Spatiotemporal analysis of Population aging based on Geographically and temporally neural network weighted regression

Zilin Shen, Xin Fu, Fan Wang, Duodan Chen

College of Science Minzu University of China, Beijing, 100081, Beijing, China

Keywords: population aging; spatiotemporal non-stationarity; GTNNWR

Abstract: Since the 21st century, the entire world has entered an era of population aging, with a marked deepening of the aging process in developed countries. Therefore, taking into account the two dimensions of time and space to study aging has gradually become a hot issue. However, to analyze the aging rate and its influencing factors from the perspective of regression analysis, we will encounter the problem of how to solve the spatiotemporal non-stationarity. The GWR and GTWR models might encounter challenges when addressing the intricate nonlinear interaction between time and space in reality. Based on this, this paper uses the newly proposed Geographically and temporally neural network weighted regression (GTNNWR) model. By constructing a fully connected neural network model with multiple hidden layers and combining with Dropout technology and other techniques, GTNNWR model can effectively generate space-time weight matrix, and then solve the spatiotemporal non-stationarity. This paper consults the statistical yearbook of the National Bureau of Statistics of China from 2000 to 2020 and used VIF test to select 7 variables without multicollinearity. GTNNWR model is used to analyze aging and its influencing factors, and the results are compared with OLS, GWR and GTWR models. It is proved that GTNNWR has better performance than other models in modelling the spatiotemporal non-stationarity of population aging, and the influencing factors of population aging are analysed according to the results of the final model.

1. Introduction

Population aging has aroused great attention in the world since it was formally proposed. Internationally, there is a general consensus that population aging is considered mild when 10% to 20% of the population is aged over 60, moderate when it ranges from 20% to 30%, and severe when it exceeds 30%. In recent years, to alleviate population aging, many countries and world organizations have formulated a series of policies on aging. For example, in 2010, France enacted the pension reform Act, which decided to gradually delay the retirement age and pension time and raise the contribution level to reduce the pressure on the welfare deficit.

If we consider both time and space, we will encounter the problem of how to solve the spatiotemporal non-stationarity when studying the aging rate and its influencing factors from the perspective of regression analysis. In dealing with spatiotemporal non-stationarity, GTWR model is commonly used at present. However, because the GTWR model is still mainly represented by the linear weighted combination of distance measures, it is challenging to fully fit the complex nonlinear characteristics of real geographical relations, which may lead to problems in solving the spatiotemporal non-stationarity. By constructing a fully connected neural network model containing multiple hidden layers and combining Dropout technology, Parametric ReLU activation function and batch normalization technique, GTNNWR model can effectively deal with complex nonlinear interactions of time and space, generate space-time weight matrix, and then solve space-time non-stationarity. At present, the relevant research results have been applied in the estimation of PM2.5 concentration, and have proved to be significantly reliable [1-2].

At present, China is not in the ranks of serious aging countries. China's aging rate remains below that of developed nations like Japan, Europe, and the United States, indicating significant room for further development. Due to China's large population, vast territory and diverse geographical environment, rapid development has been achieved in the past two decades in terms of economy, population and medical facilities. There are great differences in the situation of aging among different

  DOI: 10.25236/etmhs.2024.076

provinces in China, and the aging process is complicated and changeable in time and space. Based on the GTNNWR model and the official data from the Statistical Yearbook of the National Bureau of Statistics of China from 2000 to 2020, considering the dimensions of time and space, the process of population aging is modeled. Compared with other models such as GTWR, it is proved that GTNNWR has better performance than other models in solving the spatiotemporal non-stationarity of population aging, and the influencing factors of population aging are analyzed according to the results of the final model.

2. Data description

2.1 Dataset and variable selection

Due to the large number of provinces in China and distinct spatial layout, the development of population aging in different provinces in the past 20 years is different, which can well meet the data needs of our study. The data used in this paper are from the statistical yearbook of the National Bureau of Statistics of China, and the data of investigated provinces from 2000 to 2020 are used, so as to construct a dataset of population aging.

After searching related articles about population aging, 11 independent variables are selected under five factors: population itself, economic development, social development, culture and education, and medical and health care. All variable names and their meanings are shown in table 1 [3].

