NDVI-based vegetation in response to climate factors based on neighborhood association effect: A case study in Inner Mongolia, China

Yuwei Wang1,*, Xiaoliang Meng2, Kaicheng Wu1 and Wang Gao1
1 School of Artificial Intelligence, Jianghan University, Wuhan 430056, China
2 School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China
*Corresponding author’s email: weberwang@jhun.edu.cn

Abstract. With the combination of 2014-2018 MODIS NDVI products and climate data (precipitation and temperature) in Inner Mongolia, China, this study aims to explore and verify the effectiveness of incorporating neighborhood association effect in vegetation index modeling. A neighborhood statistical method based on Moore neighborhood was applied to update the original spatial datasets. Geographically weighted regression (GWR) was constructed to compare the model accuracy between original data and updated data. The GWR models were tested under different neighborhood sizes (3×3, 5×5, 7×7, 9×9, and 11×11 Moore neighborhood sizes). Our work compared the results of different GWR models and the original GWR model that did not consider neighborhood association effect. The results indicated that considering neighborhood association effect could improve GWR model accuracy substantially. In addition, the neighborhood sizes proved to be important factors for measuring neighborhood association effect. We conclude that neighborhood association effect should be integrated to understand vegetation changing trajectory based on climate factors.

1. Introduction
Vegetation serves as the key structural part of water conservation, soil improvement and air purification in surface ecosystem, which plays an important role in optimizing natural performance and improving environmental quality [1,2]. As a natural link between soil and atmosphere, vegetation has significant impacts on regulating climate. At the same time, climate factors such as temperature and precipitation also determine the growth and phenology of vegetation, and affect the growth and distribution of plants [3,4]. Hence, the mechanisms between vegetation and climate factors have always been a hot issue in global climate change research, and have become a crucial theoretical basis for regional vegetation protection and sustainable development of ecosystem [5,6]. Normalized difference vegetation index (NDVI), an effective indicator of vegetation coverage and growth, has been widely used to study vegetation classification and coverage at regional scale in previous researches [7,8].

In recent years, satellite remote sensing images have become one of the important data sources for monitoring large-scale vegetation dynamic change, and have been applied in the field of ecological protection and global climate change [9,10]. Simultaneously, based on geographic information system (GIS) and spatial analysis techniques, the relationships between vegetation NDVI and climate factors can be effectively quantified. According to the Tobler’s first law of geography, spatial association is...
the ubiquitous phenomenon in spatial analysis [11]. In the study of spatial association analysis, especially in the local scale analysis, scholars have paid more and more attention to the neighborhood association effect [12,13].

A number of researches have indicated that neighborhood association effect is an essential part during geographic process modeling, such as ecosystem evolution, land use and land cover change, and should be taken into consideration to better understand local spatial configurations [14,15]. At local scale, the neighborhood association effect means the underlying spatial interactions between variables of different locations. Thus, the value of spatial variable in a certain spatial location (a raster pixel), may be affected by neighborhood areas [16]. In detail, the value of spatial sample is not only related to the value of the current location, but also affected by the neighborhood association, the value may depends on the neighborhood sample values. In addition, in the process of regression analysis, the sample data may have outliers or missing values, resulting regression bias or errors. Nevertheless, existing scholars paid little attention to the neighborhood association effect of regression samples.

With these considerations in mind, this paper proposed a neighborhood statistical method that fully considering the neighborhood association effect. Under the assistance of this method, the value of geographic variables in each location could update a new value based on neighborhood association effect. As a typical region to study the relationship between vegetation and climate factors, a case study was selected in Inner Mongolia autonomous region, China. And then, the effectiveness of incorporating neighborhood association effect was tested by spatial regression models in the study area for comparison.

2. Study area, data and methodology

2.1. Study Area
Inner Mongolia autonomous region is located in the inner part of Eurasian continent and in the northern border of China (37°24′-53°23′N, 97°12′-26°04′E). The region extends with a long and narrow shape, the straight-line distance between east and west is about 2400km, while the north-south distance is about 1700km. The total land area is approximately 1.183 million km$^2$, accounting for 12.3 percent of the total area of China [17]. Influenced by the climate conditions such as temperature and precipitation, Inner Mongolia is rich in vegetation types, covered by coniferous forest, broad-leaved forest, forest grassland, temperate grassland and desert from east to west. As one of the most sensitive regions of global climate change, the ecological environment of Inner Mongolia is fragile and unstable [18]. Therefore, it is of great significance to explore the driving mechanism of climate change on vegetation growth based on neighborhood association effect.

