Eco. Env. & Cons. 26 (3) : 2020; pp. (1282-1289)
Copyright@ EM International
ISSN 0971–765X

Spatial relationship modeling for urban environmental factors analysis to population density change

Nutchanat Buasri[1], Patiwat Littidej [1], Naruset Prasertsri[1] and Benjamabhorn Pumhirunroj[2]

[1] Department of Geo-informatics, Faculty of Informatics, Mahasarakham University, Thailand 44150
[1] Research unit of Geo-informatics for Local Development
[2]Program in Animal Science, Faculty of Agricultural Technology, Sakon Nakhon Rajabhat University, Sakon Nakhon 47000

(Received 29 April, 2020; Accepted 4 June, 2020)

ABSTRACT

Changes in land use directly affect population changes. Population forecasts in grid units are critical to implementing the results of urban development planning. The study created a whole model of the synthesis of physical autonomous variables so that they could analyze the relationship to population changes attributed to the X_1 (the vegetation conservation index that calculation from difference percentage of summation of NDVI cell), X_2 (the miscellaneous conservation index that calculation from the difference percentage of miscellaneous area), X_3 (the difference percentage of built-up area), X_4 (the volume of road network of block), X_5 (the normalized distance index from center point of block) and Y = the difference value between the population density of block each. Results from the OLS model analysis showed that independent variables X_4 and X_5 correlated positively and negatively to significant population changes per grid unit (1 sq.km.).

Key words: *Spatial regression model, Population density change, Nakhon Ratchasima of Thailand*

Introduction

Migration means that the relocation or relocation of the migration is associated with the normal relocation of the place, (Haughey, 2005) is called a government unit where migrants from the same settlement and the place where the migrants are relocated to their destination. Toth (2003) studies show that the factors influencing migration, factors that are driving factors that are tensile. The big factors that are possible both push and attract are populations of both size and density, economy, tradition, culture, society and political security, utilities, transportation and the environment, and so on. In another research Cao, 2000 and his colleagues (2000) in order to study the changes in land use have a direct impact on population changes, but there are different factors in the area, which is mostly a factor in road networking, the potential of miscellaneous areas, the loss of soil cover such as vegetation and forests (Geist, 2002). The use of such factors in modeling to anticipate population changes is necessary to be appropriately established and in line with the actual relocation of the study area (Littidej, 2019). (Liu, 2003) conducted a thorough -statistical analysis of the urbanization process at the national and the provincial level. Spatial modeling to predict the population is used by multiple models. According to Littidej, (2019) findings, there are several popular environmental simulation spatial models used as models that predict data over a period of up to 10 years.

Predicting land-use and land-cover change using mathematic models and geo-information systems is one of the popular methods as it can visualize outcomes in spatial information patterns, which are not limited to the quantitative information available in other conventional methods. Land-use information is fundamental to analyzing the changes of land coverage by analyzing the various spatial software (Pontius, 2000).

This work applies a wide range of spatial models, such as Geomod2 (Eastman, 2006), CA, SLEUTH Urban Growth (Jantz, 2003), Lucas (Conclelis, 1997), and LTM (Feoli, 2004). Each model has its own characteristics and processes. However, they function in the same capacities, which are used to model the change of land- use and land-cover. Additionally, each technique is applicable to different geo-information software. CLUE-S can be used together with ArcGIS, QGIS, and IDRISI.

Many of the models mentioned above are predictable models at both grid level and quantitative data, but are complex in configuring independent variables, as well as assigning subarea units to models that, compared to the result, have a precision value not unlike ordinary least square (OLS) linear models. In this study, OLS models were used to be used as symmetrical models and were more suitable than Geographic weighted regression (GWR) models that required complex subarea units. The structure of this study, in order to create an independent, spatial variable model associated with the physicality of the city, is associated with population changes and analyzes factors influencing population changes.

