Table 8. Estimated coefficients and odds ratios for the logistic regression model containing the 7 independent variables

Variable	Coefficient (*B*)	Standard error	Odds ratio	Sig.	Exp (B)
X1 (P_URBAN)	0.685	0.012	0.984	0.000	1.984
X2 (DIST_Euu)	-0.007	0.000	-0.007	0.000	0.993
X3 (Dist_Mjrd	0.000	0.000	0.000	0.710	1.000
X4 (DIST_Euu)	0.000	0.000	0.000	0.000	1.000
X5 (Dist_Cc)	0.000	0.000	0.000	0.096	1.000
X6 (Dist_Nr)	0.000	0.000	0.000	0.000	1.000
X7 (C_Dist_Nuc)	0.000	0.000	0.000	0.000	1.000

South sector

Agriculture Land parcels with more surrounded urban areas were more likely to be converted to urban land-use. The variable of the urban cells within a neighborhood of 7 by 7 cell sizes (X1) had an odds ratio = 5.926. With an increase of 1 urban cells within the neighborhood, the odds of development will increase by 1.935. The use of a land parcel was affected by the land-use type surrounded areas. Thus, urban sprawl preferred to occur in locations surround the existing urban uses (X2). The value of odds ratio = - 0.008, or- 8/1000, referring that, the probability of sprawl in areas near to existing cores is more than areas located farther that. The odds of sprawl would decrease by 8 if the distance to existing urban areas decreased by 8 meters.

Urban sprawl prefers to direct around the nearest county urban center (X5) which are occupied by low-class residential land-use and small shops (Table 9). The odds ratio of distance to county centers (X5) = -0.001, or- 1/1000. The odds of sprawl in areas closer to urban areas were signed 1000 times more than odds of 1 km further areas. This indication that dragging force has taken impact in nearby centers where infrastructure, services, and daily needs commercial. The odds ratios of the driving factors: the distance to nearest major roads (X3), distance to the CBD (X4), cost distance to active Main urban centers (X7), and distance to Nile River (X6), The odds ratio for (X3, x4, x5, x6, and x7) = 0 for all of them, thus, there were no impacts of decentralization trends of sprawl in the South sector of GCMR. Urban sprawl tends to be located surrounded the existing urban centers to obtain infrastructure and essential services from the nearest existing urban point. In the South sector, local people use the unpaved roads for the daily trips to services and work. Therefore, the impacts of major roads were very limited on urban sprawl (Table9).

Table 9. Estimated coefficients and odds ratios for the logistic regression model containing the 7 independent variables

Variable	Coefficient (B)	Standard error	Odds ratio	Sig.	Exp (B)
X1 (P_URBAN)	1.935	0.060	5.926	0.000	6.926
X2 (DIST_Euu)	-0.008	0.000	-0.008	0.000	0.992
X3 (Dist_Mjrd	0.000	0.000	0.000	0.008	1.000
X4 (DIST_Euu)	0.000	0.000	0.000	0.000	1.000
X5 (Dist_Cc)	-0.001	0.000	-0.001	0.000	0.999
X6 (Dist_Nr)	0.000	0.000	0.000	0.323	1.000
X7 (C_Dist_Nuc)	0.000	0.000	0.000	0.301	1.000

PREDICTION OF URBANIZATION PROBABILITY

The probability of urban sprawl was predicted by connecting the coefficients of logistic regression model including the 7 considerable driving factors (M7) into Eq. (2). To take temporal dynamics into as much estimation, and raster layers were improved with newer cell values. The highly affected driving factors in forming the urbanization probability map were (Figure, 6): in north sector, number of urban cells within a 7 by 7 Neighborhood (X1), distance to the nearest Existing urban area (X2), and distance to the nearest major road (X3). In middle sector, number of urban cells within a 7 by 7 Neighborhood (X1), and distance to the nearest Existing urban area (X2). In south sector, number of urban cells within a 7 by 7 Neighborhood (X1), distance to the nearest Existing urban area (X2), and Distance to the nearest County Center (X5). The map of the predicted urbanization probability is shown in Figure 6, which is a ramp color measures from 0- 100. The darker tones indicate higher probabilities of urban Sprawl. The future urban distribution pattern is easily observable from this map. Some new urban spots far from existing urban areas can be seen. Most probable areas for urban development are closer to existing urban areas in addition to specific probable places around major roads in north sector.