2.2. Data
We obtained vegetation NDVI data during 2014-2018 from MODIS MOD13A1 product. The product provides two primary vegetation layers, including the normalized difference vegetation index (NDVI) and enhanced vegetation index (EVI). The spatial resolution of the product is 500 m and time resolution is 16 days. The data of the vegetation growth season (from June to August) were extracted, the MODIS reprojection tool (MRT) was adopted to mosaic and reproject the extracted data, and then the NDVI value at pixel level was calculated by maximum value composite (MVC) approach. Finally the annual NDVI spatial distribution maps for each year were acquired in this study. The meteorological data were from China meteorological science data sharing service platform (http://data.cma.cn/), including the monthly average temperature and cumulative precipitation of 110 meteorological stations in Inner Mongolia and its surrounding areas during 2014 to 2018. Kriging method was employed for spatial interpolation of precipitation and ANUSPLIN method was employed for spatial interpolation of temperature [19]. In order to keep consistent with NDVI data, the spatial resolution of meteorological data was also set to 500m. Then the annual average temperature and accumulative precipitation were calculated by ArcGIS (Version 10.2) for year 2014-2018 respectively. At last, the NDVI and meteorological data were clipped by Inner Mongolia boundary.
2.3. Methodology

2.3.1. Neighborhood statistical method. Spatial data are often affected by local neighborhood association effect in the spatially explicit analysis [20]. For raster dataset, the neighborhood statistical process is to use mathematical statistical method to calculate an updated value for each pixel, which represents the association effect with neighborhood pixels. To this end, the Moore neighborhood was adopted in this study. For every location in study area, the Moore neighborhood comprises eight cells immediately surrounding it, namely, a 3×3 rectangular grid neighborhood is defined. Here we extended Moore neighborhood into higher dimensions to explore the influence of different neighborhood sizes. Finally, five different Moore neighborhood sizes (3×3, 5×5, 7×7, 9×9, and 11×11) were selected (Figure 1). Essentially, 3×3 Moore neighborhood size contains 9 pixels, 5×5 Moore neighborhood size contains 25 pixels, and so on. Like focal statistic in spatial neighborhood analysis, neighborhood statistic updates an output raster where the value for each output pixel is a function of all input pixel values within its specified neighborhood. The function for input pixel values is generally pre-defined mathematical operations. Considering the data used in this study, we chose the mean mathematical function to represent the neighborhood association effect.

![Image](image.png)

Figure 1. The different Moore neighborhood sizes defined in this study.

2.3.2. Geographically Weighted Regression. In literature, the spatial nonstationary and heterogeneity relationships between vegetation productivity and climatic factors were confirmed in most of the previous studies in Inner Mongolia region [21,22]. In order to investigate the complex relationship between climate factors and vegetation index, and further examine the neighborhood statistical method introduced in this study. The geographically weighted regression (GWR) was applied to model the spatial nonstationarity of the correlation between NDVI and climate changes.

GWR is an effective local spatial analysis method, which is helpful to reveal the changes of spatial relations within the study area. GWR model is an extension of ordinary linear regression, the parameters of this method are the function of spatial position and tend to vary over the study area. The variation of the relationship between independent and dependent variables on the spatial scale is evaluated by obtaining local parameters. The model equation is defined below:

\[
y_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i)x_{ik} + \varepsilon_i
\]

(1)

Where \(y_i\) is the fitted value for the spatial unit \(i\), \((u_i, v_i)\) is the geographic coordinate of the \(i\)th spatial unit, \(\beta_0(u_i, v_i)\) is the estimated value of the constant term of the \(i\)th unit, \(\beta_k(u_i, v_i)\) is the \(k\)th regression parameter, \(x_{ik}\) is the \(k\)th independent variable, \(\varepsilon_i\) is the random error that obeys the independent normal distribution with the mean value of 0.

