Applicability analysis of multi satellite remote sensing precipitation data in Tarim river basin. *South to North Water Diversion and Water Conservancy Science and Technology, 16(5),* 1–8. DOI 10.13476/j.cnki.nsbdqk.2018.0117.
20. Brunsdon, C., Fotheringham, A. S., Charlton, M. E. (1996). Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis, 28(4),* 281–298. DOI 10.1111/j.1538-4632.1996.tb00936.x.
21. Stewart, F. A., Charlton, M., Brunsdon, C. (1996). The geography of parameter space: An investigation of spatial nonstationarity. *International Geographical Information Systems, 10(5),* 605–627. DOI 10.1080/02693799608902100.
22. Ding, Y. P., Zhang, J. H., Liu, Y. H., Lu, C. L., Wang, S. Q., Qin, J. T., Ding, S. Y. (2021). Analysis of spatial distribution characteristics and influencing factors of soil organic carbon in Yihe River Basin based on GWR model. *Journal of Ecology, 41(12),* 4876–4885. DOI 10.5846/stxb202001140109.
23. Ding, S. (2017). *Temporal and spatial variation of environmental heterogeneity in typical areas of the middle and lower reaches of the Yellow River.* Henan University, China.
24. Liu, X. F., Ren, Z. Y. (2012). Vegetation cover change and its relationship with climate factors in Northwest China. *Chinese Agricultural Sciences, (10),* 1954–1963. DOI 10.3864/j.issn.0578-1752.2012.10.008.
25. Feng, Y. (2017). *Research on terrain humidity index based on fine DEM*. Northwestern University, China.