REMOTE SENSING AND MULTIVARIATE LOGISTIC REGRESSION MODEL FOR THE ESTIMATION OF URBAN EXPANSION (CASE OF DARKHAN CITY, MONGOLIA)

P. Myagmartseren¹, D. Ganpurev¹, I. Myagmarjav^{2*}, G. Byambakhuu¹, G. Dabuxile^{3,1}

¹Dept. of Geography, National University of Mongolia, Ulaanbaatar 14201, Mongolia - (myagmartseren, ganpurev,

byambakhuu)@num.edu.mn

²Dept. of Land Management, Mongolian University of Life Sciences, Ulaanbaatar 17024, Mongolia – myagmarjav@muls.edu.mn ³College of Geographical Sciences, Inner Mongolia Normal University, Hohhot 010022, China - 391451163@qq.com

Commission III, WG III/7

KEYWORDS: Urban modelling, Urban growth prediction, Spatial analysis, Land Use and Land Cover Change LUCC

ABSTRACT:

Urban expansion and land use and land cover change (LUCC) studies are a key knowledge of efficient local governance and urban planning and a lot contributing to the future sustainable development of the city. The main goal of the paper is to model a future urban spatial expansion by 2029 and 2039 of Darkhan city using Landsat TM satellite imagery (land use and cover change map of 1999, 2009, and 2019) and multivariate logistic regression model. Clark Lab's (Clark University) IDRISI & TerrSet software applied for the urban expansion prediction and the correlation between expansion and driving factors. On account of multivariate logistics regression modelling, eight physical factors influencing urban expansion identified to predict urban expansion based on USGS Landsat TM imageries (Landsat Multispectral Scanner with 60 m resolution). The regression statistic accounted for the probability of future urban expansion was positive. Overall, the LUCC has estimated the transition of natural cover to the impervious surface in Darkhan city. Our result estimates an increase in the built-up area and slum area during the period 1999-2009 and 2009-2019, represents LUCC was characterized by an external transformation from natural to urban area. According to the future urban growth prediction, the urban area would be significantly spread into the open space and natural vegetation area. The main findings stated here are that Darkhan city is expanding in an unsystematic way, even though the urban growth has not been analysed in detail and has a bad case of urban unregulated sprawl.

1. INTRODUCTION

1.1 Study phenomenon and area

Urbanization is a relatively new phenomenon for modern Mongolia. This process began in the middle of the 20th century when just over 20% of the population lived in urban areas, and urbanization has intensified since the 1950s.

In the early 1960s, for modern industrial development, the Darkhan city initially with inhabitants of 40,000 about was established in the north of Mongolia (Figure 1) along the Trans-Siberian railroad, and over the last 60 years, the young city growing till to the second-largest urban area with an important role for the socio-economic development of Mongolia after capital city Ulaanbaatar. The number of large industries and services had been established during the socialist time and had strictly and straight forward managed by the central government to control the growth according to town planning. However, the democratic revolution in the 1990's allowed numerous domestic migrants to arrive from the neighbouring rural provinces to the city. The aftermath of domestic population migration increased inhabitants to 96,076 in 2018, which was unexpected by government town planning to sustainably develop the city. Therefore, the Darkhan city is chosen as a case area and this study was accomplished for the purpose to estimate the future spatial change of the urban extent.

1.2 Literature review

Logistic regression analysis is the most common method to measure land use and land cover change (Dietzel and Clarke, 2006), and future urban expansion trends (Gillham, 2002). Although common to international scholars and researchers are not a full-scale study in Mongolia that estimates urban sprawl using a logistic regression method. There are considerable studies carried out and published by scientists about a logistic regression analysis for urban growth study and we considered for establishing research design and method below listed kinds of literature: Firman (1997), Hu and Lo (2007), Eyoh et al., (2012), Helbich and Leitner (2010), Alsharif and Pradhan (2014). Also, Myagmartseren and Myagmarjav (2017) used population growth land supply-demand and land suitability evaluation for determining the future urban growth of Darkhan City.