Materials and Methods

Data

The creation of a land utilization data layer in Muaeng district of Nakhon Ratchasima is an extraction of high-resolution, scaled image data. 1:10000 is divided into 5 main categories: urban and built-up area, forest agriculture, water body, and miscellaneous summarized and compared the land use statistics in 2011 and 2019, respectively showed in Table 1 and Figures 1 and 2. The Urban and built-up area (U) has increased in a single category and Agriculture land (A) Forest land (F) Water body (W) Miscellaneous land (M) has steadily declined due to Nakhon Ratchasima province being the center of northeast transportation and the construction of large-scale projects including Motorway-6 and High speed train, resulting in the expansion of buildings and the expansion of population and relocation to work in the area. The population density data obtained from (Department of Provincial Administration, 2019) DOPA is processed in conjunction with Landscan satellite imagery data, showing the population density per square kilometer. As shown in

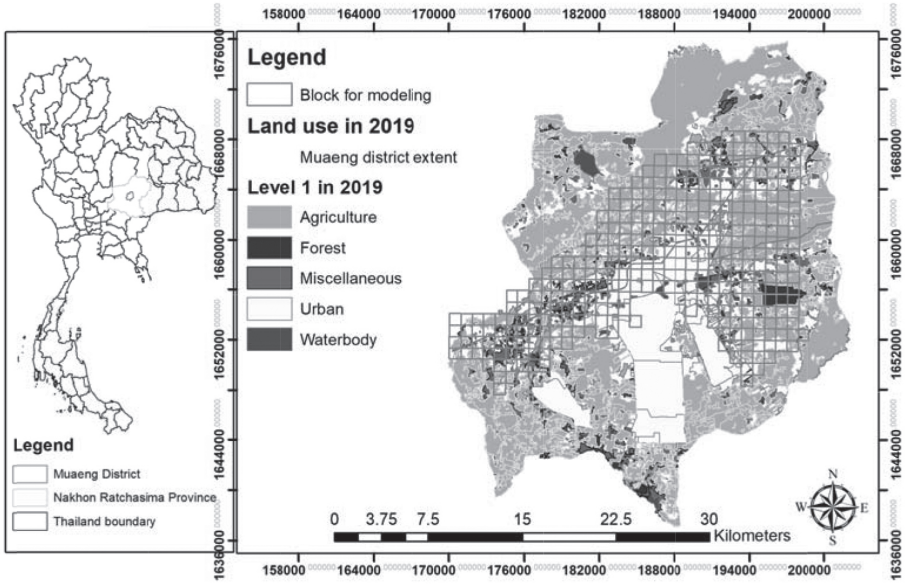

Fig. 1. Map of land use and land cover (LULC) in year 2019 of study area (Muaeng district of Nakhon Ratchasima province).

Table 1. The allocation for land-use categories in 2011 and 2019.

Land-use Types	Year (2011)		Year (2019)	
	sq. km	%	sq. km	%
Urban and built-up area (U)	202.68	31.40	225.88	34.97
Agriculture land (A)	385.01	59.61	370.06	57.30
Forest land (F)	7.12	1.10	6.52	1.01
Water body (W)	7.45	1.15	6.91	1.07
Miscellaneous land (M)	43.58	6.74	36.47	5.65
Total	645.84	100.00	645.84	100.00

Figs. 3 and 4. The population density in 2011 ranges from 7-13,383 persons/1 sq.km and in 2019 is 6-11,898 persons/1 sq.km.

Methods

In this research, the OLS model was used to determine the relationship of factors to population density changes. The adoption of OLS models to analyze spatial relationships to population changes is essential to create an appropriate spatial unit to determine the value of independent variables and dependent variable accordingly to that extent. In this study, the size grid was created. 1 km. x1 km. up to designate as the same unit as population density. How accurate is the OLS model to determine the appropriate size and shape of the spatial unit, according to (Littidej, 2019) and (Prasertsri, 2020).

