

Socio-economic factors of high trash generation in the city of Los Angeles

Xiaoyu Liang, Mi-Hyun Park  and Michael K. Stenstrom 

ABSTRACT

Trash is one of major pollutants in urban runoff. Some studies have been conducted to verify the different impacts of land use on trash generation in a qualitative way and focused on the performance of trash control measures. Few studies have explored the human impacts on trash generation or developed a quantitative model to describe the phenomenon. This paper examined the impact of human activity on trash generation. Spatial regimes on high trash generation areas were identified using the selected variables from best subset model regression and validated with Moran's I scatter plot and spatial analysis of variance. Bidirectional spatial lag regression with regimes was performed to develop the final model to explain the spatial distribution of trash generation and identify its major causes. The result showed that economic status and occupation of the population were correlated with trash accumulation and the dominant land use type, and the distance to rivers most affected trash generation. The effects of these indicators were different within and outside the high trash generation areas.

Key words | GIS modeling, spatial lag model, stormwater pollution, trash generation

Xiaoyu Liang
School of Remote Sensing Information
Engineering,
Wuhan University,
Wuhan, Hubei,
China

Mi-Hyun Park  (corresponding author)
Michael K. Stenstrom 
Department of Civil and Environmental
Engineering,
University of California,
Los Angeles, CA,
USA
E-mail: parkmh@ucla.edu

INTRODUCTION

Urban runoff discharged through drainage systems has been identified as a major vehicle of pollutants such as debris and trash to local water bodies. The previous study on the composition and distribution of marine litter on the Orange County beaches showed that water was a more important vehicle to deliver trash into the ocean than direct abandonment and wind carriage (Moore *et al.* 2001). Trash in waterways adversely affects the quality of receiving waters, aquatic life and public health (City of LA 2002; Kim *et al.* 2004). Floatable trash can reduce the growth of aquatic vegetation and habitats for aquatic life while settleable trash can contaminate sediment (City of LA 2002). To address the problem, the federal Clean Water Act of 1972 requires each state to list impaired waters and establish a pollutant-specific Total Maximum Daily Load (TMDL) (US Senate 2002). Accordingly, Los Angeles County established 'zero' trash objectives (California RWQCB 2007), which can be interpreted as no trash accumulating in deleterious amounts and no trash directly discharged into local water bodies.

Many studies have focused on the efficiency of trash capture systems based on the trash characteristics (Lau *et al.* 2001; Kostarelos *et al.* 2011; Alam *et al.* 2017; Liu *et al.* 2019). Street sweeping is effective in removing trash and has been shown to reduce stormwater contaminants such as total suspended solids and chemical oxygen demand (COD) (Kang *et al.* 2009) and prevent growth of bacteria during low flow periods. However, only a few have explored the main drivers of trash generation that could impact the receiving water quality. The Stormwater Management Division of the City of Los Angeles (2002) conducted trash generation studies to aid assessment of trash control measures. They identified the radiating spatial pattern of trash generation with qualitative analysis by comparison of thematic maps for trash generation and certain indicator distribution. McCuen *et al.* (2014) examined the seasonal pattern of trash deposition in local waters of Baltimore and developed predictive models with environmental characteristics for five trash types in Maryland. Núñez *et al.* (2019) investigated the spatial distribution, the temporal pattern and the physical features of transport of the marine litter in an estuary of the northern coast of Spain to predict the litter accumulation. Quantitative models were developed in these studies, but no indicators related to humanistic characteristics were examined and

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY-NC-ND 4.0), which permits copying and redistribution for non-commercial purposes with no derivatives, provided the original work is properly cited (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

doi: 10.2166/wst.2019.283

the models cannot be directly applied to other areas due to different natural and socio-economic conditions.