Table 1 Variable name and definition

Name	definition
Degree of population aging(AG)	The proportion of the population aged 65 and older to the total population of each province in the same year.
Birth rate (BIRTH)	The ratio of annual births in each province to the province's total population's years of survival.
Death rate (DEATH)	The ratio of annual deaths in each province to the population living within the same year.
Average household size (AHS)	The ratio of the year-end total population of each province to the number of households within the same year.
Total population density (TPD)	The proportion of each province's year-end total population to the country's total population within the same year.
Elderly population density (EPD)	The proportion of the population aged 65 and older in each province to the total population aged 65 and older in the country in a year.
GDP per capita (GDP)	The per capita GDP in each province in a year.
Net migration (NM)	The difference between the number of people moving in and out of each province in a year.
Elderly population density (EPD)	Ratio of urban population to total population of the same year in each province.
Illiteracy rate per hundred labor force (ILLITERACY)	The number of illiterate individuals per one hundred labor force in each provincial-level region in a certain year.
Number of hospital beds per thousand people (NHB)	The number of hospital beds available per one thousand people in each provincial-level region in a certain year.

2.2 Data preprocessing

Due to the presence of missing data for certain years in the dataset, data preprocessing is conducted. Considering the potential spatial correlation among the development of population aging across different provinces in China, in order to mitigate the influence of spatial location and spatial correlation on interpolation results, we employ the Kriging interpolation method to seek smoother and more reasonable interpolation outcomes.

Because of the large number of selected independent variables, it is necessary to assess the degree of correlation among them before modeling. Strong multicollinearity among the selected variables can lead to model instability and inaccurate estimation. We conducted a variance inflation factor (VIF) test to examine the correlation among the variables. The results are shown in table 2.

Table 2 VIF for all variables

Variables	VIF
Constant	-
BIRTH	2.707
DEATH	1.53
AHS	3.456
TPD	17.351
EPD	18.88
GDP	3.379
NM	1.255
EPD	5.529
ILLITERACY	1.492
NHB	1.781

The results indicate that the VIF values for Total population density, Elderly population density, and Urbanization rate are all greater than 5, suggesting the presence of multicollinearity among these three types of independent variables. Therefore, we remove these three variables. After conducting another VIF test, the updated results are shown in table 3:

Table 3 VIF after removing variables

Variables	VIF
Constant	-
BIRTH	2.226
DEATH	1.049
AHS	2.989
GDP	1.779
NM	1.106
ILLITERACY	1.302
NHB	1.598

Due to significant differences in the scales of the selected data, it often leads to a greater influence of features with larger numerical values on the model during the training process. Therefore, we standardize the data to ensure a relatively balanced weighting of each feature. As the selected data consist of panel data, we currently standardize the data for each year within the dataset.

3. Methods

3.1 OLS, GWR and GTWR

To verify the performance of GTNNWR when dealing with population aging modeling that considers both time and space, this paper uses common OLS, GWR and GTWR models for comparison and verification.

The OLS population aging regression model constructed in this paper is as follows:

$$\begin{aligned} AG_i = \beta_0 + \beta_1 \times BIRTH_i + \beta_2 \times DEATH_i + \beta_3 \times AHS_i + \beta_4 \times GDP_i + \beta_5 \times NM_i + \\ \beta_6 \times ILLITERACY_i + \beta_7 \times NHB_i + \varepsilon_i, i = 1, 2, \cdots, n \end{aligned} \tag{1}$$

where $\beta_1, \cdots, \beta_n$ are the regressive coefficients of the independent variables, β_0 represents the regression constant, and ε_i represents the error term. For the estimation of regression coefficient, the ordinary least square method is usually used, and the estimated value of the dependent variable

$\widehat{AG} = H \times AG$ is finally obtained through the hat matrix $H = X(X^T X)^{-1} X^T$.

where,

$$AG = \begin{vmatrix} AG_1 \\ AG_2 \\ \cdots \\ AG_n \end{vmatrix}, X = \begin{bmatrix} 1 & BIRTH_1 & DEATH_1 & NHB_1 \\ 1 & BIRTH_2 & DEATH_2 & NHB_2 \\ \cdots & \cdots & \cdots & \cdots \\ 1 & BIRTH_n & DEATH_n & NHB_n \end{bmatrix} \tag{2}$$