The factors used in the modeling in this study were conducted by the relevant authorities in Nakhon Ratchasima province, based on the population changes in the last 10 years. The factors used in the modeling are divided into variables based on Y (the difference value between the population density), independent variables that focus on creating models based on spatial data derived from data collected from local authorities that indicate that these five factors affect population changes: X_1 (the vegetation conservation index that calculation from dif-

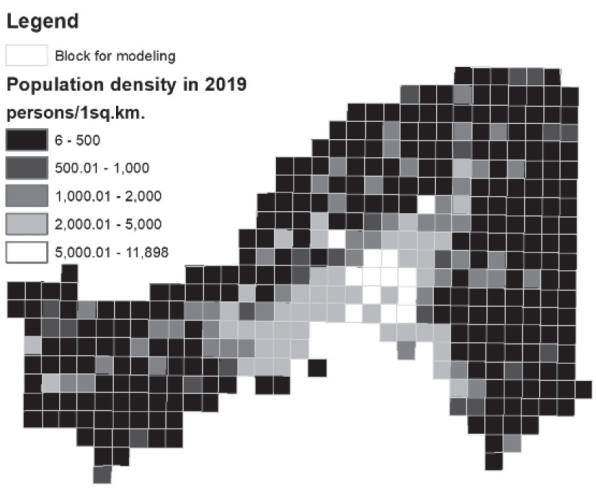

Fig. 3. Map of population density in 2019.

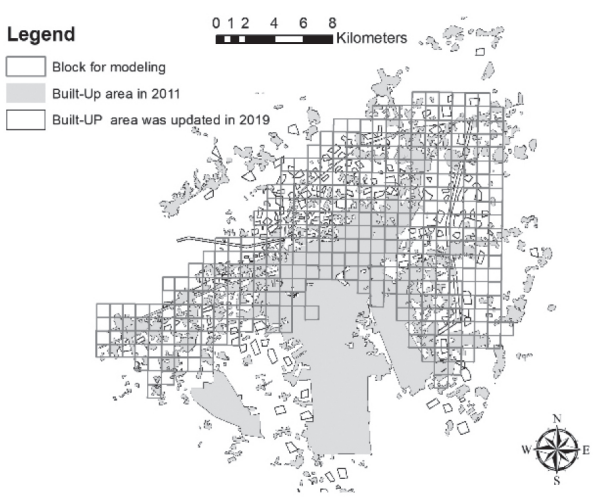

Fig 2. Map of built-up polygon in 2011 and 2019.

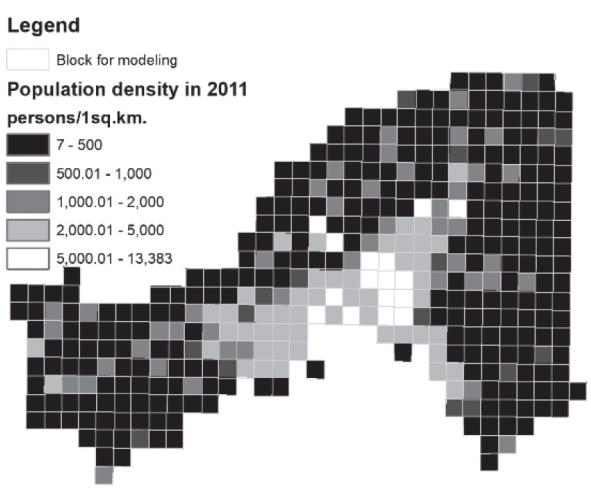

Fig 4. Map of population density in 2011.

ference percentage of summation of *NDVI*), X_2 the miscellaneous conservation index that calculation from the difference percentage of miscellaneous area, X_3 the difference percentage of built-up area, X_4 the volume of road network of block and X_5 the normalized distance index from center point of block, respectively, and display modeling to calculate the value of variables as equations 1 through 7.

