Responses of NDVI to climate factors in Inner Mongolia using geographically weighted regression

Yuwei Wang \textsuperscript{1a,7}, Wang Gao\textsuperscript{1b}

\textsuperscript{1}School of Mathematics and Computer Science, Jianghan University, Wuhan 430056, China
\textsuperscript{a}weberwang@jhun.edu.cn, \textsuperscript{b}gaow@jhun.edu.cn

*Corresponding author

Abstract. Using the MODIS normalized difference vegetation index (NDVI) datasets along with the climate data (precipitation and temperature), geographically weighted regression (GWR) was constructed to explore the spatial nonstationarity and heterogeneity relationships between NDVI and climate factors in Inner Mongolia, China. Our work compared the results of GWR model accuracy with ordinary least squares (OLS) model. The results indicated that GWR method yielded better goodness of fit and higher model accuracy than OLS. Moreover, the GWR model could deeply reveal the complex relationship between NDVI and climate factors. At the same time, the research results could also provide scientific basis for vegetation modeling in Inner Mongolia and similar areas.

Abbreviations
NDVI Normalized difference vegetation index
GWR Geographically weighted regression
OLS Ordinary least squares
AP Annual precipitation
AAT Annual average temperature
AICc Corrected Akaike information criterion

1. Introduction
Vegetation is the indispensable part of terrestrial ecosystem, which plays an important role in environment and climate change\cite{1}. At present, the rapid development of high-resolution remote sensing technologies allow sensors to observe vegetation cover comprehensively. Hyperspectral and multispectral vegetation remote sensing data have been widely used in the study of vegetation productivity\cite{2,3}. Normalized difference vegetation index (NDVI) is a classical vegetation index for detecting vegetation cover\cite{4–6}. NDVI is calculated by near infrared band and red band, while the normalized index is used to reflect the surface characteristics of vegetation.

Inner Mongolia Autonomous Region is a typical arid and semi-arid area in China, which is one of the most sensitive areas of global climate change\cite{7}. Therefore, it is of great meaning to study the change of vegetation cover caused by climate factors. Based on geographic information system (GIS) and regression analysis, many studies have explored the corresponding relationship between NDVI and climate factors\cite{8–10}. 
The ordinary least squares (OLS) regression analysis method is a global modeling technique to study the driving forces of NDVI change, which estimates each force from a global perspective. However, previous research had reported that the relationships between NDVI and environment exhibited spatial nonstationarity and heterogeneity in Inner Mongolia\cite{11,12}. As a matter of fact, the influence of meteorological factors on NDVI varies in different spatial locations. Geographically weighted regression (GWR) model is the extension of traditional regression analysis, allowing parameter estimation at local scale. According to the spatial location information of the sample data, by means of the distance-decayed weight method, GWR model can generate the local regression coefficient varying with the spatial location, making regression results more credible. With these considerations in mind, the GWR model was introduced to explore the spatial heterogeneity between multiple factors. We compared the GWR model performance with OLS model, in order to verify the effectiveness of incorporating spatial heterogeneity in this study.

2. Study area, data and methodology

2.1. Study Area
Inner Mongolia Autonomous Region is located in the northern part of China, it serves as an important agricultural and animal husbandry production base in China. It is located between 37°24′ to 53°23′ North and 97°12′ to 126°04′ East, extending from northeast to southwest in a narrow shape\cite{13}. The annual average temperature of the area is 0-8 °C, while the annual total precipitation is between 50-450mm. The study area is rich in vegetation resources with a grassland coverage rate of 44%, including a large number of forests and grasslands\cite{14}. The GWR model on the response of vegetation NDVI to climate factors in Inner Mongolia can reveal the NDVI-environment varying relationships, which is of great significance to assess the impact of climate factors on ecosystem. In order to implement regression analysis and verification, 4000 points were generated randomly across whole region as samples (Figure 1).

