GEOGRAPHICALLY WEIGHTED REGRESSION ON THE ECOLOGICAL FACTORS OF HUMAN LONGEVITY IN GANG-WON PROVINCE, KOREA

Don-Jeong Choi[*a], Yong-Chul Suh[b]

[a] Dept. of Spatial Information Engineering , 599-1 Daeyeon 3-Dong, Nam-Gu, Busan 608-737 South Korea; Tel; +82-51-629-6654, Pukyung National UNIV. E-mail: choidonjeong@gmail.com

[b] Dept. of Spatial Information Engineering, 599-1 Daeyeon 3-Dong, Nam-Gu, Busan 608-737 South Korea; Tel; +82-51-629-6658, Pukyung National UNIV. E-mail: suh@pknu.ac.kr

KEY WORDS: Human Longevity, Ordinary Least Square, Geographically Weighted Regression, Ecological Factors

Abstract: Ordinary least square regression model(OLS) through assumptions the correlations between distribution of longevity population and ecological factors are same in all regions. Therefore it cannot explain the continuous characteristics of ecological data and its spatial variation. Geographically Weighted Regression(GWR) model can quantitatively calculate the spatial similarity of adjacent areas through Geographical Weighting Function. In addition GWR can be described local spatial variation of the distribution of longevity population reflect the environmental characteristics. In this paper we performed a comparative analysis between OLS and GWR model about ecological factors of human longevity that proposed in previous studies.

INTRODUCTION

Existing research on the longevity factor shows that the individual's genetic and social factors are most closely with human longevity. However, in recent years, the regional ecological factors also closely associated with longevity revealed (Pesa et al., 2011). The regional ecological condition has associated with the sanitary conditions of the social system. It has gaining attention from the point of impact on the health status of the elderly people. In also, some specific factors are being raised that associated with human longevity. Previous studies related to ecological factors showing OLS based analysis patterns in the grasp of the longevity factor. However, The basic assumption of OLS model that the quartered acid of residual cause for errors in the spatial phenomena. In addition, the derived residual by regression equation assumed homoscedasticity and sum of variance is to be zero. Which mean every spatial unit has equal constant value and it causes the same influence value about factors of human longevity in the study region. Ecological variables associated with longevity are distributed continuously in space and may have spatial variability depending on the conditions of the terrain or distance. OLS regression model has risk of under-estimate when the variables make spatial autocorrelation (Griffith, 1996). When considering the first law of geography that everything is interrelated, but near objects are more related than distant objects (Tobler, 1974). Therefore, the method considering spatial characteristics required when analyzing the ecological factors of human longevity. GWR model derived different regression model about each spatial location and can be reflect the spatial variation of factors applying geographical weighted function. In this paper we performed a comparative analysis between OLS and GWR model about ecological factors of human longevity that proposed in previous studies.

STUDY AREA

The study area coverd Gang-neung census administrative in Gang-won province, KOREA. This area lies between 37°27'N and 37°54'N and 129°04'E and 128°35'E(Fig. 1). The total terrestrial area of Gang-neung census administrative is 1,039.82㎢.

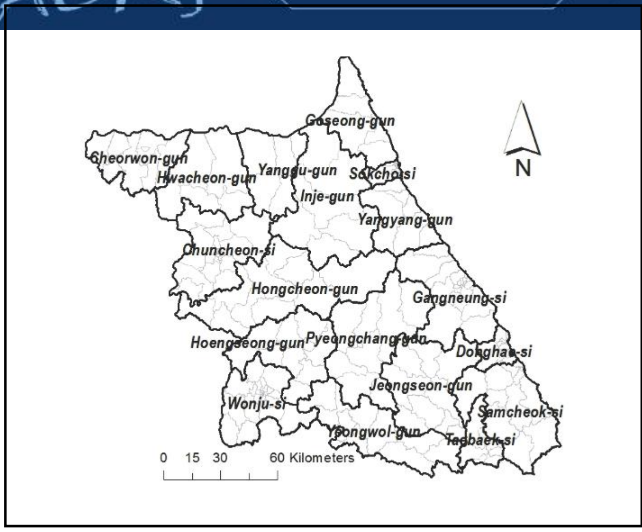

Fig. 1 Study area

The temporal range limited 2010 because the focus of paper was just methodological comparison. The dependent and independent variables used in the analysis of existing case studies. Land cover data using the data from the Ministry of Environment of KOREA and meteorological data were used for 2010 automatic weather system (AWS). Meteorological point data was reprocessed spatial statistical method, kriging.

METHODLOGY

The longevity index used in this study, equation 1(Kim, 2005).

$$\text{Longevity Index} = \frac{Over\ 80\ population}{Over\ 65\ population} \times 100(\%) \qquad (1)$$

GWR model applied the regression coefficient as a function of location, none constant value. Therefore, it can be identified the spatial variability of effects of the independent variables each location. The GWR model process is Eq. 2.

$$y_i = \beta_{oi}(u) + \beta_{1i}(u)x_{1i} + \quad {}_{2i}(u)x_{2i} + \cdots.. + \beta_{mi}(u)x_{mi} \quad (2)$$

Equation 2 is similar with general regression model equation. However, GWR model is weighted least method included function of distance such as u. Different weights applied according to the location in the study area and calculate different estimates such as Eq. 3.

$$\hat{\beta}(u) = [X^T W(u) X]^{-1} X^T W(u) Y \qquad (3)$$

GWR model fit assessed carried out generally mixed AICc (corrected Akaike Information Criterion : AICc).