Spatial regression model in the study using ordinary least square (OLS) depicted as equation (1).

$$Y = a_0 + b_1X_1 \pm ..... \pm b_nX_n \pm \epsilon \quad .. (1)$$

Where:
Y = dependent variable
a_0 = intercept
b_1 = coefficient of X_1
X_1 = the independent variable of parameter *1*
b_n = coefficient of X_n
X_n = the independent variable of parameter each *n*
ϵ = error of the model

Y (dependent variable) of spatial relationship analysis for population density change was showed in equation (2).

$$Y = P_{2019(i)} - P_{2011(i)} \quad .. (2)$$

Where:
Y = the difference value between the population density of block each (*i*) in year 2019 with 2011 (unit: persons/1 sq.km.)
$P_{2011(i)}$ is the population density (persons/1sq.km.) in year 2019 in block each (*i*)
$P_{2011(i)}$ is the population density (persons/1sq.km.) in year 2011 in block each (*i*)

$$X_1 = \left(1 - \left|\left(\frac{\sum_{i=1}^{n} NDVI_{2019(i)}}{n_i}\right) - \left(\frac{\sum_{j=1}^{m} NDVI_{2011(j)}}{n_j}\right)\right|\right) x100 \quad .. (3)$$

Where:
X_1 = the vegetation conservation index that calculation from difference percentage of summation of *NDVI* cell each (*i*) in block of year 2019 and 2011 (unit: 0-1)
NDVI is the difference normalize vegetation index *(NIR-RED/(NIR+RED)*
n_i is a number of *NDVI* that value >0.2 of cell each *(i)* in block of year 2019
n_j is a number of *NDVI* that value >0.2 of cell each *(i) in* block of year 2011

$$X_2 = 100 - \left[\frac{(M_{2019} - M_{2011})x100}{10^6}\right] \quad .. (4)$$

Where:
X_2 = the miscellaneous conservation index that calculation from the difference percentage of miscellaneous area in year 2019 with 2011 (unit: 0-100 percent)
M_{2019} = the miscellaneous area in year 2019 (sq.m.)
M_{2011} = the miscellaneous area in year 2011(sq.m.)
10^6 = the area in unit (sq.m.) of 1 sq.km.

$$X_3 = \left[\frac{(BU_{2019} - BU_{2011})x100}{10^6}\right] \quad .. (5)$$

Where:
X_3 = the difference percentage of built-up area in year 2019 with 2011 (unit: 0-100 percent)
BU_{2019} = the built-up area in year 2019 (sq.m.)
BU_{2011} = the built-up area in year 2011 (sq.m.)
10^6 = the area in unit (sq.m.) of 1 sq.km.

$$X_4 = \Sigma_{R=1}^{r} A_{R(i)} L_{R(i)} \quad .. (6)$$

Where:
X_4 = the volume of road network of block in year 2019 with 2011 (unit: m^3)
$A_{R(i)}$ = the area of road link number *R* in each block (*i*) in year 2019 (unit: sq.m.)
$L_{R(i)}$ = the length of road link number *R* in each block (*i*) in year 2019 (unit: meters)

$$X_5 = \frac{DC_{ij}}{DC_{max}} \quad .. (7)$$

Where:
X_5 and DC_{ij} = the normalized distance index from center point of block each (i) to nearest road (j) (meters)
DC_{max} = the maximum distance from center point of block each (i) (meters)

The calculation of the grid value of each variable is calculated using the raster calculator function, which is a tool in Arcmap 10.X. The data used to calculate the model's calculation is the NIR and RED wave range of the Landsat 8 OLI satellite image. Built-up land use data is used to calculate the basis of population, with the assumption that the population will outnumber other types of land use areas

Results and Discussion

Dispersion of the variables

The effect of the distribution of the *Y* (dependent variable) shown in Figure 5 shows the change in population density in the central area and the com-

mercial district and the major economic areas will change more than others. In addition, the urban areas of the area have also decreased due to the change in relocation to most of the city's residents.