International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 10 (2018) pp. 8467-8481

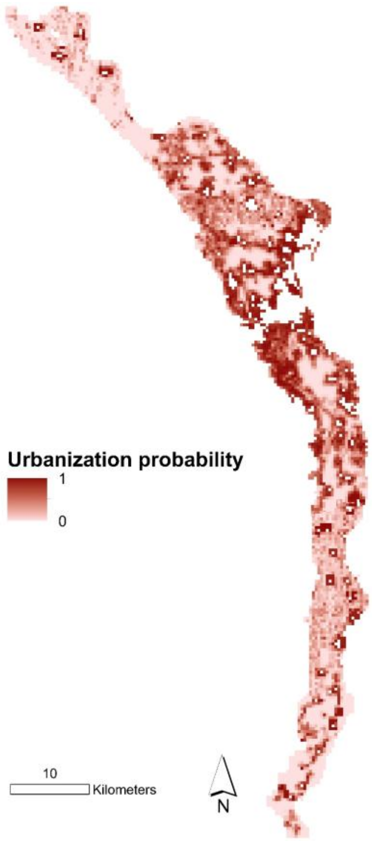


Figure 6. Urbanization probability map of Cairo

MODEL EVALUATION

The simulated urban sprawl cells were compared with the real urban cells that indeed converted from non-urban to urban land-use 2004-2013. The ROC compares binary data over the entire range of simulated probabilities. It assembles into one index of agreement to represent the model probability to expect the potential urban land-use distribution in study sectors. ROC method is a prime method to examine the efficacy of logistic models by expecting the occurrence of a phenomenon through comparing the probability image with a binary image displaying the actual occurrence of this phenomenon. (Schneider & Pontius, 2001).

To carry out this model evaluation, the map of urban sprawl probability predicted from the logistic regression model was compared against that of actual urban sprawl (reference image) created by comparison of the 2004, and 2013 land-use maps derived from satellite images. Tables 11, 12, and 13 show the number of true positive cells, which are representing the potential urban sprawl cells and the actual urban sprawl in the reference image. By Considering B as the number of false positive cells, C as the number of false negative cells, and D as the number of true negative cells, one data point (x, y) was generated where x is the rate of false positives (false positive % = B / B+C) and y is the rate of true positives (true positive % = A / A+C):

Table 10. Contingency table showing the comparison of the expected urban sprawl probability with the reference image

		Predicted probability	
		No urban sprawl (0)	Urban sprawl (1)
	No urban sprawl (0)	D (true negative)	C (false negative)
Observed (Reference image)	Urban sprawl (1)	B (false positive)	A (true positive)

Table 11. shows the comparison of the predicted urban Sprawl probability with the reference points in the North Sector

Observed (Reference image)		Predicted probability		
		Urban Sprawl (Y)		Percentage Correct
		(Y=0)	(Y=1)	
Urban Sprawl (Y)	(Y=0)	5701	61	98.9%
	(Y=1)	209	197	48.5%
Overall Percentage				95.6%

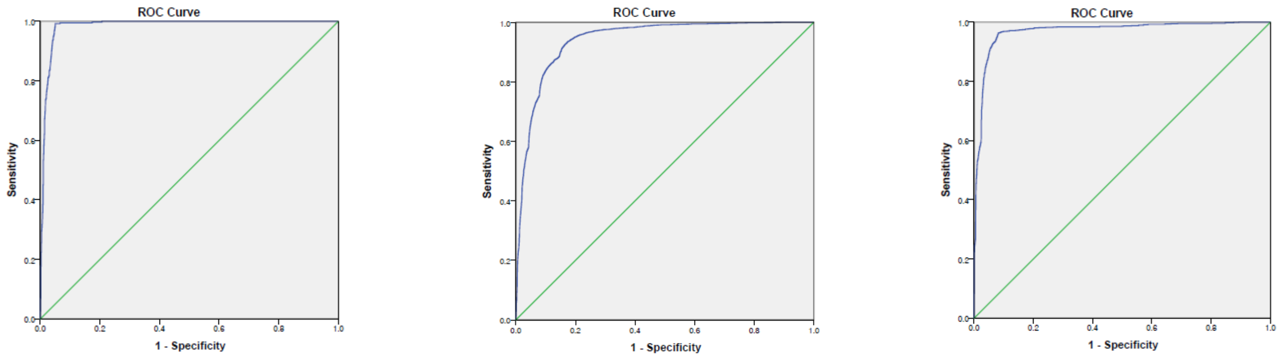
Table 12. shows the comparison of the predicted urban Sprawl probability with the reference points in middle Sector