Figure 1. Map over the study area. (a) Location of Darkhan; (b) Map with a satellite image.

^{*} Corresponding author

2. MATERIALS AND METHOD

This research was based on data from Landsat TM-5 satellite imageries of Darkhan city in terms of 1999, 2009 and 2019. The urban expansion model was defined using the Logistics Regression Model of the Geographic information system and remote sensing software IDRISI & TerrSet which is developed by Clark Labs at Clark University (Eastman, 2016).

2.1 Remote sensing dataset and image processing

Remote sensing provides admirable tools for handling all issues related to land resource management. One of the significant features of the remote sensing technology is capable of handling and combining different types of data very efficiently (Myagmartseren et al., 2018). For this study, multi-resolution remote sensing data of USGS Landsat imageries (a series of earth observation satellites) for the period 1999-2019 has been used. Landsat TM (Thematic Mapper with multispectral scanner image with bandwidth 0,5-1,1 μ m and 60 m horizontal resolution) and Landsat 7 ETM + (Landsat Enhanced Thematic Mapper Plus with 30 m horizontal resolution and bandwidth 0,45-2,35 μ m) imageries acquired in September 2019 and downloaded from http://earthexplorer.usgs.gov/ (Table 1).

#	Dataset	Date	Reference coordinate system
1	Landsat TM 5	8/10/1999	WGS84 UTM Zone 48N
2	Landsat ETM+7	8/20/2009	WGS84 UTM Zone 48N
3	Landsat OLI 8	8/21/2019	WGS84 UTM Zone 48N

Table 1. List of satellite data

For accurate satellite image classification, the maximum likelihood classification technique (in our case supervised classification method based on reference area polygons as a training area) was used for land use and land cover main class delineation (Al-shalabi et al., 2012). At the end of the classification, land use and land covers have been determined (Table 2). Landsat satellite images supervised classification was performed to extract the urban area layer of 1999, 2009, and 2019. A total of 8 classes were generated; e.g., residential, vegetation, slum, industrial, open space and water. Other subsequent classes such as forest (except intra-urban area forest classes counted as green construction or gardens and merged into the open space class) were not further used (cf. Table 2). After the supervised classification and several image manipulations, the city's urban area (in our study context named urban footprint) was separated from other categories of satellite images. Using the cadastral map attribute data with parcel land-use identity purpose, we segregated and re-classed the urban area into residential, industrial and mining sub-classes. By the reason, 30 m resolution of Landsat dataset has difficulty for detailed classification of urban land uses residential, slum or industrial, etc (Myagmartseren et al., 2018).

#	Land use and land cover classes	Explanation
1	Residential	A mid- and high-rise building and construction area for housing, communal utilities and services, cultural and enlightenment, commercial, office and other facilities.
2	Slum	"Ger district" (circular nomad's tent- yurts detached to the land lot is a type of

		slum settlement district in Mongolia) low density and single houses, dwellings occupied area
3	Vegetation	A grassland with natural vegetation cover for grazing animals and livestock.
4	Water	A body of water such as rivers, streams, springs, lakes, ponds, wetlands
5	Infrastructure	Physical and organizational structures and facilities (road, railway, their stripes, dams, etc.)
6	Mining	Mine production facilities, mine operation construction, and mining areas
7	Industrial	Light food factory and heavy industries construction and areas
8	Open space	The land is not provided for possession and uses, vacant lands and in urban footprint/ urban envelop

Table 2. Classified images of Land use *Source: Myagmartseren et al., 2018*

Have to note the residential land areas were further divided by two sub-categories: residential (area with up to two stories: middle and high-rise buildings) and sprawl residential/slum area: low-rise area with detached "ger" tents surrounded by fences or with handmade dwellings (Myagmartseren et al., 2018). The reason for dividing the residential class into sub-categories was to quantify the urban extent and further in the paper named "residential" class means a mid- and high-rise building and construction area for housing. Each land use class temporal and spatial changes have been calculated after the land use classification.