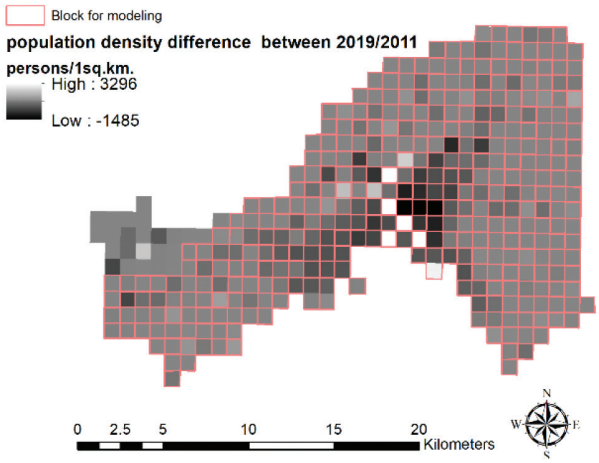

Fig. 5. Y (dependent variable)

The X_1 (the vegetation conservation index) shown in Figure 6 indicates that the preserve of vegetation in the area is much higher, but if the lower the index indicates that the plant's retention rate is relatively low, the area of the northern city is relatively low due to the rapid change in land cover, which makes the population change in the area quite high.

The miscellaneous conservation index (X_2) shown in Figure 7 represents the persistence of miscellaneous areas, the lower the value of the index, the lower and the rate of miscellaneous areas to the higher the utilization of other types of land. The area has changed a lot in the area with X_1, but there

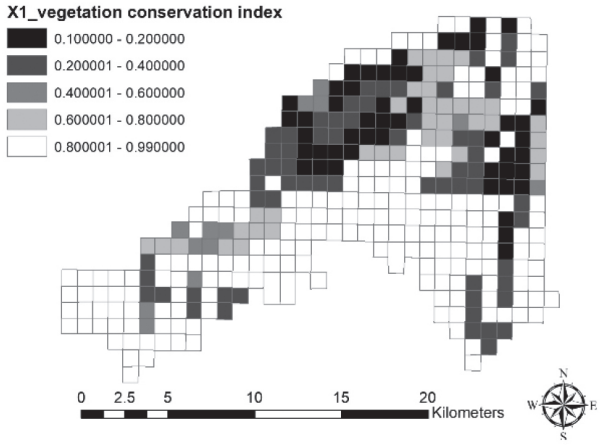

Fig. 6. Map of variable (X_1)

will be more radical changes because this type of land cover is already being prepared for development.

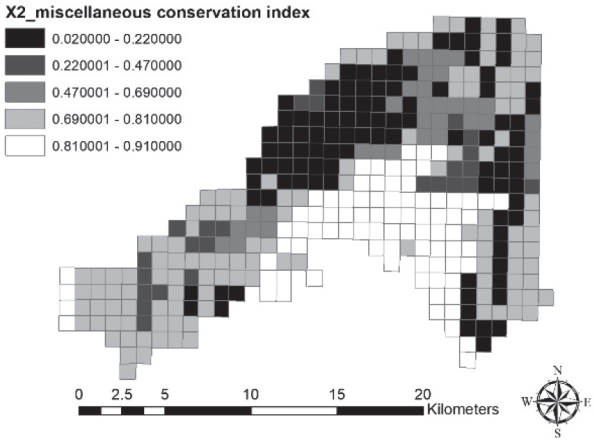

Fig. 7. Map of variable (X_2)

X_3 factor (the difference percentage of built-up area), an index that shows the area, the rate of changes in construction area. The index is very high, showing a high increase in buildings and other buildings such as shopping malls. High-indexed areas are scattered along the northbound lanes and new east intersections and are shown as white ranges in Figure 8.

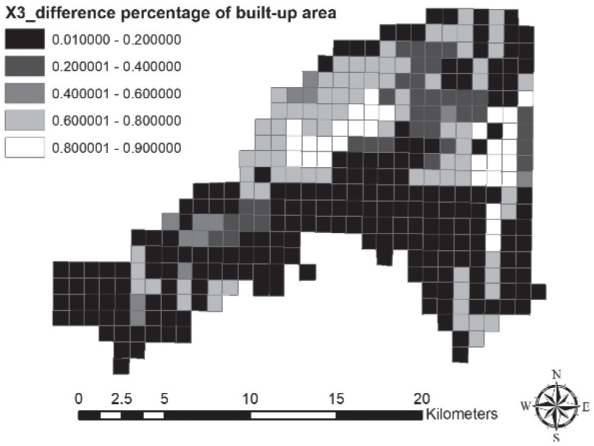

Fig. 8. Map of variable (X_3)

The volume of road network shows the volume of road networks in sub-areas of 1 km.x1 km. The volume of roads is greater due to the large population of the road as it is moved to business and other activities where high indices are scattered along the northbound road, in addition to the longer road areas, and the area is developed as a building along

the side of the road about 1 km away. As shown, the distribution of the index in Figure 9.