Observed (Reference image)		Predicted probability		
		Urban Sprawl (Y)		Percentage Correct
		(Y=0)	(Y=1)	
Urban Sprawl (Y)	(Y=0)	15849	1042	93.8%
	(Y=1)	1439	3455	70.6%
Overall Percentage				88.6%

Table 13. shows the comparison of the predicted urban Sprawl probability with the reference points in the South Sector

Observed (Reference image)		Predicted probability		
		Urban Sprawl (Y)		Percentage Correct
		(Y=0)	(Y=1)	
Urban Sprawl (Y)	(Y=0)	9508	207	97.9%
	(Y=1)	383	588	60.6%
Overall Percentage				94.5%

North Sector | Middle Sector | South Sector



Area Under Curve

.985 | .941 | .970

Figure 7. ROC curves of the logistic regression model for the submitted three study sectors

The data points were linked to generate an ROC curve from which the ROC value was estimated. As shown in figure 7, the ROC value is the area under the curve that links the plotted points. Sensitivity is the proportion of true positives or the proportion of cases correctly identified by the test as meeting a certain condition while Specificity is the proportion of true negatives or the proportion of cases correctly identified by the test as not meeting a certain condition.

DISCUSSION AND CONCLUSION

For getting a better understanding about the interactions between the mutable types of urban sprawl and related physical driving factors in GCMR, a binary logistic regression model has been established for 2004-2013. The model was submitted to multicollinearity analysis and detected to yield VIF <10. The significance of proportion of urban cells, distance to Major roads, and distance to Existing urban users grown over time in the whole sectors, while the significance of distance to CBD, distance to County Centers, and distance to Nile river are quite decreased in the middle and increment in fringe zones. This agreed with Osman et al., 2016 study of spatial metrics that show external sprawl followed by fragmentation in the fringe areas and intensification of the core.

Moreover, the logistic model has been applied in discussing the role of spatial driving factors in formulating urban sprawl characteristics in GCMR, and to generate a probability map to present where urban sprawl would take place in the future. The outcome pointed out that proportion of the urban cells within 7 by 7 neighborhood force has the highest odds ratio, which means that the urban cells sprawl will occur in a proliferation pattern in lack of obvious strategies of development in GCMR. That explained the importance of urban cells proportion as a driving force. Therefore, urbanization was more potentially to occur close to existing urbanized areas and major roads and to avert isolated areas with higher living expenses and transportation costs which imply the importance of these factors in this study.

The model findings reported that proportion of the urban cells within a 7 by 7 Neighborhood (+ve), distance to nearest excising urban areas (-ve) in whole study sectors, distance to nearest major roads (-ve) in north and middle sectors, and distance to nearest county centers (-ve) in south sector were the most active driving factors in during 2004 -2013. Most of this study findings are consistence with similar study findings of Hu & Lo, 2007; B. Huang et al., 2009; Vermeiren et al., 2012.

The logistic regression model had a low performance to specify where urbanization occurrence in north sector but it was highly effective for middle and south sectors. Therefore, potential urbanization types for north sector could to be predicted by various techniques like distribution cells that constitute the predestined size of the desired urban land on probability map. This method could identify information about urbanization tendencies in high accuracies which will be more helpful for decision makers to manage urban sprawl.