In order to deal with the spatial non-stationarity in geographical relations, GWR model was introduced by Brunsdon *et al*. (1998) [4]. The regression coefficients of the GWR model can vary with spatial position changes. The GWR population aging regression model constructed in this paper is as follows:

$$\begin{aligned} AG_i = \beta_0(u_i,v_i) + \beta_1(u_i,v_i) \times BIRTH_i + \beta_2(u_i,v_i) \times DEATH_i + \beta_3(u_i,v_i) \times AHS_i + \beta_4(u_i,v_i) \times \\ GDP_i + \beta_5(u_i,v_i) \times NM_i + \beta_6(u_i,v_i) \times ILLITERACY_i + \beta_7(u_i,v_i) \times NHB_i + \varepsilon_i, i = 1,2,\cdots,n \end{aligned} \tag{3}$$

where $\beta_j(u_i,v_i)$ represents the regressive coefficient of j of point i, the value of which is related to the spatial coordinates of the sample point, namely the latitude and longitude (u_i,v_i). The error term ε_i satisfies $\varepsilon_i \sim N(0,\sigma^2)$ and $Cov(\varepsilon_i,\varepsilon_j) = 0(i \neq j)$.

If we consider the dimension of time, we need to solve the spatial non-stationarity and the temporal non-stationarity at the same time. In order to solve this problem, Huang *et al*. (2010) constructed a metric expression of spatial-temporal distance and proposed GTWR model [5]. The GTWR population aging regression model constructed in this paper is as follows:

$$\begin{aligned} AG_i = \beta_0(u_i,v_i,t_i) + \beta_1(u_i,v_i,t_i) \times BIRTH_i + \beta_2(u_i,v_i,t_i) \times DEATH_i + \beta_3(u_i,v_i,t_i) \times AHS_i + \beta_4(u_i,v_i,t_i) \\ \times GDP_i + \beta_5(u_i,v_i,t_i) \times NM_i + \beta_6(u_i,v_i,t_i) \times ILLITERACY_i + \beta_7(u_i,v_i,t_i) \times NHB_i + \varepsilon_i, i = 1,2,\cdots,n \end{aligned} \tag{4}$$

where $\beta_j(u_i,v_i,t_i)$ indicates the coefficient of regression for j of point i. The value of $\beta_j(u_i,v_i,t_i)$ is related to the spatial coordinates (u_i,v_i) and temporal coordinates t_i of the point i. The error term satisfies $\varepsilon_i \sim N(0,\sigma^2)$ and $Cov(\varepsilon_i,\varepsilon_j) = 0(i \neq j)$.

For GWR and GTWR models, the estimation ideas of regressive coefficient are approximate. The following takes GTWR model as an example, and the regressive coefficient is estimated as follows:

$$\hat{\beta}(u_i,v_i,t_i) = (X^T W(u_i,v_i,t_i) X)^{-1} X^T W(u_i,v_i,t_i) AG \tag{5}$$

where $W(u_i,v_i,t_i) = diag[w_0(u_i,v_i,t_i), w_1(u_i,v_i,t_i), \cdots, w_7(u_i,v_i,t_i)]$ represents the space-time weight matrix of the sample point i, and its diagonal element represents the geographical weight. The fitting accuracy of the model estimate is determined by $W(u_i,v_i,t_i)$. The solution of the weight matrix is caculated from the weight kernel function, which is divided into fixed type and adaptive type. After comparing the performance of model established by the two types of kernel functions, the kernel function selected in this paper is a fixed Gaussian kernel function:

$$w_{ij} = \exp(-(d_{ij}^S)^2 / b^2) \tag{6}$$

where b represents bandwidth and characterizes the attenuation relationship of weight with distance $d_{ij}^S = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2}$. In this paper, the AICc criterion is employed to select the optimal bandwidth.

3.2 Geographically and Temporally Neural Network Weighted Regression

Due to the GTWR model's reliance on a simplistic weighted combination of time and space distance for measuring spatiotemporal distance, it may face challenges when dealing with intricate nonlinear spatiotemporal interactions. This paper uses the GTNNWR model proposed by Wu *et al.* (2021). The GTNNWR model extends the GNNWR model to include time to estimate the spatial non-stationarity, so as to model the spatiotemporal non-stationarity relationship.