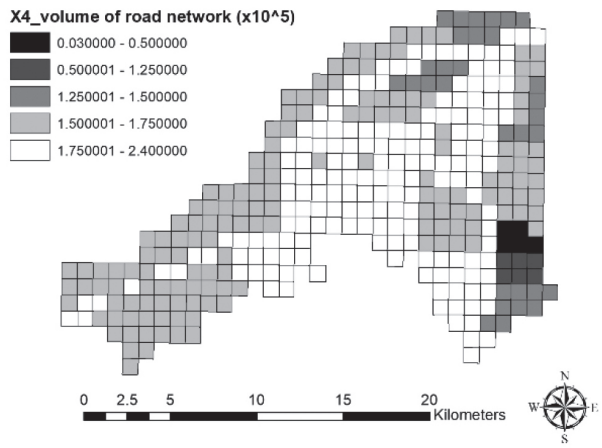

Fig. 9. Map of variable (X_4)

The Normalized index from center point of block displays the value of the standard distance from the center of each grid to the intersection of the nearest road line, an index that indicates the influence of population settlement gravity. The index values are consistent and similar to the X_5 variable, but are different in the central area of the city, displayed as low values, since there is no link of the large road network to most grids in that area and the index is displayed in Figure 10.

Spatial model for population density analysis

The OLS models used in this study were used by Arcmap 10.X in function of ordinary least square (OLS) to display the results as Table 2. The results of the model showed that the five independent variables that were modeled on both positive and negative relationships with population changes. The X_1 to X_3 factor was found to have a negative correlation to the population, which when the decrease in the number of vegetation, miscellaneous and built-up areas contributed to the increase in the number of populations, but the results of the model showed an insignificant correlation of population changes, showing a significant level of *-6.26*, -4.64 and *–7.28* n/s and coefficients equal to -5.821x10^{15}, 1.087x10^{15} and -6.233x10^{15,} respectively.

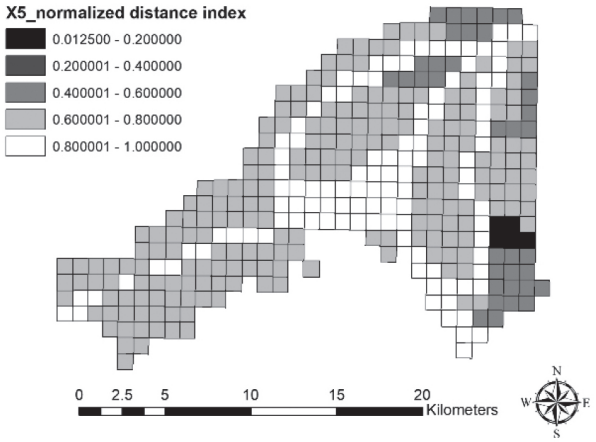

Fig. 10. Map of variable (X_5)

The OLS models reject the adoption of X_1 to X_3 variables, as well as the intercept values, but observations from this analysis value are useful to define independent variables in similar modeling. The X_4 (the volume of road network of block) and X_5 (the normalized distance index from center point of block) factors are two factors that are characterized by similar population changes, namely, with the relocation characteristics of the population in this area, they choose to move to an area near the main

Table 2. Summary of the results of the OLS model.

Variables	OLS coefficients (â)	p-value[a]
Intercept	5.821x10^{15}	9.48 n/s
X_1 the vegetation conservation index that calculation from difference percentage of summation of NDVI cell	-1.087x10^{15}	-4.64 n/s
X_2 the miscellaneous conservation index that calculation from the difference percentage of miscellaneous area	-5.146x10^{15}	-6.26 n/s
X_3 the difference percentage of built-up area	-6.233x10^{15}	-7.28 n/s
X_4 the volume of road network of block	3392	0.01***
X_5 the normalized distance index from center point of block	-7936	0.00***
N	366	
Adjusted R^2	0.724	

***significant at 1% level

n/s = not significant, [a] Results of the Monte Carlo test for spatial non-stationary

road and build houses along the road.