In this study, the GWR model was conducted at pixel level. The annual NDVI value was taken as dependent variable, while the climate factors were regarded as the corresponding independent
variables. For each year, the GWR model was tested with the raw data firstly. The next step is therefore to apply neighborhood focal statistic with five different Moore neighborhood sizes. As a result, five groups of updated variables were generated to represent the effect of neighborhood association. In order to verify the effectiveness of incorporating neighborhood association effect, the different GWR model with raw data and updated data were compared. Guided by previous studies, the indexes to evaluate the goodness of fit of GWR model were Akaike information criterion (AIC) and Adjusted $R^2$. Without loss of generality, the difference between the AIC values of the two models is greater than 3, which indicates that there are obvious differences between the two models, and the model with lower AIC value has better fitting effect [23]. Moreover, Adjusted $R^2$ refers to the accuracy of model fitting. The larger the Adjusted $R^2$ value, the better the model goodness of fit. The combination of the two indicators can reflect the explanatory power of different models.

3. Result

3.1. Spatial distribution of NDVI in Inner Mongolia during 2014 to 2018

With the help of remote sensing images, the spatial distribution of average NDVI value from 2014 to 2018 is shown in Figure 2. There was a clear regional and spatial difference of NDVI distribution across the Inner Mongolia, which demonstrated a decreasing trend from east to west in the study area. This is mainly ascribed to complex geomorphic types in this area, such as hills, plains, deserts, rivers and lakes. As a result, the vegetation types are divided into forest in the east, grassland in the middle and desert in the west, this results in uneven distribution of NDVI from the east to the west.

![Figure 2. Spatial distribution of average NDVI in Inner Mongolia during 2014 to 2018.](image)

3.2. Spatial distribution of climate factors in Inner Mongolia during 2014 to 2018

The average value of annual accumulated precipitation and annual mean temperature during 2014-2018 were calculated and the spatial distributions were displayed in Figure 2 respectively. The annual precipitation during five years ranged from 95.00mm to 543.24mm, while the annual mean temperature value was from -3.62°C to 11.14°C. Like vegetation index NDVI, the climate factors in the study area presented significant spatial variation. In Figure 2 for specific performance: the precipitation showed a decreasing trend from northeast to southwest, the temperature increased from northeast to southwest conversely. Overall, the NDVI and meteorological data exhibited obvious spatial nonstationarity in Inner Mongolia. Here the GWR model was introduced to explore the spatial nonstationarity between multiple factors.
3.3. Impact of neighborhood association effect by GWR Models

Five groups of updated data with different Moore neighborhood (3×3, 5×5, 7×7, 9×9, and 11×11 sizes) statistical method, as well as the raw data without any neighborhood statistic, were examined with GWR models for each year. The AIC value and Adjusted $R^2$ for different GWR models were summarized in table 1 for comparison.

According to the model fitting evaluation indexes, the regression effect of most GWR models using annual precipitation and mean temperature to fit NDVI is significant (Adjusted $R^2 > 0.8$). For every year, by comparing the goodness of fit of different GWR models, it was found that the AIC value decreased significantly when neighborhood statistical method were included in the model, the Adjusted $R^2$ increased at the same time. Besides, changing the neighborhood sizes also affected the model accuracy. As can be noticed, with the increase of neighborhood size, the model accuracy enhanced. The majority of Adjusted $R^2$ of GWR models considering neighborhood association effect were greater than 0.9. For example, based on neighborhood statistical method in 2014, the AIC value decreased from -1446.693 to -2763.662, the Adjusted $R^2$ increased from 0.853 to 0.955. For other years, the similar results were confirmed for model evaluation indexes. The result suggests that the updated data with neighborhood association effect are more effective for studying the NDVI modeling response to climate variability.

4. Discussion

This paper modeled NDVI-based vegetation in relationship to climate factors in Inner Mongolia, using MOIDS and meteorological data for 2014-2018. The neighborhood association effect was included in order to access the additional benefit when accounting for local spatial association. Firstly, vegetation plays important roles in green space ecosystem, thus understanding its pattern and trajectory in response to climate factors can monitor and predict vegetation dynamics. Through estimating different model metrics, the model accuracy result in this paper proved to be tremendous, illustrated that it was effective to apply GWR model on the influencing factors of vegetation change in Inner Mongolia.

| Year | Neighborhood Size | AIC       | Adjusted $R^2$ |
|------|-------------------|-----------|----------------|
| 2014 | 5×5               | -2216.178 | 0.925          |
|      | 3×3               | -1974.804 | 0.907          |
|      | -                 | -1446.693 | 0.853          |

Table 1. The comparison of goodness of fit between different GWR models.