We construct the GTNNWR population aging spatiotemporal regression model:

$$AG_i = w_0(u_i, v_i, t_i) \times \beta_{0OLS} + \sum_{j=1}^{7} w_j(u_i, v_i, t_i) \times \beta_{jOLS} x_{ij} + \varepsilon_i, i = 1, 2, \cdots, n \tag{7}$$

where $\varepsilon_i \sim N(0, \sigma^2)$ and $Cov(\varepsilon_i, \varepsilon_j) = 0 (i \neq j)$. $w_j(u_i, v_i, t_i)$ represents the spatiotemporal non-stationary weights of OLS coefficients β_{jOLS} x_{ij} denotes the value of variable *j* at point *i*. Similar to GTWR, the validity of model estimation depends on the solution of space-time weight matrix $W(u_i, v_i, t_i) = diag[w_0(u_i, v_i, t_i), w_1(u_i, v_i, t_i), \cdots, w_7(u_i, v_i, t_i)]$. In this regard, Wu *et al.* (2021) proposed generating spatiotemporal proximity STPNN based on spatiotemporal distance, and constructed spatiotemporal weighted neural network STWNN for solving

$$W(u_i, v_i, t_i) = STWNN([STPNN(d_{i1}^S, d_{i1}^T), \cdots, STPNN(d_{in}^S, d_{in}^T)]^T) \tag{8}$$

Where $d_{ij}^T = \sqrt{(t_i - t_j)^2}$ denotes the time distance between sample points. STPNN is a neural network structure composed of an input layer, a hidden layer and an output layer. For estimated point *i*, after the spatial and time distances between *i* and other points $[(d_{i1}^S, d_{i1}^T), \cdots, (d_{in}^S, d_{in}^T)]$ are entered, STPNN outputs the spatiotemporal proximity vector of point *i* and all other points $[d_{i1}^{ST}, \cdots, d_{in}^{ST}]^T$ .Then, with spatiotemporal proximity vector being the input layer, the STWNN will obtain the space-time weight matrix of point i through multiple hidden layers, and then the population aging regression estimation of point *i* is completed:

$$\widehat{AG_i} = x_i^T W(u_i, v_i, t_i)(X^T X)^{-1} XAG \tag{9}$$

The optimization algorithm for each hidden layer of STPNN and STWNN incorporates Dropout technology, Parametric ReLU activation function, and batch normalization technique for training.

3.3 Method evaluation

In this paper, R^2 and RMSE are utilized to assess the performance of the model:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (\widehat{y_i} - y_i)^2}{\sum_{i=1}^{n} (y_i - \overline{y})^2} \tag{10}$$

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(y_i - y_i)^2}{n}} \tag{11}$$

4. The modeling process of dataset

4.1 Model Construction for population aging regression using GTNNWR

Based on the construction approach of the GNNWR model by Du et al. (2020), the GTNNWR model in this paper splits the dataset into training, validation, and test sets with proportions of 0.75, 0.15, and 0.1, respectively. The training set uses the mini-batch stochastic gradient descent algorithm, with MSE serving as the model's training loss function [6]. The validation set is employed to record the optimal prediction accuracy and model parameters based on MSE in each epoch. Moreover, if the MSE of the validation set fails to decrease for 1000 consecutive epochs, it is interpreted as a sign of overfitting, leading to the stop of training. The most recently recorded optimal model is considered the final model, which is used to fit the test set and conduct model evaluation.

We configure the batch size as 64 and the Dropout ratio as 0.1. After experiments, we find that when the number of epochs exceeds 2000, the model results basically does not change, so we set the maximum number of epochs to 2500. Three hidden layers are constructed in GTNNWR in this paper, and the size of neurons is determined as [1024,256,32]. The optimal learning rate of each epoch is traversed by 0.001 to 0.05, and the optimal learning rate of the final model is 0.0033.

4.2 Model evaluation and result analysis

To comprehensively compare model performance, we fit the dataset using GTNNWR model and then similarly fit the dataset using other models, obtaining fitting results for all four models.

Table 4 Optimization parameters

Models	R^2	RMSE
OLS	0.464	0.179
GWR	0.669	0.141
GTWR	0.912	0.073
GTNNWR	0.928	0.065

As shown in Table 4, both from the perspective of R^2 and RMSE, the performance of OLS is the poorest. The performance of GWR is slightly better than OLS, but overall, there is a significant gap compared to the GTWR and GTNNWR models, which consider the dimension time. This indicates that the dimension time, as a fundamental geographical attribute, is crucial in population aging analysis. GTWR achieves a R^2 as high as 0.912 and an RMSE of 0.073, demonstrating its strong capability in handling temporal and spatial non-stationarity. However, GTNNWR reaches a R^2 of 0.928 and an RMSE of 0.065, indicating superior performance compared to GTWR. This suggests that GTNNWR exhibits the optimal ability to handle spatiotemporal non-stationary features in population aging spatiotemporal regression analysis.