Many studies used regression analysis to identify socio-economic factors (Bandara *et al.* 2007; Kannangara *et al.* 2018; Ramachandra *et al.* 2018). However, these studies referred to previous literature or surveys to select indicators for developing their regression models. A study was conducted with no priori, extracting key socio-economic factors of municipal solid waste generation in Styria, Austria, from 116 variables of census data to develop an efficient model to validate previous studies (Lebersorger & Beigl 2011). However, all these studies did not consider the influence of neighboring areas and littering behaviors that are largely influenced by neighborhoods (Roales-Nieto 1988). Geographic information systems (GIS) are frequently used to identify the spatial patterns of some geographic phenomena combining regression analysis, although few have employed GIS methods to trash generation problems. Favot & Grasseti (2017) assigned dummy variables to three macro regions in regression analysis to explore the regional characteristics of E-waste collection. Ghermandi (2016) analyzed the hot and cold spots of public use in two frequently visited natural treatment systems to identify its spatial distribution with GIS tools. Keser *et al.* (2012) compared the results of spatial autoregression and geographically weighted regression to examine the difference of socio-economic factors influencing municipal waste generation at the local and global scales. Conley *et al.* (2019) employed spatial analysis

and considered the spatial dependency to investigate the trash condition in an improved model for tracking trash reduction. This study will combine the aforementioned GIS techniques to exploit the influence of each individual indicator and the spatial relations on the trash generation pattern.

This study aims to fill in the gap between socio-economic analysis and trash generation from the geographic and quantitative perspectives. The main objectives are two-fold: (1) to identify the trash generation drivers in human society; (2) to construct a quantitative model to describe the causal relationship. To achieve the first goal, best subset model regression and spatial regimes will be employed to select the most statistically significant combination of indicators to predict trash generation. To achieve the second goal, bidirectional spatial lag regression with regimes will be exerted to bridge the trash generation and the potential indicators.

MATERIAL AND METHODS

Study area

The City of Los Angeles covers a total area of 1,302 km², comprising 1,214 km² of land and 88 km² of water with the Los Angeles River as the main drainage channel and a flood control channel. The city is intersected by four watersheds (Figure 1): 744 km² of Los Angeles Rivers Watershed,

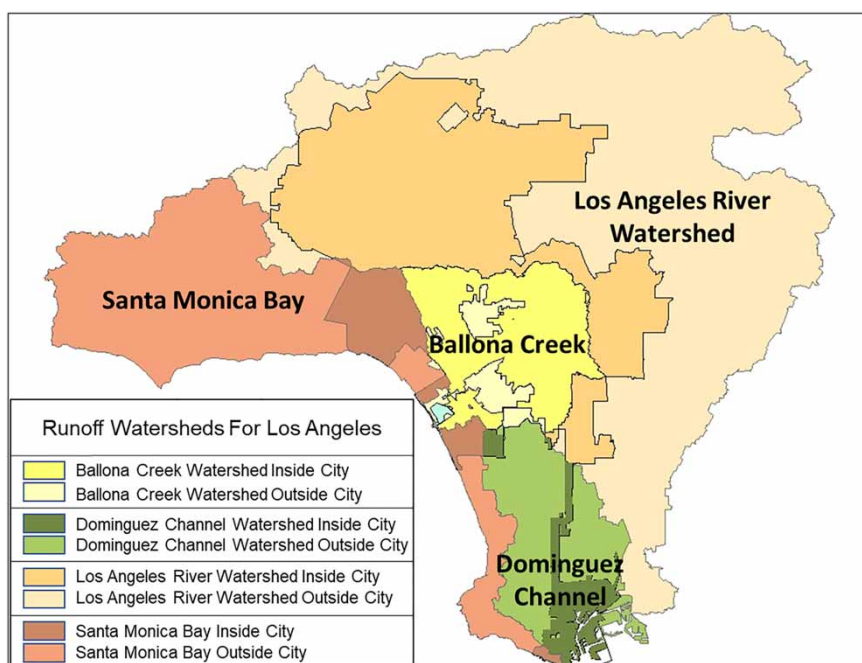


Figure 1 | Runoff watersheds for the City of Los Angeles (modified from <http://www.lastormwater.org/about-us/about-watersheds/>).