Figure 3. Spatial distribution of annual precipitation (a) and annual mean temperature (b) in Inner Mongolia during 2014-2018.
Secondly, the regression results of vegetation NDVI with and without neighborhood association effect were analyzed and compared. Many studies have confirmed that neighborhood association effects are important driving forces during local spatial modeling [24,25]. It is therefore necessary to use regression modeling technology based on local effect to quantify the causal relationship between NDVI and meteorological forces in different spatial locations. In this study, compared with the model that did not consider neighborhood association effect, the updated model had evident improvement in the prediction power. It was worth noting that the inclusion of neighborhood association effect for spatial datasets was significant and necessary. The results could be explained by the following aspects. As a matter of fact, the sample value for spatial regression model is relatively unstable, especially when there are some missing values or outliers in the sample data, the fitting results of regression will be affected accordingly [26]. To cope with this, the samples in neighborhood area could be used to rectify the missing values or outliers. Furthermore, rely on spatial association theory, the value of spatial variable is not only affected by the current location, but also related to the neighborhood area. The study integrated neighborhood statistical method for every location, aiming to reflect the neighborhood association effect and understand the existing spatial configurations in local scale. The GWR model accuracy (both AIC and Adjusted R² value) corroborated the effectiveness of accounting for neighborhood association effect. The data containing neighborhood information could reduce the influence of outliers and missing values in samples, and better fit the relationship between NDVI and climate factors.

In addition, the size of neighborhood area was discovered as another important factor for measuring neighborhood association effect. Accompanied by larger neighborhood size in this study, the goodness of fit of the model increased. The expansion of neighborhood size introduced more local samples and abundant spatial contexts in the neighborhood statistical process, which can provide enriched neighborhood information and local configuration to a certain extent. As a result, building GWR models based on larger neighborhood size exhibited better explanatory power.
Although this study integrates neighborhood statistical method and different neighborhood sizes to improve the model goodness of fit, it still has several limitations. Future studies may incorporate more neighborhood statistical methods of local spatial configuration, which can result in improved model performance. Moreover, more detail work should be input to the neighborhood shapes and sizes.

5. Conclusion
This work provided a neighborhood statistical method to integrate neighborhood association effect in spatial data. The neighborhood association configurations, as expressed by five different Moore neighborhood sizes (3×3, 5×5, 7×7, 9×9, and 11×11) were decided. The results confirmed that neighborhood association effect could affect GWR model accuracy. The case study in Inner Mongolia highlighted the data containing neighborhood information were more stable with less randomness, and the regression model was more reliable after considering the neighborhood association effect. Meanwhile, the model accuracy performances further verified the effectiveness of including larger neighborhood sizes in the study. Therefore, the combination of spatial neighborhood association effect and GWR can effectively analyze and simulate the dynamic changes of vegetation based on climate factors in Inner Mongolia. The proposed method can be a useful tool for modeling vegetation dynamics.

Acknowledgments
This research was funded by the National Natural Science Foundation of China (NSFC): 41971352.