Regarding to previous urban plans for GCMR which performed starting from 1956 and followed by new plans in 1973, 1982, 1991, and 2006 (Osman et al.,2015b) However, urban sprawl rates increased over time from 3.4% annually in 1990s to more than 6.3% in 2010s (Osman et al.,2016). The fail of authorized urban plans to find the real driving forces of local communities of buildable-lands demand and use them as pillars to prepare and enforce plans was the main reason for this severe increase in sprawl rates. The GCMR local community has their needs of buildable-lands demands for residential and business activities which particularly ignored in authorized urban plans .Therefore, Local community meet these needs by urban sprawl activities out of the authorized plans. Strict urban planning procedures should be taken to use the active driving factors as tools to direct the urban plans in GCMR to the right way. Otherwise, the negative impacts of urban sprawl will be persisted in a haphazard way. Consequently, the outcome of this model can highly assist decision makers in manage development by formulating various options of future urban sprawl scenarios.

The urbanization history of GCMR indicated a persistent increment of Agriculture land, and the potential urbanization findings in this study proved that with more potentials for more environmental deterioration in the future. This should increase the warnings for decision makers and urban planners. Therefore, the GCMR desperately requires of a real urban plan, and rigorous urban development rules to mitigate the urban sprawl rates for saving Agriculture lands and conserve natural environment.

Diverse considerations have been grasped from this study for future research. First, when utilizing a logistical regression model to investigate urban sprawl, researchers must be attentive about spatial autocorrelation that predominantly occurs in spatially referenced data which overstep the hypothesis of the model. Second, to curb the demerit of logistic regression modeling in transacting with temporal dynamics, further research have to examine the self-modifying method for model variables to be able to improve themselves automatically. Finally, future research have to find the multi-scale attributes of land-use systems by utilizing multi-level statistics or a hierarchical modeling method. Moreover, The modeling findings proved that the variables of roads and existing urban areas had a considerable impact on formulating the future urbanization in GCMR. Therefore, these variables should be regarded in a new study with higher modeling accuracy rates to catch measuring indicators could be used to manage the future development by local urban decision makers.

REFERENCES

[1] 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 the Sana'a metropolitan city, Yemen. Environmental Earth Sciences. doi:10.1007/s12665-012-2137-6.

[2] Angotti, T. (1993). Metropolis 2000: Planning, poverty and politics. London: Routledge.

[3] Bailey, T. C., & Gatrell, A. C. (1995). Interactive spatial data analysis. Harlow Essex/New York, NY: Longman Scientific & Technical/Wiley.

[4] Bian, L. (1997). Multiscale nature of spatial data in scaling up environmental models. In D. A. Quattrochi & M.

[5] Cheng, H. Q., & Masser, I. (2003). Urban growth pattern modeling: a case study of Wuhan city, PR China. Landscape and Urban Planning, 62 (4), 199-217.

[6] Cheng, J., & Masser, I. (2003). Urban growth pattern modeling: a case study of Wuhan city, PR China. Landscape and Urban Planning, 62 (4), 199e217.

[7] Chuvieco, J. Li & X. Yang (Eds.), Advances in Earth Observation of Global Change (pp. 27-42): Springer Netherlands.

[8] Congalton, R. G. (1988). A comparison of sampling schemes used in generating error matrices for assessing the accuracy of maps generated from remotely sensed data. Photogrammetric Engineering and Remote Sensing, 54 (5), 593–600.

[9] Domencich, T. A., & McFadden, D. (1975). Urban travel demand: behavioral analysis. Amsterdam: North- Holland.

[10] Dubovyk, O., Sliuzas, R., & Flacke, J. (2011). Spatio-temporal modelling of informal settlement development in Sancaktepe district, Istanbul, Turkey. ISPRS Journal of Photogrammetry and Remote Sensing, 66 (2), 235e246.

[11] F. Goodchild (Eds.), Scaling in remote sensing and GIS (pp. 13–25). Boca Raton, FL: CRC/Lewis Publishers.

[12] Field, A. (2009). Discovering Statistics Using SPSS: SAGE Publications.

[13] Firman, T. (1997). Land conversion and urban development in the northern region of West Java, Indonesia. Urban Studies, 34, 1027–1046.

[14] Fu, B., Zhang, Q., Chen, L., Zhao, W., Gulinck, H., Liu, G., et al. (2006). Temporal change in land use and its relationship to slope degree and soil type in a small catchment on the Loess Plateau of China. CATENA, 65 (1), 41e48.

[15] Gillham, O. (2002). The limitless city: A primer on the

urban sprawl debate (pp. 328). Washington, DC, USA: Island Press.