In this paper, two major analyzes were performed for accomplish the purpose of research. First, the major statistical indicators were compared between OLS and GWR model for identify the effective model in the field of research on the longevity factors. Another purpose of this research is visualization to the spatial variability of regression coefficient of independent variables from GWR model.

RESULT

Table. 1 provides a comparison of the OLS and GWR models. Variance Inflation Factor (VIF) shows none multi-collinearity among the independent variables. Only one of the regression coefficients value was calculated by OLS

	OLS	GWR			VIF[1)]
		Min	Median	Max	
constant	11.082*	20.539*	21.721*	22.048*	
Mean temperature	0.496	0.141	0.569	0.984	2.146
Rain fall	0.002	-0.005**	0.0001**	0.003**	1.591
Elevation	-0.005**	-0.007	-0.006	-0.003	3.129
Forest rat.	3.178**	2.646	4.764	5.791	1.112
Agriculture rat.	0.290	-0.3	1.252	2.679	1.906
R^2adj	0.15	0.2			
AICc	953.699	947.39			
σ	2.994	2.89			
Residual sum of square	1631.474	1476.727			

*** p < 0.05, ** p<0.01**

Table. 1 Results of OLS & GWR

model. On the other hand, The GWR model could be calculated different regression coefficient values for each local study area. Major statistics on the results from each of the regression model shows that the GWR model was calculated more effective model estimation results than OLS model, in the R^2 and AICc. R^2 and AICc. The confidence of the estimation results from regression model can be determined through the R^2 and AICc.

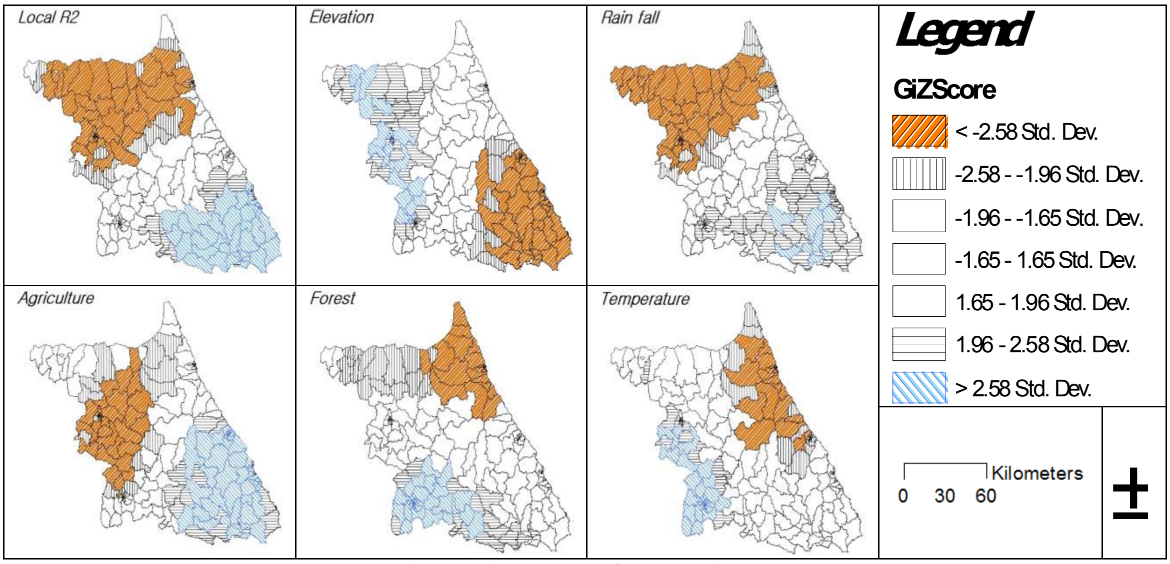

Figure.2 cluster map of GWR estimates

Figure. 2 is the visualization of $\text{Getis} - \text{Ord}\,Gi^*$ analysis results to identifying the spatial variability of local R^2 and estimated regression coefficient from GWR model. The clusters can be represents spatial relationships between variables.

DISCUSSION

In this paper we performed a comparative analysis between OLS and GWR model about ecological factors of human longevity that proposed in previous studies. Major statistics on the results from each of the regression model shows that the GWR model was calculated more effective model estimation results than OLS model, in the R^2

and AICc. Therefore, GWR model can be explained the spatial variability of specific environment factors.

ACKNOWLEDGEMENT

This work was researched by **the supporting project to educate GIS experts**

REFERENCES

Kim, H., 2007, Comparison of Three Dasymetric methods for population density mapping, The Geographical Journal of Korea, 41(4), 411-419.

Griffith, Daniel A, 1996. Introduction: The need for spatial statistics. *Practical Handbook of Spatial Statistics*. Boca Ration, FL: CRS Press, pp. 1-15.

Pesa, G.M., Tolu, F., Poulain, M., Errigo, A., Masala, S., Pietrobelli, A., Battistini, N.C., and Maioli, N.C., 2011, Lifestyle and nutrition related to male longevity in Sardinia: An ecological study. Nutrition, Metabolism & Cardiovascular Disease, Vo. 10, pp. 1-8.

Tobler, W., 1970, A Computer Movie Simulating Urban Growth in The Detroit Region, Economic Geography, Vol. 46, No. 2, pp. 234-240.