276 km² of Ballona Creek, 111 km² of Santa Monica Bay Watershed and 95 km² of Dominguez Channel. The southern portion of the Los Angeles River Watershed captures runoff from urbanized areas surrounding downtown Los Angeles. The major land uses include residential (36%), open space and agricultural (44%) (<http://www.lastormwater.org/about-us/about-watersheds/los-angeles-river/>). The Ballona Creek flows from mid-Los Angeles to the Pacific Ocean at Playa del Rey. The watershed is composed of residential (59%), open space (17%) and commercial (14%), resulting in 49% being impervious surfaces (<http://www.lastormwater.org/about-us/about-watersheds/ballona-creek/>). The Santa Monica Bay Watershed is located to the east of Pacific Ocean, with 89 km of coastline and beaches. The watershed has around 200 separate storm drain outlets and contributes over 114 million m³ of runoff to the bay per year. The highly urbanized areas within the watershed are dominated by residential (32%) and open space (48%) (<http://www.lastormwater.org/about-us/about-watersheds/santa-monica-bay/>). The Dominguez Channel Watershed is located in the south of the Los Angeles Basin. Approximately 176 km² of the watershed drains to the Dominguez Channel and the remaining drains directly to the Los Angeles Harbor. The watershed is highly impervious (61%), dominated by residential use (41%) and other urban land uses such as industrial, commercial, and transportation-related land use (44%) (<http://www.lastormwater.org/about-us/about-watersheds/dominguez-channel/>).

The population of the city is diverse. According to Census Bureau 2000 (<https://www.census.gov/quickfacts/fact/table/losangelescitycalifornia/PST045217>), the city has population of approximately 3.7 million with Whites (46.9%), African Americans (11.2%), Native Americans (0.8%), Asians (10.0%), Pacific Islanders (0.2%), other races (25.7%), and 5.2% from two or more races in 2000.

Data acquisition and preprocessing

The trash data as the response variable were obtained from the City of Los Angeles (2002) and their trash data were collected mostly during catch basin cleaning for storm-water runoff mitigation. The trash volume was normalized to the area of each sub-basin. The explanatory variables were divided into two perspectives: (1) static population composition; (2) dynamic human activities. Population composition came from census data of 2000 Census Bureau (<https://www.census.gov/census2000/states/ca.html>), which is the nearest survey year to the trash collection date. Human activities can be indirectly reflected by the

land use types and proximity to rivers. Land use maps were downloaded from Los Angeles County GIS Data Portal (<https://egis3.lacounty.gov/dataportal/>). The river vector feature was obtained from the United States Geological Survey (<https://water.usgs.gov/maps.html>).

The trash and the river layers were re-projected to be on the same projection and coordinate system as the land use and census data. Then, all data were combined into the same geographic unit for further analysis. The attributes with missing information in some census tracts were removed. Overlap analysis was performed to attach trash volume to each census tract. All the lines were transformed into distance raster layers by averaging the distance from each feature to the investigating analysis unit. The proportion of each land use type was calculated within census tracts. The processes were run on ESRI ArcGIS v. 10.5.

GIS and statistical modeling

The model was developed through two processes: indicator selection and model building (Figure 2). Indicator selection was performed using exploratory regression in ArcGIS and best subsets regression to remove redundant or unrelated explanatory variables. Spatial analysis of variance (ANOVA) was performed to validate the existence of spatial

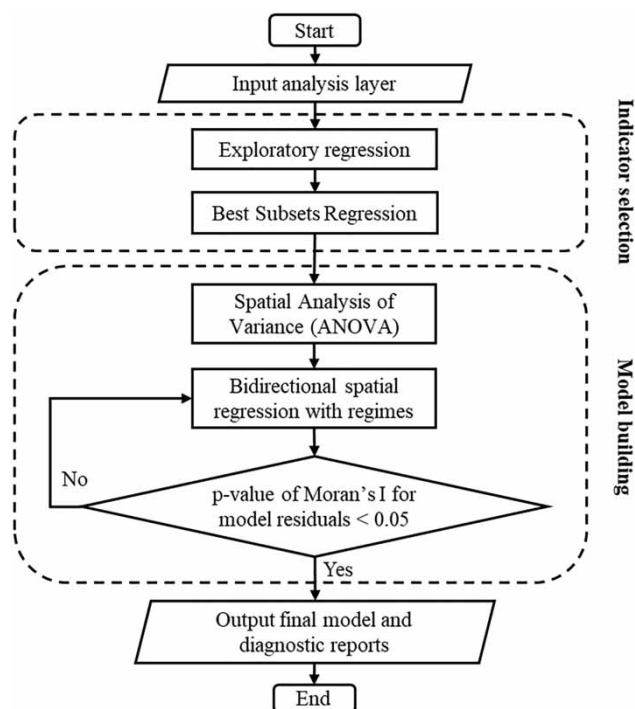


Figure 2 | Flowchart of model specification.