[16] GOPP (2005). The national project for preparing the strategic development plans for Egyptian villages. Cairo: GOPP (Last accessed 20.10.13) www.gopp.gov.eg.

[17] Grimm, N., Faeth, S., Golubiewski, N., Redman, C., Wu, J., Bai, X., et al. (2008). Global change and the ecology of cities. Science, 319 (5864), 756e760.

[18] Hosmer, D. W., & Lemeshow, S. (1989). Applied logistic regression. New York: Wiley.

[19] Hu, Z., & Lo, C. P. (2007). Modeling urban growth in Atlanta using logistic regression. Computers, Environment and Urban Systems, 31 (6), 667e688.

[20] Huang, B., Zhang, L., & Wu, B. (2009). Spatiotemporal analysis of rural urban land conversion. International Journal of Geographical Information Science, 23 (3), 379e398.

[21] Huang, B., Zhang, L., &Wu, B. (2009). Spatiotemporal analysis of rural–urban land conversion. International Journal of Geographical Information Science, 23, 379–398.

[22] Jacobson, C. R. (2011). Identification and quantification of the hydrological impacts of imperviousness in urban catchments: a review. Journal of Environmental Management, 92 (6), 1438e1448.

[23] Jiang, B., & Yao, X. (2010). Geospatial analysis and modelling of urban structure and dynamics (Vol. 99, p. 440). Springer: Netherlands.

[24] JICA (2008). The strategic urban development master plan study for sustainable development of the Greater Cairo region in the Arab Republic of Egypt.The final report summary, volume 1. Cairo, Egypt.

[25] Jokar Arsanjani, J. (2011). Dynamic land use/cover change modelling: Geosimulation and multiagent-based modelling (hardback) (series: springer theses) (XVII, p. 139), Springer: Berlin, Heidelberg.

[26] Kaufmann, R. K., Seto, K. C., Schneider, A., Liu, Z., Zhou, L., & Wang, W. (2007). Climate response to rapid urban growth: evidence of a human-induced precipitation deficit. Journal of Climate, 20 (10), 2299e2306.

[27] Kipper, R. And Fischer, M., (eds.), 2009: Cairo's Informal Areas: Between Urban Challenges and Hidden Potentials: Facts. Voices. Visions. Chapter 2: Daily Life in Informal Areas, 49-83.

[28] Kleinbaum, D. G. (1994). Logistic regression: A self-learning text. New York: Springer.

[29] Lambin, E. F., Turner, B. L., Geist, H. J., Agbola, S. B., Angelsen, A., Bruce, J. W., et al. (2001). The causes of land-use and land-cover change: moving beyond the myths. Global Environmental Change, 11, 261–269.

[30] Lin, Y.-P., Chu, H.-J., Wu, C.-F., & Verburg, P. H. (2010). Predictive ability of logistic regression, auto-logistic regression and neural network models in empirical land-use change modeling – a case study. International Journal of Geographical Information Science, 25, 65–87.

[31] Long, Y., GU, Y., & Han, H. (2012). Spatiotemporal heterogeneity of urban planning implementation effectiveness: evidence from five urban master plans of Beijing. Landscape and Urban Planning108 (2e4), 103e111.

[32] López, E., Bocco, G., Mendoza, M., & Duhau, E. (2001). Predicting land-cover and land-use change in the urban fringe: a case in Morelia city, Mexico. Landscape and Urban Planning, 55, 271–285.

[33] McKinney, M. (2008). Effects of urbanization on species richness: a review of plants and animals. Urban Ecosystems, 11 (2), 161e176.

[34] Miller, M. D. (2012). The impacts of Atlanta's urban sprawl on forest cover and fragmentation. Applied Geography, 34, 171e179.

[35] Moellering, H., & Tobler, W. (1972). Geographic variances. Geographical Analysis, 4, 34–50.

[36] Nong, Y., & Du, Q. (2011). Urban growth pattern modeling using logistic regression. Geo-spatial Information Science, 14 (1), 62-67. doi: 10.1007/s11806-011-0427-x.

[37] O'brien, R. (2007). A Caution Regarding Rules of Thumb for Variance Inflation Factors. Quality & Quantity, 41 (5), 673-690. doi: 10.1007/s11135-006-9018-6.