2.2. Data
In order to investigate the vegetation NDVI in relationship to climate factors (including annual precipitation and average temperature), we collected NDVI and climate data in Inner Mongolia in 2018. The NDVI data were retrieved from MOD13A1 dataset of MODIS/Terra satellite. MOD13A1 were provided every 16 days at 500-meter spatial resolution. Maximum value composite (MVC) method was adopted to calculate the annual NDVI max value, representing the climax productivity of

![Figure 1. Location and sample points of the study area.](image-url)
the vegetation annually. The meteorological data originated from China meteorological data sharing service system (http://data.cma.cn/), the system includes 610 meteorological stations in China. The average monthly temperature (°C) and precipitation (mm) in the system were acquired. Based on the data of each station, spatial kriging interpolation was carried out to obtain continuous raster data nationwide. Then the annual precipitation (AP) and annual average temperature (AAT) were calculated in this study. Finally, NDVI and climate data were clipped by the boundary shape file of the study region.

2.3. Methodology

2.3.1. Geographically weighted regression. Inner Mongolia has a wide variety of vegetation species, and the distribution of vegetation productivity is uneven in the whole study area. Furthermore, on the basis of the accumulation of previous research, the AP and AAT exist spatial heterogeneity[15]. In order to simulate the relationship between vegetation NDVI and climate factors, geographically weighted regression (GWR) was applied to conduct multivariate association analysis between different variables[16,17]. The equation of GWR is defined as follows[18]:

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

(1)

where \( y_i \) is the dependent variable, \((u_i, v_i)\) is the geographical coordinates of location \( i \), \( x_{ik} \) is the \( k \)th independent variable in regression, \( \varepsilon_i \) is the error term of the observation value at location \( i \), the parameter \( \beta_k (u_i, v_i) \) changes in space to measure the nonstationarity of different locations.

2.3.2. Model accuracy test. The goodness of fit of regression model was evaluated and compared by \( R^2 \), adjusted \( R^2 \) and corrected Akaike information criterion (AICc)[19] in this research. \( R^2 \) and adjusted \( R^2 \) vary from 0 to 1, the larger value indicates better model accuracy. Corrected Akaike information criterion (AICc) is followed by the equation:

\[ AIC_c = 2n \ln(\hat{\sigma}) + n \ln(2\pi) + n \left( 1 + \frac{\text{tr}(S)}{n - 2} \right) \]  

(2)

where the subscript \( c \) represents the "corrected" estimate value, \( n \) is the size of the sample, \( \hat{\sigma} \) is the standard deviation of error term estimation, \( \text{tr}(S) \) is the trace of matrix \( S \) of GWR. If the difference of AICc value between the two models is greater than 3, the model with lower AICc value will be considered as a preferred model.

In this study, the NDVI was taken as dependent variable, while AP and AAT were two independent variables. In order to verify the effectiveness of GWR method, for the study year, the OLS and GWR model was tested, respectively. A total of 4000 points were generated randomly to fit and test different models. Then the indexes to evaluate the model were compared.

3. Result

3.1. Spatial heterogeneity of NDVI in relation to climate factors

The result of GWR outputs, including regression coefficient of AP, coefficient of AAT, local \( R^2 \) and intercept were visualized in Figure 2.
As shown in Figure 2, there existed obvious spatial heterogeneity between NDVI and climatic factors in Inner Mongolia. The AP coefficients were always positively related with NDVI in study area (Fig. 2a). It could be interpreted that higher NDVI value emerged with abundant precipitation. Meanwhile, the coefficients of AAT, i.e., the relationship between NDVI and AAT, varied across the whole region and was negatively correlated with NDVI in most areas. Specially, AAT was positively correlated with NDVI in areas of Xilingol, Chifeng and Baotou (Fig. 2b). In addition, the local $R^2$ (Fig. 2c) and intercept (Fig. 2d) of the GWR model further confirmed the spatial heterogeneity among multivariable. The higher local $R^2$ values ($R^2 > 0.5$) mainly occurred in the northeast part of Inner Mongolia.