heterogeneity in the dependent variable over the City of Los Angeles. Bidirectional spatial regression with regimes was performed to reduce the residual autocorrelation to a statistically accepted level. The final model can explain the geographic characteristics of trash generation over the City of Los Angeles and provide trash generation prediction and practical understanding. We used ESRI ArcGIS v. 10.5 and R v. 3.4.1 for our analyses and modeling.

Indicator selection

Exploratory regression and best subset regression were used in sequence to select the proper indicators for modeling to provide a baseline for spatial analysis. Exploratory regression was first employed to test all possible combination of the variables according to the expected sizes of the model. During the process, multicollinearity and significance analyses among all explanatory variables were conducted to remove redundant or unrelated variables while ensuring statistical significance. This step was initialized with the Ordinary Least Squares (OLS) regression with two explanatory variables. Then the number of variables was increased in the next iterations, retaining the most significant ones among the multicollinear group of variables. Since the significance of each variable changed as the model size became larger, the threshold for variable removal was set to be 90% significant to prevent removing meaningful variables with low significance at the beginning of the process. The redundant variables were removed when the value of variation inflation factor (VIF) is high (>7.5). The process stopped when the number of the remaining variables was below 50. Based on the output from the previous analysis, best subsets regression (Furnival 1971) was then performed on the subsets of the selected variables to identify the most affecting combination for model building.

Model building

The final model was developed based on the OLS regression model using the selected indicators. The test results of OLS regression indicated whether and what spatial effects exist. Autoregressive models, i.e. spatial regime models and spatial regression with regimes, were employed to remove such spatial effects in the final model, including spatial heterogeneity and spatial autocorrelation.

Spatial heterogeneity cannot be sufficiently explained from a global model such as OLS regression because the coefficients of the explanatory variables should be calculated differently for areas with different influential

mechanisms. Spatial regime addresses such difference by assigning the regime to the model as a dummy variable, where the geographic units within the regime are set to be 1 while the rest are set to be 0 (Ertur et al. 2006). When calculating the spatial regime model, the models within and outside the regime differ as the equation below.

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix} \quad (1)$$

where y_1 , y_2 and X_1, X_2 represent the response and the explanatory variables within and outside the regimes, respectively. $\beta_1, \beta_2, \epsilon_1, \epsilon_2$ are the coefficients and the random errors within and outside the regime, reflecting the changing effects of the indicators (Anselin 2005).

ANOVA was performed to test the significance of the regime. The dummy variable was introduced into the model to interact with each explanatory variable. Then, a spatial regime model was developed using the dummy and the previously selected explanatory variables. High trash generation areas (trash volume $>0.1 \text{ L/m}^2$) were designated as the regimes to test how the explanatory variables influence trash generation differently within and outside the hotspot.

Spatial autocorrelation can be explained by Tobler's First Law of Geography in that 'near things are more related to each other than the distant ones' (Tobler 1970). Each geographic unit is not separable from its neighbors but is influenced by them, and therefore such influence should be addressed in the final model. Spatial lag was calculated for each geographic unit using a spatial weight matrix. The matrix defines which neighbors influence the unit and how such influence becomes weaker with the distance (Cliff & Ord 1982). Moran's I scatter plot was used to assess spatial autocorrelation. The spatial autocorrelation can be introduced into the OLS model as a separate spatial lag or spatial error. Based on spatial regime model, the spatial regression model (i.e. spatial lag model (Equation (2)) and spatial error model (Equation (3))) with regimes are shown below

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \rho W \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix} \quad (2)$$

$$\begin{aligned} \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix} &= \lambda W \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \\ &= \lambda W \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} - \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} \end{aligned} \quad (3)$$

where ρ is the coefficient for spatial lag, W is the spatial weights matrix, λ is the coefficient for spatial lag in the

error and u_1 , u are another error term after the spatial lag for error is introduced (Anselin 2005).