2.2 Detecting the land use changes and future prediction

Land use and land cover changes are determined by measuring the different time areal differences between the same geographical location land patches. Land use change for each land use/cover class i was calculated as Zhang et al., 2002:

$$RCi = \left((S + i/Si +)^{1/T2 - T1} - 1 \right) \times 100\%$$
(1)

where RCi- is the annual rate of change;

```
S- is the area (in ha);
```

```
i- is the years;
```

S+i and Si - are the areas of a land use/cover class at years;

T1- (initial time) and T2- (next time).

Multivariate logistic regression analysis was used to map the probability of urbanization. Based on the urbanization probability map produced by the model, knowing where urban expansion will happen, we can demonstrate future urban distribution patterns based on the current change of urban land use.

Logistic regression is a commonly accepted method to predict the urban spatial change process when using a raster binary map of the dependent variable (main driving factors of urban land use change) of urban LUCC (Alsharif and Pradhan, 2014). The urban land conversion probability of each cell in the raster map was represented on the following logistic regression equation by Alsharif and Pradhan (2014):

$$f(z) = 1 / (1 + e^{-z})$$

$$P(Y = 1 | X_1, X_2, \dots, X_k) = 1 / (1 + e^{-(\alpha + \sum \beta_i X_i)})$$
(2)

where Xi is an independent variable representing driving factors of the urban process; α is the coefficient of the model formula;

P(Y=1|X1, X2,...,Xk) is the probability of the dependent variable Y being 1 given (X1, X2,...,Xk), i.e., the probability of a cell being changed to urban land; and β i is the coefficient of variable Xi. The regression coefficient β i reflects the function of independent explanatory variables (Alsharif and Pradhan, 2014).

The literature review indicates several common explanatory variables used by international and domestic scholars to predict urbanization by the different methods. Based on our chosen previous urban growth prediction and modelling study literature (cf. Section 1.2) which are best suited and similar to the urbanization process of the pilot area. Some variables were not able for us due to the time constraint and high requirement permission demand (classified files such as topographic maps, etc.) for permission access due to privacy of entity according to the Mongolian law on State and Official Secrets. Therefore, remotely sensed data and tools are most preferable source. We selected probable 8 variables: a distance to the active economic center; distance from the residential area; distance to the nearest urbanization; distances to the education zone; distance from the road; terrain elevation; surface slope; and population density. These 8 independent variables chosen as driving forces contributing to the expansion of urban areas were at first provided as raw data by several institutions, then we have been processed it as a raster map and these are illustrated in Figure 3. Land-use zoning, demography and road network data were acquired as vector data in September 2019 from the Land Management and Urban Development Authority of Darkhan city and urban land use planning private company "Lanres" at a scale of 1:25,000. The Agency of Land Affairs, Geodesy and Cartography provided anonymous data (excluded all private entity's personal and business data which are prohibited by privacy law to open access) on cadastral mapping (not only display land fee and taxation information but also provides data on land-use types, building type and height, etc.) with a 1:2000 scale of onsite survey. Using the cadastral digital attribute data with parcel land-use purposes, we segregated and re-classed the urban area into residential, industrial and slum settlement use.

The next table explains a summary of the different raster map used as the independent variable's base data for this study. Two raster maps were generated from the GDEM. These were slope and elevation. Global Digital Elevation Model (GDEM) v.1 at 30 m horizontal resolution generated from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) satellite instrument. This is distributed at no charge for scientific purposes by USGS (http://earthexplorer.usgs.gov/). All data processing was done with ESRI ArcGIS (with 3D and Spatial Analyst extensions) and ERDAS Imagine 8.1 and additional details on how these data were created are presented in Table 3.