References
[1] Liu Y, Li L, Chen X, Zhang R and Yang J 2018 Temporal-spatial variations and influencing factors of vegetation cover in Xinjiang from 1982 to 2013 based on GIMMS-NDVI3g Glob. Planet. Change 169 145–55
[2] Wang H, Liu G, Li Z, Zhang L and Wang Z 2020 Processes and driving forces for changing vegetation ecosystem services: Insights from the Shaanxi Province of China Ecol. Indic. 112 106105
[3] Chu H, Venevsky S, Wu C and Wang M 2019 NDVI-based vegetation dynamics and its response to climate changes at Amur-Heilongjiang River Basin from 1982 to 2015 Sci. Total Environ. 650 2051–62
[4] Gottfried M, Pauli H, Futschik A, Akhalkatsi M, Barančok P, Alonso J L B, Coldea G, Dick J, Erschbamer B and Kazakis G 2012 Continent-wide response of mountain vegetation to climate change Nat. Clim. Chang. 2 111–5
[5] Wu D, Zhao X, Liang S, Zhou T, Huang K, Tang B and Zhao W 2015 Time-lag effects of global vegetation responses to climate change Glob. Chang. Biol. 21 3520–31
[6] Pound M J and Salzmann U 2017 Heterogeneity in global vegetation and terrestrial climate change during the late Eocene to early Oligocene transition Sci. Rep. 7 1–12
[7] Gandhi G M, Parthiban S, Thummalu N and Christy A 2015 Ndvi: Vegetation change detection using remote sensing and gis–A case study of Vellore District Procedia Comput. Sci. 57 1199–210
[8] Wen Y, Liu X, Yang J, Lin K and Du G 2019 NDVI indicated inter-seasonal non-uniform time-lag responses of terrestrial vegetation growth to daily maximum and minimum temperature Glob. Planet. Change 177 27–38
[9] Fang X, Zhu Q, Ren L, Chen H, Wang K and Peng C 2018 Large-scale detection of vegetation dynamics and their potential drivers using MODIS images and BFAST: A case study in Quebec, Canada Remote Sens. Environ. 206 391–402
[10] Hao L, Pan C, Fang D, Zhang X, Zhou D, Liu P, Liu Y and Sun G 2018 Quantifying the effects of overgrazing on mountainous watershed vegetation dynamics under a changing climate Sci. Total Environ. 639 1408–20
[11] Bivand R S and Wong D W S 2018 Comparing implementations of global and local indicators of spatial association TEST 27 716–48
[12] Munibah K, Widiatmaka and Widjaja H 2018 Spatial autocorrelation on public facility availability index with neighborhoods weight difference J. Reg. City Plan. 29 18–31
[13] Anselin L 1995 Local Indicators of Spatial Association—LISA Geogr. Anal. 27 93–115
[14] Anputhas M, Janmaat J J A, Nichol C F and Wei X (Adam) 2016 Modelling spatial association in pattern based land use simulation models J. Environ. Manage. 181 465–76
[15] Verburg P H, de Nijs T C M, van Eck J R, Visser H and de Jong K 2004 A method to analyse neighbourhood characteristics of land use patterns Comput. Environ. Urban Syst. 28 667–90
[16] Kern K B, Colberg T P, Wunder C, Newton C and Slepian M J 2019 A local neighborhood volunteer network improves response times for simulated cardiac arrest Resuscitation 144 131–6
[17] Cai Y, Zhao M, Shi Y and Khan I 2020 Assessing restoration benefit of grassland ecosystem incorporating preference heterogeneity empirical data from Inner Mongolia Autonomous Region Ecol. Indic. 117 106705
[18] Qiao Y, Zhu H, Li Y, Shao X, Zhong H, Shi H and Wu Z 2020 Large-Scale Spatial Patterns of Grassland Community Properties in the Inner Mongolia Autonomous Region, China Rangel. Ecol. Manag. 73 560–8
[19] Guo B, Zhang J, Meng X, Xu T and Song Y 2020 Long-term spatio-temporal precipitation variations in China with precipitation surface interpolated by ANUSPLIN Sci. Rep. 10 1–17
[20] Cromley E K, Wilson-Genderson M and Pruchno R A 2012 Neighborhood characteristics and depressive symptoms of older people: Local spatial analyses Soc. Sci. Med. 75 2307–16
[21] Guo L, Zuo L, Gao J, Jiang Y, Zhang Y, Ma S, Zou Y and Wu S 2020 Revealing the fingerprint of climate change in interannual NDVI variability among biomes in Inner Mongolia, China Remote Sens. 12 1332
[22] Wang Z, Zhong J, Lan H, Wang Z and Sha Z 2019 Association analysis between spatiotemporal variation of net primary productivity and its driving factors in inner mongolia, china during 1994–2013 Ecol. Indic. 105 355–64
[23] Sakamoto Y, Ishiguro M and Kitagawa G 1986 Akaike information criterion statistics Dordrecht, Netherlands D. Reidel 81 26853
[24] Anselin L 2019 A local indicator of multivariate spatial association: extending Geary’s C Geogr. Anal. 51 133–50
[25] Yu Y, Gorman B P and Hering A S 2020 Objective identification of local spatial structure for material characterization Stat. Anal. Data Min. ASA Data Sci. J. 13 377–93
[26] Yu Y, Workman A, Grasmick J G, Mooney M A and Hering A S 2018 Space-time outlier identification in a large ground deformation data set J. Qual. Technol. 50 431–45