Lagrange multiplier diagnostics were performed to determine which is more statistically significant between spatial lag and spatial error. The diagnostics consist of two tests, the Lagrange Multiplier (LM) test and Robust Lagrange Multiplier (RLM) test. They were performed sequentially: LM is first calculated for the models and the one with significant p -value is chosen; if both of them have significant p -value for LM test, the RLM test is then performed and the statistically significant model is chosen as the final choice. If spatial lag is insufficient to cover the residual autocorrelation, more variables should be added into the model to address the unexplained spatial effects (Miron 1984).

Bidirectional spatial regression was applied to decide the final selection of indicators considering the spatial regimes and spatial dependency term. The objective was to address the residual autocorrelation in the initial model so as to fully explain the spatial pattern of trash generation. The variable in the output of the exploratory regression that greatly reduced residual autocorrelation was added to or removed from the model. Like stepwise regression, it started with the model obtained in the indicator selection process. The model with the best model performance was retained for the next iteration. Moran's I test for checking residual autocorrelation was used as the measure metric. When the metric was not statistically significant (p -value > 0.05), the process ceased. Each iteration resulted in five models for selection. The models improving the metric with practical meaning were included, or the variables with little influence on the metric were removed from the existing models. The quantile map for each variable was

plotted to see whether it had a similar spatial pattern to the residuals of the current model.

Akaike information criterion (AIC) was chosen as the measure for model performance. Adjusted coefficient of determination (R^2) is reliable for the fitting performance of OLS regression but inappropriate for both the spatial regime model and spatial lag model while AIC is suitable for all these models.

RESULTS AND DISCUSSION

Indicator selection

Exploratory regression reduced the original 282 variables obtained from census data and environmental variables to 37 variables. Best subsets regression as the initial model generated 100 models, consisting of five models with highest adjusted R^2 for each with the number of variables from 1 to 20. The change in R^2 and VIF was caused by the increase in number of indicators (Figure 3). The increase in indicators to over six did not improve the adjusted R^2 and the increase in indicators to over 12 led to a significant increase in VIF. Therefore, the number of indicators was selected to be six for best subsets regression to balance the model performance and simplicity.

The result of the initial OLS regression is shown as follows:

$$\begin{aligned} \text{Trash} = & -3.8227 + 0.0909 \times \text{HUOld} + 0.1601 \times \text{HUwSRO} \\ & + 0.2525 \times \text{HUwNoVeh} + 0.6122 \times \text{DistRiv} \\ & - 0.2087 \times \text{HHwSSI} + 0.1261 \times \text{HHwRI} \end{aligned} \quad (4)$$

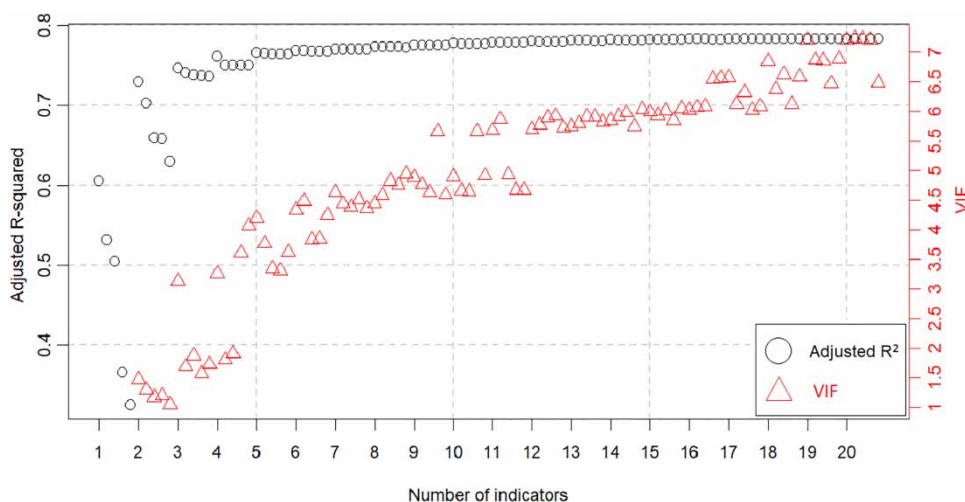


Figure 3 | Plot of R^2 and VIF versus number of indicators.