The accuracy level of the model is R^2=0.724, which is not very high compared to spatial statistical models, which mostly require that the model should have an R^2 value greater than 0.8 to be able to predict the effect of the variable accordingly (Patiwat, 2019). Constructing a road volume index is used as a model for population changes studies, predicting trends of population changes.

The greater the volume of roads, the greater the population of the area, the less likely the road area is to be the area of the road, which is much greater than the area near the road. Models can predict the effect of population close to real values from Lanscan data, which models predict well in areas where there are 2 factors with uniform index fragmentation.

The model given to this study is $Y = 3392(X_4)$-$7936(X_5)$, which can be used to predict other areas with similar land utilization changes, such as provinces influenced by the construction of large roads such as motorways or dual rail trains, as well as high-speed trains, for example, mainly provinces located in the northeastern region of Thailand where all three indexes are able to study the relationship of population changes, but they must be used to model in smaller and more appropriate areas.

Adaptation in the study area

The implementation of OLS models accurately predicts the likely areas of population changes and shows the discrepancies shown in Figs. 11 and 12. Using SR (standard residual) to help make decisions allowing the model to be more predictable from image 11 is a display of SR, based on a coefficient of 5 factors (X_1 to X_5).

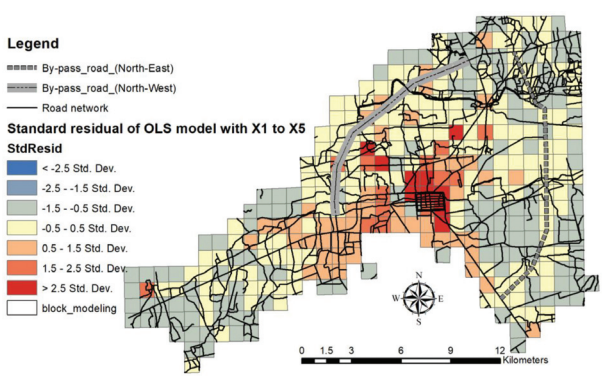

Fig. 11. Map of standard residual of OLS model using independent variables of X_1 to X_5.

The SR value from the model shows the range between (-2.5 to>2.5). The appropriate values to be used are -0.5 to 0.5 when the grids that represent these are areas where population changes can be predicted more accurately and less misleading than other areas.

Grids in areas with city bypass roads cut through both roads (By-pass road NE direction) and (By-pass road NW direction) are areas where models have low tolerances due to the consistent distribution of X_4 and X_5 variable values. Although the model only applies to the X_4 and X_5 variables, the SR display model displays a lower error than other zone areas as shown in Fig. 12.

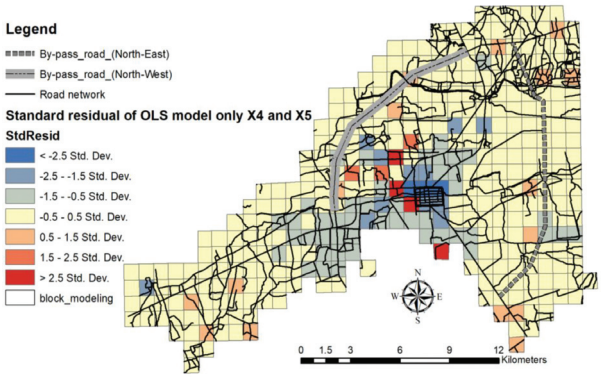

Fig. 12. Map of standard residual of OLS model using independent variables of X_1 to X_4.

These areas have changed land use from miscellaneous areas and agriculture to buildings, resulting in a number of population changes in the range of 1000 persons/1sq.km. The area within a range of up to 2 kilometers along the city's bypass road also has the potential to predict better models than other zones, namely that land use changes contribute greatly to attracting an increased population or a reduction in numbers.