[38] Osman, T.; Arima, T.Divigalpitiya, P., Measuring urban Sprawl patterns in Greater Metropolitan Cairo Region, Journal of the Indian Society of Remote Sensing, 2016. Springer Publications. DOI: 10.1007/s12524-015-0489-6.

[39] Osman, T.; Divigalpitiya, P.; Arima, T., Modeling urban growth scenarios in Cairo Metropolitan Region 2050, Proceedings of the 14th International Conference of Computers in Urban planning and Urban management CUPUM 2015a, Massachusetts institute of technology MIT, Boston, USA.

[40] Osman, T.; Divigalpitiya, P.; Arima, T., Effect of Governmental Housing Regulations on the Egyptian Housing Market: Focusing on Greater Cairo Metropolitan Region, Journal of Architecture and urban Design, Kyushu University, No. 28, pp. 1- 9, July 2015b.

[41] Overmars, K. P., & Verburg, P. H. (2005). Analysis of land use drivers at the watershed and household level: linking two paradigms at the Philippine forest fringe. International Journal of Geographical Information Science, 19, 125– 152.

[42] Schneider, L., & Pontius, R. G. (2001). Modeling land

use change in the Ipswich watershed, Massachusetts, USA. Agriculture, Ecosystems and Environment, 85, 83–94.

[43] Séjourné, M. (2006). Les politiques récentes de “traitement” des quartiers illégaux au Caire. PhD Thesis, Université de Tours, France.

[44] Seto, K. C., Fragkias, M., Güneralp, B., & Reilly, M. K. (2011). A meta-analysis of global urban land expansion. PLoS ONE, 6 (8), e23777.

[45] Sims, D. (2003). The case of Cairo, Egypt, UN habitat, global report on human settlements 2003, the challenge of slums, Earthscan, London.

[46] Sims, D., & Séjourné, M. (2008). The dynamics of peri-urban areas around greater Cairo: A preliminary reconnaissance. Washington, D.C.: World Bank.

[47] Tayyebi, A., Delavar, M., Yazdanpanah, M., Pijanowski, B., Saeedi, S., & Tayyebi, A. (2010). A Spatial Logistic Regression Model for Simulating Land Use Patterns: A Case Study of the Shiraz Metropolitan Area of Iran. In E.

[48] Thapa, R. B., & Murayama, Y. (2010). Drivers of urban growth in the Kathmandu valley, Nepal: examining the efficacy of the analytic hierarchy process. Applied Geography, 30 (1), 70e83.

[49] United Nations. (2012). World urbanization prospects: The 2011 revision. Available at: (Last accessed 25.09.14) http://esa.un.org/unpd/wup/index.htm.

[50] Veldkamp, A., & Lambin, E. F. (2001). Predicting land-use change. Agriculture, Ecosystems & Environment, 85, 1–6.

[51] Verburg, P. H., Kok, K., Pontius, R. G., & Veldkamp, A. (2006). Modeling land-use and land-cover change. Land-Use and Land-Cover Change, 117–135.

[52] Weigel, S. J. (1996). Scale, resolution and resampling: representation and analysis of remotely sensed landscapes across scale in geographic information systems. Ph.D. Dissertation, Louisiana State University.

[53] Wu, K., & Zhang, H. (2012). Land use dynamics, built-up land expansion patterns, and driving forces analysis of the fast-growing Hangzhou metropolitan area, eastern China (1978e2008). Applied Geography, 34 (0), 137e145.

[54] Yang, X., & Lo, C. P. (2003). Modeling urban growth and landscape changes in the Atlanta metropolitan area. International Journal of Geographical Information Science, 17 (5), 463–488.

[55] Youssef, A. M., Pradhan, B., & Tarabees, E. (2011). Integrated evaluation of urban development suitability based on remote sensing and GIS techniques: contribution from the analytic hierarchy. Arabian Journal of Geosciences, 4 (3–4). doi:10.1007/s12517 -009-0118-1.

[56] Zhao, Y., &Murayama, Y. (2011). Urban dynamics analysis using spatial metrics geosimulation. Spatial Analysis and Modeling in Geographical Transformation Process, 153–167.