Computer Mathematical Statistics Analysis of downscaling data of different vegetation index inversion TRMM

Shuangbao Qu¹,*, Miaoxing Zhao¹ and Shuo Deng²

¹Hope College of Southwest Jiaotong University, Chengdu, China
²Surveying and mapping technology service center of Sichuan surveying and mapping Geographic Information Bureau, Chengdu, China

*Corresponding author: shuangbaoqu@swjtuhc.cn

Abstract. This paper uses enhanced vegetation index (EVI) data, normalized vegetation index (NDVI) data, DEM, aspect data, and TRMM3B43 (V7) data, based on a geographically weighted regression model (GWR), and uses a statistical downscaling method to achieve Central China Downscaling of regional TRMM data from 2010 to 2019. The research results show: (1) TRMM data has good applicability in Central China, and the R² of TRMM data and weather station measured data is above 0.8. (2) Improve the ground resolution from 0.25°×0.25° (approximately 27.5km×27.5km) to 1km×1km while ensuring the same accuracy as the original data. (3) Overall, the accuracy of EVI downscaled precipitation data in Central China is better than that of NDVI downscaled precipitation data.

Keywords: EVI; NDVI; TRMM; GWR; downscaling; Central China.

1. Introduction
Precipitation is a key part of the global water cycle and plays an important role in regulating the global climate. It promotes global material circulation and energy exchange, and it plays a very important role in atmospheric processes at different time and space scales [1]. Precipitation data is the basic data for hydrological model simulation, water resource evaluation and flood disaster management. A large number of studies have shown that traditional point measurements based on meteorological stations cannot effectively reflect the spatial changes of precipitation [2]. With the rapid development of remote sensing geographic information, people have developed a series of precipitation data products based on satellite observations, and TRMM data has been widely used.

In recent years, the adaptability of TRMM data has attracted much attention. Ning [3] analyzed the adaptability of TRMM data in Tianshan area and found that the correlation coefficient between it and the measured data is above 0.7. The research results of Wang Kai et al. showed that: TRMM data and measured site data have a good linear regression relationship [4].

However, the terrestrial resolution of TRMM data is low, which cannot meet the needs of large-scale analysis, and it is necessary to perform spatial downscaling on it [5]. Most scholars at home and abroad have realized the downscaling of TRMM data based on factors such as NDVI and DEM [6, 7]. However, Wong [8] performed linear regression on EVI and precipitation data on various temporal and spatial scales, and the study showed that EVI and precipitation are more closely related than NDVI.
Recently, the method of downscaling TRMM data based on the GWR model has been very popularly.

Aifeng [9] estimated the spatial distribution of TRMM in the Qaidam Basin through the GWR model; the research results of Ji Shibao [10] show that the correction accuracy of the geographically weighted regression method is better than that of the multiple linear regression method. Therefore, based on the GWR model, this paper uses EVI or NDVI data to downscale the TRMM data in Central China to explore the impact of different vegetation indices on the downscaling results.

2. A Research area profile and data

All manuscripts must be in English, also the table and figure texts, otherwise we cannot publish your paper. Please keep a second copy of your manuscript in your office.

When receiving the paper, we assume that the corresponding authors grant us the copyright to use the paper for the book or journal in question. Should authors use tables or figures from other Publications, they must ask the corresponding publishers to grant them the right to publish this material in their paper.

2.1. Overview of the research area

Central China consists of three provinces: Hunan, Hubei and Henan. The latitude and longitude range is 24°38′-36°24′N, 108°21′-116°39′E, and the area is about 5.6×10^5 km^2[11]. Central China is located in a transportation fortress, with developed light industry and a high level of economic development. The terrain is mainly composed of plains, hills, basins, and lakes. The annual precipitation is 800-2000mm. summer is dominated by short-term precipitation and rain. Short-term precipitation and rain have a large impact on the surface, which can easily trigger floods and cause soil erosion. The average annual temperature is 14-21°C, the humidity is relatively high, the temperature is suitable, and the regional climate is conducive to the growth of vegetation.

2.2. Data source and pre-processing

The NDVI and EVI data used in this paper are MODIS13A3 products, the remote sensing precipitation data is TRMM3B43 (V7), the DEM data is the SRTM data set, and the weather station data uses the

Figure 1. Distribution of study areas and meteorological stations.