Data Considered	Source	Method of Creation	
Urban areas	Landsat satellite images	Supervised classification	
Forest areas (urban internal part reclassed as open space)	Landsat 7 satellite image	Supervised classification	
Residential, industrial (urban area reclassed)	Land use classification map Land use zoning map	Created by Land Management and Urban Development Authority of Darkhan city with the participation of authors	

Slope and	Aster GDEM of	Geo-
Vegetation, forest (reclassed as open space)	Landsat satellite image	Normalized Difference Vegetation Index (NDVI)
Slum /sprawl (ger area), residential- middle/high rise	Cadastral map	Created by Agency of Land Affairs, Geodesy and Cartography and Lanres company
Road network	Road map	Digitizing/georeferen cing

Table 3. Data sources used for the study Source: Myagmartseren et al., 2018

3. RESULT AND DISCUSSION

3.1 Land use and land cover change

The three temporal images of 1999, 2009 and 2019 were classified and used to estimate the change of 8 classifications (Figure 2). This analysis was used to identify changes in several land use classes, such as an increase of an urban slum area, a cut of an open space land, etc.



Figure 2. LUCC of Darkhan city by 1999, 2009 and 2019

The greater area of Darkhan city is 9980.66 ha which has the residential area 409 ha (4%), vegetation 4038.97 ha (40.4%), slum/ger area 370.89 ha (3.7%), water area 334.28 ha (3.3%), infrastructure 88.1 ha (0.8%), mining lands 25.53 ha (0.2%), industrial sites 165.58 ha (1.6%), open space 4451.16 ha (44.5%), respectively as shown in Table 4.

	1999		2009		2019	
Land use	Area	Area	Area	Area	Area	Area
	(na)	70	(na)	70	(na)	70
Residential	409.71	4	542.9	5.4	596.4	5.9
Vegetation	4038.97	40.4	3818.76	38.2	3600.54	36
Slum/Ger	370.89	3.7	772.67	7.7	1180.12	11.8
Water	334.28	3.3	305.76	3	331.46	3.3
Infrastructure	88.1	0.8	93.6	0.9	93.14	0.9
Mining	25.53	0.2	73.9	0.7	365.34	3.6
Industrial	165.58	1.6	292.9	2.9	372.36	3.7
Open space	4451.16	44.5	3957.4	39.6	3318.57	33.2
Total	9980.66	100	9980.66	100	9980.66	100

Table 4. LUCC of Darkhan city (1999, 2009 and 2019)

The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLIII-B3-2020, 2020 XXIV ISPRS Congress (2020 edition)

In comparison to the urban area of the Darkhan city between 1999 and 2019, the spatial distribution of the expansion to the natural vegetation cover and into open space has intensified (Figure 3). In other words, natural land cover and vacant land or spaces had been transformed to the urban extent. Between 2009 and 2019, the residential area increased by 0.5%, slum ("ger") area by 4%, mining area by 2.9%, industrial area by 0.7% of the total area and controversially, pasture land decreased by 2%, and open space reduced by 6.4%.



Figure 3. Land use changes in Darkhan city

Compared to the increased expansion of slum areas between 1999 and 2019, the open space of the city and vicinity's vegetation areas have been rapidly diminished. Residential, industrial and mining land use as the main demand for socioeconomic growth has increased by 1-4% and it is expected to increase further (Table 5).

	Change of	area (ha)	Change of area (%)	
Land Use	1999/	2009/	1999/	2009/
	2009	2019	2009	2019
Residential	133.19	53.5	1.3	0.5
Vegetation	-220.21	-218.22	2.2	-2
Slum/Ger	401.78	407.45	4	4
area	401.78	407.45	4	4
Water	-28.52	25.7	0.2	-0.2
Infrastructure	5.5	0.46	0	0
Mining	48.37	291.44	0.4	2.9
Industrial	127.32	79.46	1.2	0.7
Open space	-493.76	-638.83	4.9	-6.4

Table 5. LUCC of Darkhan city

3.2 Urbanization probability prediction

This study examines in more detail the importance and application of logistic regression analysis. In order to analyse urban expansion factors in a multivariate logistic regression analysis, probable and statistically significant factors were selected that influence the expansion of cities. The describing raster maps of eight variables were used to create the predicted urbanization map is illustrated in Figure 4.