where HUOld is the percentage of housing units built before 1939, HUWSRO is the percentage of housing units with single room occupancy, HUwNoVeh is the percentage of housing units with no vehicle, DistRiv is the average distance to the rivers, HHwSSI is the percentage of households with supplemental security income, and HHwRI is the percentage of households with retirement income. HHwRI, HUwNoVeh, and HHwSSI represented population characteristics, which were related to retirement and income level. HUwNoVeh and HHwSSI were positively correlated with the indicator of poverty (i.e. the percentage of population under poverty level), with the correlation coefficients (r) of 0.82 and 0.55 respectively. This implied that places with low income tend to generate more trash. HHwRI seemed negatively related to trash generation, but the result of the OLS regression model showed its effects on trash generation varied over the city. HUOld and HUWSRO, related to housing characteristics, showed that more trash was generated from places with old or small-size houses. The DistRiv, representing an environmental factor, presented a tendency for less trash generation in places near rivers. The model showed high adjusted R^2 (0.77) with relatively small VIF. Each variable was statistically significant (p -value < 0.001).

Spatial regime model

Moran's I scatterplot (Figure 4) shows clear difference in spatial distribution of trash generation levels. All the high trash generating areas (i.e. red circles in the first quadrant) displayed a high-high clustering pattern, which means high trash generating areas were surrounded by high trash generating areas. Low trash generation areas (i.e. green circles in the third quadrant) showed a low-low clustering pattern, which means low trash generating areas were surrounded

by low trash generating areas. The spatial pattern of trash generation closely related to the absolute trash volume, showed three different spatial distributions of trash generation. The influential points for low or medium trash generation level slightly influence the value of the Moran's I while the influential points for high trash generation areas mainly had extreme trash generation values but did not influence the statistic.

To distinguish how the effects of the indicators varied across different trash levels, the high trash generation areas were selected as spatial regimes. The reasons for choosing the spatial regimes include the completely different spatial distribution pattern and different influential points type of the high trash level data compared to those from other levels. The result of the model with high trash generation areas as spatial regimes (Equation (5)) shows the significance of the spatial regime for high trash generation areas, proving that spatial heterogeneity exists within and outside the trash hotspot. The relatively high adjusted R^2 value ($=0.71$) indicates that the dummy variable for high trash level can explain the trash generation pattern. Moran's I on the residual of the spatial regime model (0.52 with p -value < 0.001) indicates strong spatial autocorrelation in the model.

$$\text{Trash} = 4.1106 + 19.3742 \text{ High_trash} \quad (5)$$

Spatial lag model with regimes

Spatial lag was introduced to remove spatial autocorrelation in the model. The results of Lagrange Multiplier diagnostics for the spatial regime model (Table 1) for determining the spatial dependency term showed that the LM values for both spatial error and spatial lag were statistically significant with p -value less than 0.001. The result of the RLM test showed that spatial lag was significant while spatial error was insignificant (p -value > 0.05). Therefore, spatial lag was selected as the spatial dependency term in the final model.

To address the residual autocorrelation in the spatial lag model with regimes using the initial six indicators in Equation (4), the bidirectional spatial regression process

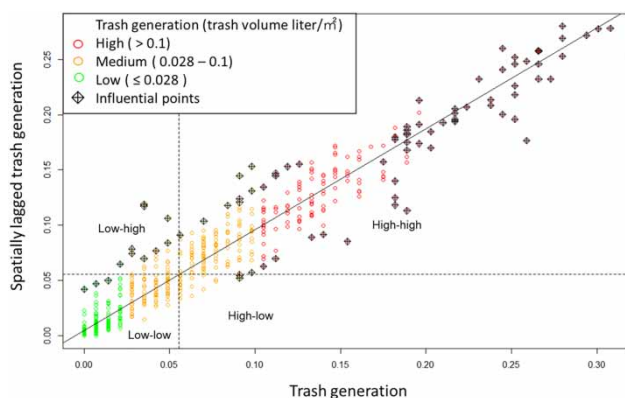


Figure 4 | Moran's I plot for trash volume.