2.2. Data source and pre-processing

The NDVI and EVI data used in this paper are MODIS13A3 products, the remote sensing precipitation data is TRMM3B43 (V7), the DEM data is the SRTM data set, and the weather station data uses the
data of 53 stations in Central China from January 2010 to December 2019. MODIS data undergoes a series of preprocessing such as band extraction, image mosaic, projection coordinate transformation, and data format conversion. TRMM data is preprocessed for batch projection conversion, rotation, format conversion, and unit conversion with the help of ENVI software.

3. Research methods

3.1. GWR model principle
Scholars such as Brunsdon [12] proposed a GWR model in which the regression parameters change with the change of geographic location. According to the distance of the distance, the distance weight ratio is established to obtain the best estimate. The data obtained by combining the geographic location is often more fit the local change characteristics. The basic formula (1) of the GWR model is as follows

\[
y_i = \beta_0(u_i, v_i) + \sum_{t=1}^{n} \beta_t(u_i, v_i) x_{it} + \varepsilon(u_i, v_i) \quad i = 1, 2, \ldots, m
\]

The \( y_i \) in the formula is the precipitation data of the \( i \) point, which is used as the dependent variable. There are \( m \) points in total. \( \beta_0(u_i, v_i) \) is the constant value of the \( m \) point in the geographical regression weighting, \( (u_i, v_i) \) represents the latitude and longitude coordinates of the \( i \) point, \( \sum_{t=1}^{n} \beta_t(u_i, v_i) x_{it} \) is the coefficient of the linear regression independent variable \( x_{it} \) of the \( t \) impact factor with a total of \( n \) on the \( i \) point, and \( \varepsilon(u_i, v_i) \) is the residual calculated by the model at the \( i \) point.

3.2. TRMM downscaling steps
In this paper, three evaluation indicators of correlation coefficient (R²), relative error (BIAS) and root mean square error (RMSE) are used to evaluate the accuracy of TRMM downscaling data.

4. Results and analysis

4.1. TRMM data applicability analysis
Before using the TRMM precipitation data, first conduct a feasibility analysis on the TRMM precipitation data and the meteorological station data. As shown in Table 1, there is a clear correlation between the TRMM data and the meteorological station data in Central China from 2010 to 2019, and most of the data are related the coefficient is above 0.8, and the minimum correlation coefficient is 0.544. It can be seen that the use of TRMM data to achieve the research direction of precipitation downscaling in Central China is reliable.

| Year | \( R \) | Year | \( R \) |
|------|--------|------|--------|
| 2010 | 0.843  | 2015 | 0.804  |
| 2011 | 0.544  | 2016 | 0.857  |
| 2012 | 0.827  | 2017 | 0.839  |
| 2013 | 0.861  | 2018 | 0.821  |
| 2014 | 0.842  | 2019 | 0.805  |

Average the 10-year TRMM data to obtain the annual average data of TRMM. Similarly, obtain the annual average data of EVI downscaling and the annual average data of NDVI downscaling, and use IDW interpolation to obtain the annual average data of the site. It can be seen from Figure 2 that the EVI downscaling TRMM data, NDVI downscaling data and station average data have the same spatial distribution law. The precipitation decreases from south to north. The downscaling TRMM data is compared with the pre-downscaling TRMM data and after interpolation. The site data has more detailed information.
Figure 2. TRMM annual average data, EVI downscaling TRMM annual average data, NDVI downscaling TRMM annual average data, and weather station annual average data.

The annual TRMM data, annual EVI downscaling data, and annual NDVI downscaling data of corresponding locations are extracted through meteorological stations, and the correlation analysis is carried out with the meteorological station data respectively. It can be seen from Figure 3 that the \( R^2 \) of the EVI downscaled TRMM data and the site measured data is 0.927, which is an increase of 0.116 compared to the original data. The \( R^2 \) of the NDVI downscaled TRMM data and the site measured data is 0.918, which is slightly higher than the TRMM data. The accuracy of the downscaled TRMM data is improved, and the EVI downscaled TRMM data is closer to the measured data than the NDVI downscaled TRMM.

Figure 3. Linear correlation diagram of annual precipitation and measured rainfall of TRMM, NDVI downscaling TRMM, EVI downscaling TRMM.
Calculate three kinds of data correlation coefficient ($R^2$), relative deviation (BIAS), and root mean square error (RMSE) based on site data, the accuracy of the three types of data is evaluated. It can be seen from Figure 4 that the TRMM data has the lowest $R^2$ in the middle of each year, and the root mean square error reaches the highest peak of the year, indicating that the quality of the TRMM data around July is poor. The reason is that there is more extreme daily precipitation in summer in Central China [17], and short-term heavy precipitation has a great impact on the quality of TRMM data. The accuracy of the data before and after downscaling shows a consistent trend, and all the indicators of the data after downscaling have been improved to a certain extent. Overall, the $R^2$, BIAS, and RMSE of the EVI downscaling TRMM data are better than the NDVI downscaling TRMM data, and the improvement is more obvious in the months with heavy rainfall.