3.2. Model accuracy comparison

The model accuracy indexes for OLS and GWR models, expressed by AICc, $R^2$ and adjust adjusted $R^2$, were listed in Table 1. With smaller AICc values, higher $R^2$ and adjust adjusted $R^2$ values in all situations, the goodness of fit of GWR model was distinctly better than OLS model. For instance, when employing AP+AAT as independent variables, the AICc was -6309.19 for GWR model and -4639.84 for OLS model; the $R^2$ of GWR model was 0.857, compared with 0.782 for OLS. Besides, applying AAT as independent variable could explain the 45% variation of the NDVI in OLS model, while this value increased to 80% in GWR model.

We also tested three different combinations of independent variables, namely AP, AAT, and AP+AAT simultaneously. The $R^2$ value for AP in OLS were 0.71 and 0.85 in GWR. In the same way,
the values for AAT in OLS and GWR were 0.45 and 0.80. Results showed, relative to temperature, the precipitation had more significant effect on NDVI.

Table 1. Model accuracy results between OLS and GWR under different regression conditions.

| Model-Index       | AP     | AAT    | AP+AAT  |
|-------------------|--------|--------|---------|
| OLS-AICc          | -3554.28 | -943.18 | -4639.84 |
| GWR-AICc          | -6192.01 | -5122.17 | -6309.19 |
| OLS-R²            | 0.7153  | 0.4531  | 0.7831  |
| GWR-R²            | 0.8533  | 0.8083  | 0.8578  |
| OLS-Adjusted R²   | 0.7152  | 0.4530  | 0.7829  |
| GWR-Adjusted R²   | 0.8530  | 0.8079  | 0.8574  |

Model-Index: Regression model and the index selected to evaluate the model accuracy.

4. Discussion
In this paper, geographically weighted regression was used to estimate the complex relationship between vegetation NDVI and climate factors. The spatial nonstationarity and heterogeneity were fully considered in the process of spatial regression modeling analysis.

Firstly, NDVI is an important index to reflect vegetation cover. Accurate prediction of the impact of climate change on vegetation distribution is of vital importance for both scientific researches and policymakers[20,21]. Secondly, the spatial variables in Inner Mongolia, such as vegetation coverage, precipitation and temperature, have significant spatial heterogeneity. In this research, the model accuracy result proved the GWR model was more effective than OLS model. Also, the complex NDVI-environment relationships were observed by GWR. It was found that AP was positively related to NDVI in the whole study area. In contrast, AAT demonstrated a complex relationship with NDVI, both positive and negative correlations were confirmed. Overall, it is necessary to use regression modeling technology based on local effect to quantify the causal relationship between NDVI and meteorological forces in different spatial locations.

This current work indicated that GWR was superior to OLS in many aspects. The future work will integrate more related predictor variables, including socio-economic data, which can theoretically result in improved model performance.

5. Conclusion
Many studies have confirmed that climate factors are important driving forces for NDVI. This study adopted geographically weighted regression to explore the spatial heterogeneity relationships between NDVI and climate factors in Inner Mongolia, China. Compared with OLS model, our work verified that GWR model improved the model performance substantially. Due to the existing spatial heterogeneity in the study area, space-varying relationships between NDVI and climate factors were observed. The precipitation and temperature appeared to be critical factors for modeling NDVI in Inner Mongolia.

References
[1] Richardson A D et al 2018 Ecosystem warming extends vegetation activity but heightens vulnerability to cold temperatures Nature 560 368–71
[2] Liu Y, Zhang F, Wang C, Wu S, Liu J, Xu A, Pan K and Pan X 2019 Estimating the soil salinity over partially vegetated surfaces from multispectral remote sensing image using non-negative matrix factorization Geoderma 354 113887
[3] Liu N, Budkewitsch P and Treitz P 2017 Examining spectral reflectance features related to Arctic percent vegetation cover: Implications for hyperspectral remote sensing of Arctic tundra Remote Sens. Environ. 192 58–72
[4] Baeza S and Paruelo J M 2020 Land use/land cover change (2000-2014) in the rio de la plata grasslands: An analysis based on MODIS NDVI time series Remote Sens. 12 381
[5] Duan T, Chapman S C, Guo Y and Zheng B 2017 Dynamic monitoring of NDVI in wheat agronomy and breeding trials using an unmanned aerial vehicle \textit{F. Crop. Res.} 210 71–80