Figure 4. Variables affecting the urban expansion: a. Distances to the active economic center; b. Distance from the residential area; c. Distance to the nearest urbanization; d. Distances to the education zone; e. Distance from the road; f. Elevation; g. Slope; h. Population density

These variables were analysed by multivariate logistic regression analysis and in table 6 presented the eight driving factors coefficient and odds ratio. Altogether the 8 variables are significant (p-value <0.05 in a two-tailed test at 95% confidence level) in the logistic regression model. According to our model, the explanatory variable representing physical constraints to urban construction development: the slope and elevation with the highest coefficient (1.9 and 1.4) are most influential to urbanization. The urban expansion in a lower and flat terrain (in our case territories along the river valleys and upper terraces) has a higher probability than in the area has a steep slope degree and high-altitude value (mountain range and hillock). For example, the terrain slope degree has a coefficient equal to 1.9 and an odds ratio of 1.008. This means the odds of the urbanization process in the flat and low slope area is 1.9/1.088 or 1.75 times greater than those terrains with a steep slope. In the locational factors, distance to education zone and distance to the active economic center is more and less important among the socioeconomic explanatory variables. The study result presents the distance to education zone has an urbanization coefficient 0,9 and odds ratio equivalent to 1.0009, which means the urbanization probability of the places nearer to kindergarten, schools and universities are 0.89 times greater than the area more remotely situated.

Variables	Coefficient	Odds ratio
Slope	1.9	1.008
Distance to education zone	0.952	1.000
Distance to the nearest urbanization	-0.014	0.999
Population density	-0.078	0.999
Distance from the road	-0.289	0.999
Distance from the residential area	-0.546	0.999
Distance to the active economic center	-0.854	0.999
Elevation	-1,4	0.992

Table 6. Estimated coefficients and odds ratio for the multivariate logistic regression analysis

The other variables with continuous value have a middling coefficient such as a distance from the road and distance from the residential area has an intermediate influence on urbanization. Hu and Lo (2007) stated the population density variables odds ratio less than one (0.99) indicate the higher population area has less probability than an area with a low density of population. The distance to the nearest urbanization has the lowest coefficient and counted as the least significant variable. The least significant coefficient of distance to the nearest urbanization could be explained by the monocentric character of Darkhan city.

As a result of the regression statistics, R^2 was 0.2285 and Chi² (Chi-square) was 2567.4, which suggested the potential for future urban expansion was positive (Table 7).

Total number of observations	260559
Number of 0	245445
Number of 1	15114
At least 1	9.1994
Percentage of 0	5.8006
Percentage 1	24903
-2logL0	11237.2
-2log (likelihood)	8669.7
R ²	0.2285
Chi ² (Chi-square) (df=8)	2567.4402

Table 7. Regression statistics

The spatial expansions of Darkhan city based on the logistic regression analysis are shown in Figure 5. Darker tones present a higher probability value of urban growth. By the visualization of the urbanization probability maps of Darkhan city, open space and the vegetation areas will be gradually reduced by urban areas. Due to the high coefficient of slope degree, urban probability along the riparian zone or near the riverbank areas are predicted to be urbanized. The rest of the possible areas is mapped close to the recent urban area or adjacent to them. That kind of probability locational reason is the odds ratio equivalent to 1.000 of the distance to the education zone which is located in existing urban sub-centers of the city.



Figure 5. Urbanization prediction map of Darkhan city

4. CONCLUSION

In this study, the logistic regression model was used to estimate the process of urban expansion and future trends of Darkhan city. The study concludes:

Our result estimates note an increase in the total land area (9980.66 ha) of the Darkhan city the built-up area by 1.3% and 0.5%; slum area by 4.03% and 4.08%: mining area by 0.5% and 2.9%; the industrial area by 1.3% and 0.8%, while indicated decrease of the pastureland by 2.2% and 2.19%; open space (vacant land) by 4.9% and 6.4% during the period 1999-2009 and 2009-2019 respectively.