![Figure 4. TRMM data, EVI downscaling TRMM and NDVI downscaling TRMM correlation evaluation index changes.](image)

5. Conclusions

Based on EVI, NDVI, DEM and aspect data, this paper downscaling the TRMM data in the study area through the GWR model, thereby improving the ground resolution of the precipitation data in the study area, and using the station data in the study area to verify the accuracy of the downscaling data. Concluded as follow:

TRMM data is reliable to study the direction of precipitation downscaling in Central China. The $R^2$ of annual TRMM data and station data is generally high. The maximum $R^2$ is 0.94 in 2012 and the minimum is 0.544 in 2011. Other years all are above 0.8.

The ground resolution of the TRMM data is reduced to 1km×1km by downscaling, and the precipitation can be more detailed reflected in the area with relatively few station data distribution, which has more detailed information compared with the original TRMM data. And based on the analysis of three correlation evaluation indicators, it is concluded that the accuracy of downscaling data has a certain improvement compared with the original TRMM data.

Through the internal comparison of the three correlation evaluation indicators of EVI downscaled precipitation data and NDVI downscaled precipitation data, overall, the EVI downscaled precipitation data in the study area has a higher accuracy than the NDVI downscaled precipitation data, and is closer to the true value. The applicability is better, and the downscaling effect is better in months with more rainfall.
Acknowledgments
This work was financially supported by School level quality engineering project of Hope College of Southwest Jiao tong University (2019030); School level quality engineering project of Hope College of Southwest Jiao tong University (2018049) fund.

References
[1] Wang Y Y,Zhao Y J,Xie Q Y,et al. Downscaling method of TRMM satellite precipitation data in Beibu Gulf Economic Zone in Southwest China.[J].Journal of Guangxi Normal University(Natural Science Edition),2019,37(4):163-169. (in Chinese).
[2] D. D. Alexakis,I. K. Tsanis. Comparison of multiple linear regression and artificial neural network models for downsampling TRMM precipitation products using MODIS data[J]. Environmental Earth Sciences,2016,75(14):
[3] Ning, L. K.,Bao, A. M.,Pan, X. L., et al. Adaptability evaluation of TRMM over the Tianshan Mountains in central Asia[J]. Mausam: Journal of the Meteorological Department of India,2016,67(3):625-632.
[4] Wang K,Cen L,Feng J H, et al. Calculation of rainfall erosivity in China with TRMM Data [J]. Arid Land Geography,2015,38(05):948-959.
[5] Fan X W,Liu H L. Downscaling Method of TRMM Satellite Precipitation Data over the Tianshan Mountains [J]. Journal of Natural Resources,2018,33(3):478-488.
[6] Qiang Zhang,Peijun Shi,Vijay P. Singh, et al. Spatial downscaling of TRMM-based precipitation data using vegetative response in Xinjiang, China[J]. International Journal of Climatology,2017,37(10):3895-3909.
[7] Zhang Yueyuan,Li Yungang,Luo Xian, et al. Fine-Resolution Precipitation Mapping in a Mountainous Watershed: Geostatistical Downscaling of TRMM Products Based on Environmental Variables[J].,2018,10(1).
[8] WongWFJ. Spatial and temporal analysis of MODIS EVI land TRMM3B43 rainfall retrievals in Australia[C]. International Conference on Geoinformatics.IEEE,2011:1-7.
[9] Lv Aifeng,Zhou Lei,Koch Magaly, et al. A Rainfall Model Based on a Geographically Weighted Regression Algorithm for Rainfall Estimations over the Arid Qaidam Basin in China[J]. Remote Sensing,2016,8(4).
[10] Ji S B,Du J K,Qiu K Q, et al. Spatiotemporal distribution of precipitation in Guizhou and Guangxi karst mountainous regions based on satellite earth fusion data [J]. South-to-North Water Transfers and Water Science & Technology,2019,17(6):28-36.
[11] Brunsdon C, Fotheringham A S, Charlton M E. Geographically weighted regression: A method for exploring spatial nonstationarity [J]. Journal of the Royal Statistical Society,2017,47(3):431 - 443.