[6] Zhang Y, Ling F, Foody G M, Ge Y, Boyd D S, Li X, Du Y and Atkinson P M 2019 Mapping annual forest cover by fusing PALSAR/PALSAR-2 and MODIS NDVI during 2007–2016 \textit{Remote Sens. Environ.} 224 74–91

[7] DU B, ZHEN L, HU Y, YAN H, Rudolf D G and Rik L 2018 Comparison of ecosystem services provided by grasslands with different utilization patterns in China’s Inner Mongolia Autonomous Region \textit{J. Geor. Sci.} 28 1399–1414

[8] 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 \textit{Remote Sens.} 12 1332

[9] 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 \textit{Sci. Total Environ.} 650 2051–62

[10] Pang G, Wang X and Yang M 2017 Using the NDVI to identify variations in, and responses of, vegetation to climate change on the Tibetan Plateau from 1982 to 2012 \textit{Quat. Int.} 444 87–96

[11] Pei Z, Fang S, Yang W, Wang L, Wu M and Zhang Q 2019 The Relationship between NDVI and Climate Factors at Different Monthly Time Scales : A Case Study of Grasslands in Inner Mongolia, China (1982-2015) \textit{Sustainability} 11 7243

[12] Miao B, Li Z, Liang C, Wang L, Jia C, Bao F and Jiang C 2019 The Relationship between NDVI and Climate Factors at Different Monthly Time Scales : A Case Study of Grasslands in Inner Mongolia, China (1982-2015) \textit{Sustainability} 11 7243

[13] Wei J, Yang Y, Hobbie E A and Yan W 2019 Ectomycorrhizal fungal resources of Inner Mongolia Autonomous Region: above ground macrofungi and belowground ectomycorrhizae \textit{New For.} 50 521–37

[14] Hu X, Hirota M, Wuyunna, Kawada K, Li H, Meng S, Tamura K and Kamiyo T 2019 Responses in gross primary production of Stipa krylovii and Allium polyrhizum to a temporal rainfall in a temperate grassland of Inner Mongolia, China \textit{J. Arid Land} 11 824–36

[15] Zhao Z, Gao J, Wang Y, Liu J and Li S 2015 Exploring spatially variable relationships between NDVI and climatic factors in a transition zone using geographically weighted regression \textit{Theor. Appl. Climatol.} 120 507–19

[16] Wang D \textit{et al} 2020 Modeling soil organic carbon spatial distribution for a complex terrain based on geographically weighted regression in the eastern Qinghai-Tibetan Plateau \textit{Catena} 187 104399

[17] Li Y, Liu X, Han Z and Dou J 2020 Spatial proximity-based geographically weighted regression model for landslide susceptibility assessment: A case study of Qingchuan area, China \textit{Appl. Sci.} 10 1107

[18] Fotheringham A S, Charlton M E and Brunsdon C 1998 Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis \textit{Environ. Plan. A} 30 1905–27

[19] Akaike H 1974 A New Look at the Statistical Model Identification \textit{IEEE Trans. Automat. Contr.} 19 716–23

[20] Zhang H, Chang J, Zhang L, Wang Y, Li Y and Wang X 2018 NDVI dynamic changes and their relationship with meteorological factors and soil moisture \textit{Environ. Earth Sci.} 77 582

[21] Zewdie W, Csaplovics E and Inostroza L 2017 Monitoring ecosystem dynamics in northwestern Ethiopia using NDVI and climate variables to assess long term trends in dryland vegetation variability \textit{Appl. Geogr.} 79 167–78