Briefly, expansion of the urban area is expanded by the slum district area, which is gradually increasing as well as high and medium-rise built-up area is expected to growth.

The take-home message of this research is that comparing the urbanization in Darkhan between 1999 and 2019, there has been

a rapid expansion of built-up area into the pasture and open space that have not been allocated for use. Urbanization is a form of anthropogenic land degradation in the research area.

ACKNOWLEDGMENTS

Our sincere gratitude to Ms. T. Solongo, Mr. T. Enkh-Amgalan and colleagues of the Laboratory of Land Survey and Planning, National University of Mongolia, the Land Management and Urban Development Authority of Darkhan and Lanres company for the provision of digitized data and research materials. The research has received funding from the National University of Mongolia under base research grant for professors.

REFERENCES

Al-shalabi, M., Billa, L., Pradhan, B., Mansor, S., Al-Sharif, A. A. A., 2012. Modelling urban growth evolution and land use changes using GIS based cellular automata and SLEUTH models: the case of Sana'a metropolitan city, Yemen. Environmental Earth Science. Springer, 70(1), 425–437. https://doi:10.1007/s12665-012-2137-6.

Al-Sharif, A. A. A., Pradhan, B., 2014. Urban Sprawl Analysis of Tripoli Metropolitan City (Libya): Using Remote Sensing Data and Multivariate Logistic Regression Model. J Indian Soc Remote Sens, 42(1), 149-163. https://doi:10.1007/s12524-013-0299-7.

Dietzel, C., Clarke, K., 2006. The effect of disaggregating land use categories in cellular automata during model calibration and forecasting. *Computers, Environment and Urban Systems*, 30, 78–101. https://doi.org/10.1016/j.compenvurbsys.2005.04.001.

Eastman, J.R., 2016. TerrSet: Geospatial Monitoring and Modeling System. *Tutorial*, Clark Labs, Clark University.

Eyoh, A., Olayinka, D. N., Nwilo, P., Okwuashi, O., Isong, M., Udoudo, D., 2012. Modelling and predicting future urban expansion of Lagos, Nigeria from remote sensing data using logistic regression and GIS. *International Journal of Applied Science and Technology*, 2(5), 116–124.

Firman, T., 1997. Land Conversion and Urban Development in the Northern Region of West Java, Indonesia. *Urban Studies*, 34(7), 1027–1046. https://doi.org/10.1080/0042098975718.

Gillham, O., 2002: The limitless city: A Primer on the Urban Sprawl Debate. 1st edition. Island Press, Washington.

Helbich, M., Leitner, M., 2010. Post suburban Spatial Evolution of Vienna's Urban Fringe: Evidence from Point Process Modeling. *Urban Geography*, 31(8), 1100–1117. https://doi.org/10.2747/0272-3638.31.8.1100.

Hu, Z., Lo, C. P., 2007. Modeling urban growth in Atlanta using logistic regression. *Computers, Environment and Urban Systems*, 31(6), pp. 667–688. https://doi.org/10.1016/j.compenvurbsys.2006.11.001.

Myagmartseren, P., Myagmarjav, I., 2017. The results of Darkhan city land use classification, zoning and planning. *Agroecology* 2017 (2), 102-120. Hard copy. Mongolian University of Life Science. Scientific bulletin of School of Agroecology. Ulaanbaatar.

Myagmartseren, P., Bazarkhand, Ts., Myagmarjav, I., Munkhnaran, S., 2018. The Fractal Geometry of Urban Land Use: The Case of Ulaanbaatar City, Mongolia. *Land*, 7(2), 67-77. https://doi.org/10.3390/land7020067.

Zhang Q., Wang J., Gong P., Shi P., 2002. Urban built-up land change detection with roads density and spectral information from multi-temporal Landsat TM data. *International Journal of Remote Sensing*. 23(15): 3057-3078.