The application of the model predicts the population in other zone areas, it is necessary to study other independent variables in addition to this study, because that each zone has factors that contribute to different population changes, such as government centers, shopping malls, sports centers, transportation centers that are critical to change.

However, in this study, independent variable factors can predict changes in population density in areas where roads are constructed at a satisfactory level and are likely to be used as a model for education in other provincial areas.

Conclusion

The OLS models developed from the use of the prototype area of Nakhon Ratchasima province found that based on the SR test, there are grid areas that show acceptable tolerances scattered mostly along the road. In a study of (Littidej, 2019) and (Prasertsri, 2020) a study on land utilization predictions in the future years of the same area, it was found that local statistic approach (GWR) modeling can provide a higher R^2 value than the OLS method. The previous research (Littidej, 2019) and (Prasertsri, 2020) explained that the study area must be divided into sub-units in order to achieve a change in the orientation of independent variables, so in this study only the OLS model was used because it was analyzed and modeled from variables within the grid. This study did not compare the results with the GWR model, but the tendency of selecting variables that influence population changes was analyzed from this OLS model.

Acknowledgements

This Research was Financially Supported by Faculty of Informatics, Mahasarakham University Grant Year 2019.

Conflict of Interest

All the authors declare no conflict of interest.

References

Cao, G.Y. 2000. *The Future Population of China: Prospects to 2045 by Place of Residence and by Level of Education*. Laxenburg, Austria: International Institute for Applied Systems Analysis. Report No. IR-00-026.

Conclelis, H. 1997. From cellular automata models to urban models: New principles for model development and implementation. *Environ. Plan.* 24 : 165–174.

Department of Provincial Administration (DOPA), 2019. *Final report*. Https://multi.dopa.go.th/tspd/tpad/policy/plan[3]/topic[4]

Eastman, J.R. 2006. *IDRISI Andes Guide to GIS and Image Processing*. Clark Labs, Clark University: Worcester, MA, USA.

Feoli, E. 2004. *Migration towards the Cities: Measuring the Effects of Urban Expansion in Rural-Urban Interface by GIS and RS Technology*. Department of Biology University of Trieste: Trieste, Italy,

Geist, H.J. and Lambin, EF. 2002. Proximate causes and underlying driving forces of tropical deforestation. *Bioscience*. 52(2) : 143–150.

Haughey, Richard, M. 2005. *Higher-Density Development: Myth and Fact*. Washington, D.C.: ULI–the Urban Land Institute.

Jantz, C.A., Goetz, S.J. and Shelley, M.K. 2003. Using the SLEUTH urban growth model to simulate the impacts of future policy scenarios on urban land use in the Baltimore-Washington metropolitan area. *Environ. Plan. B Plan. Des.* 31 : 251-271.

Littidej, P. and Buasri, N. 2019. Built-Up Growth Impacts on Digital Elevation Model and Flood Risk Susceptibility Prediction in Muaeng District, Nakhon Ratchasima (Thailand). *Water.* 11(7) : 1496.

Liu, S., Li, X. and Zhang, M. 2003. Scenario Analysis on Urbanization and Rural-Urban Migration in China. Laxenburg, Austria: International Institute for Applied Systems Analysis. IIASA Land Use Change Report IR-03-036.

Prasertsri, N. and Littidej, P. 2020. Spatial Environmental Modeling for Wildfire Progression Accelerating Extent Analysis Using Geo-Informatics. *Polish Journal of Environmental Studies*. 29 (5) : 3249-3261.

Pontius, R.G., Jr. 2000. Quantification error versus location error in comparison of categorical maps. *Photogramm. Eng. Remote Sens.* 66 : 1011–1016.

Toth, F.L., Cao, G.Y. and Hizsnyik, E. 2003. Regional Population Projections for China (IIASA Land Use Change Report, IR-03-042). Laxenburg, Austria: *International Institute for Applied Systems Analysis*.