To further examine the overall spatiotemporal non-stationarity of the spatiotemporal regression for provinces investigated in the dataset, we refer to the F_1-statistics established by Du et al. (2020):

$$F_1 = \frac{RSS_{GTNNWR} / \delta_1}{RSS_{OLS} / (n-p-1)} \tag{12}$$

where, RSS_{GTNNWR} and RSS_{OLS} respectively represent the Residual Sum of Squares of GTNNWR and OLS. The mean of RSS_{GTNNWR} / σ^2 is δ_1 and the variance is $2\delta_2$, so is approximately an F-distribution with degrees of freedom δ_1^2 / δ_2 and $n-p-1$.If the value of F_1 approaches to 1, it indicates that there is insignificant disparity between RSS_{GTNNWR} and RSS_{OLS}.Given the significance level α=0.05, the F_1 value of the model output result is 0.09984, with a P-value of 0.01. Therefore, we reject the null hypothesis that there is no significant difference between RSS_{GTNNWR} and RSS_{OLS}, and conclude that compared to OLS, the GTNNWR model significantly improves the goodness of fit, indicating significant spatiotemporal non-stationarity.

Table 5 shows the regressive coefficients of each variable predicted by the GTNNWR model. The coefficients of each variable in the model show variations in direction, implying that the impact of independent variables on the aging rate could differ across various spatial and temporal contexts. From the perspective of the average value of coefficient, among the seven independent variables included in the model, the effect of average household size on the extent of population aging is notably negative. This indicates that at a given sampling time, provinces with higher average number of people per household tend to have a lower proportion of elderly population. The influence of Death rate on the extent of population aging is notably positive. This indicates that in a given sampling time, the percentage of elderly residents in provinces with high death rates is frequently higher, which is consistent with the reality. GDP per capita has a positive impact on population aging, and the birth rate has a negative effect on population aging. Other variables such as net migration, illiteracy rate per hundred labor force, the number of hospital beds per thousand people have a relatively insignificant effect on population aging.

Table 5 Parameter estimates of GTNNWR model

Variables	M	Max	Mean
Constant	0.542	1.556	0.937
BIRTH	-0.420	0.069	-0.106
DEATH	-0.226	0.776	0.265
AHS	-0.789	0.039	-0.270
GDP	-0.110	0.534	0.111
NM	-0.112	0.085	-0.006
ILLITERACY	-0.046	0.081	0.009
NHB	-0.147	0.073	0.014

5. Conclusions

In this paper, the GTNNWR model is utilized to analyze population aging and its influencing factors in terms of time and space. The model test result shows that the GTNNWR model has significant spatiotemporal non-stationarity. The fitting results show that the R square is 0.928 and the RMSE is 0.065. After comparison, the GTNNWR model is superior to OLS, GWR and GTWR both from the point of view of R square and RMSE. It shows that GTNNWR still has a good ability to deal with spatiotemporal non-stationarity in the spatiotemporal regression analysis of population aging, and solves the possible problems of GWR and GTWR in effectively dealing with complex temporal and spatial nonlinear interactions in reality. According to the coefficient of independent variables, death rate and GDP per capita have a positive correlation with the degree of population aging. Birth rate and average household size have a negative impact on the degree of population aging. Net migration, illiteracy rate per hundred labor force and number of hospital beds per thousand people have a relatively insignificant impact on the degree of population aging.

References

[1] Wu, Sensen, et al. "Geographically and temporally neural network weighted regression for modeling spatiotemporal non-stationary relationships." International Journal of Geographical Information Science 35.3 (2021): 582-608.

[2] Wang, Weilin, et al. "Estimation of PM2. 5 concentrations in China using a spatial back propagation neural network." Scientific reports 9.1 (2019): 13788.

[3] Anderson, Gerard F., and Peter Sotir Hussey. "Population Aging: A Comparison Among Industrialized Countries: Populations around the world are growing older, but the trends are not cause for despair." Health affairs 19.3 (2000): 191-203.

[4] Brunsdon, Chris, Stewart Fotheringham, and Martin Charlton. "Geographically weighted regression." *Journal of the Royal Statistical Society: Series D (The Statistician)* 47.3 (1998): 431-443.

[5] Huang, Bo, Bo Wu, and Michael Barry. "Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices." *International journal of geographical information science* 24.3 (2010): 383-401.

[6] Du, Zhenhong, et al. "Geographically neural network weighted regression for the accurate estimation of spatial non-stationarity." International Journal of Geographical Information Science 34.7 (2020): 1353-1377.