Table 1 | Lagrange multiplier diagnostics on spatial regime model

Spatial dependency term	Lagrange Multiplier	p -value	Robust Lagrange Multiplier	p -value
Spatial error	234.00	<0.001	0.87606	>0.05
Spatial lag	501.25	<0.001	268.13	<0.001

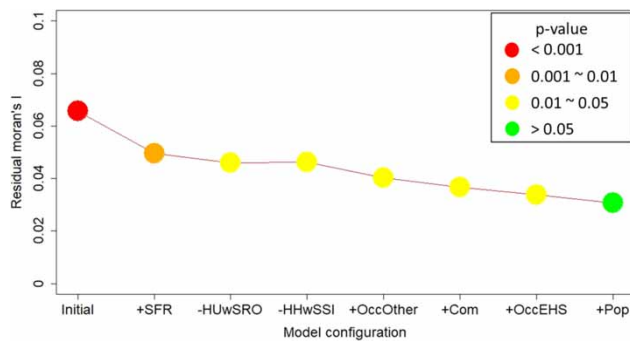


Figure 5 | Moran's I test in each step of bidirectional spatial regression with regimes.

removed two initial variables (HUwSRO, HHwSSI) and added five additional variables: OccEHS (percentage of employed civilian population occupied in educational, health and social services), OccOther (percentage of employed civilian population occupied for other services, single family residential (SFR) (percentage of lands for single family residual), Com (percentage of lands for commercial use) and Pop (total population). The selecting process stopped when p -value was greater than 0.5, as plotted in Figure 5. The figure also shows decreasing significance of Moran's I from extremely significant to insignificant. This proves that the modification of the independent variables improves the explanatory ability of the model. The correlation between each pair of variables is shown in Figure 6. Circles with darker color and bigger size show higher correlation. The correlation matrix

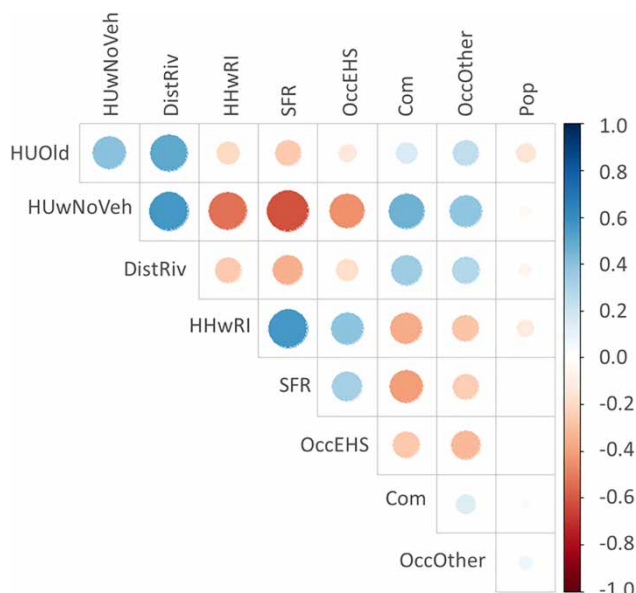


Figure 6 | Correlation matrix for each pair of explanatory variables.

confirmed that each pair of variables were not significantly correlated ($< |\pm 0.6|$), which means these variables were not collinear.

The quantile maps of the dependent variable and independent variables in the final model are shown in Figures 7 and 8, respectively, for comparison. OccEHS, SFR and Com presented an associated spatial pattern for trash generation. Less trash was generated in the areas with SFR, which was negatively related to population density and had relatively low imperviousness among all land uses, infiltrating runoff and preventing it from carrying trash to nearby catch basins. The areas dominated by people employed in educational, health and social services (OccEHS) tended to generate less trash, while the commercial areas tended to generate more trash. Total population and percentage of population employed in other services did not show a parallel pattern to trash generation distribution.

The resulting spatial lag model with regimes (Table 2) presented changes by adding the spatial lag and spatial regime. As for the spatial regime, the influence of each variable was different within and outside the regime. Some

Trash Generation in the City of Los Angeles (Trash)

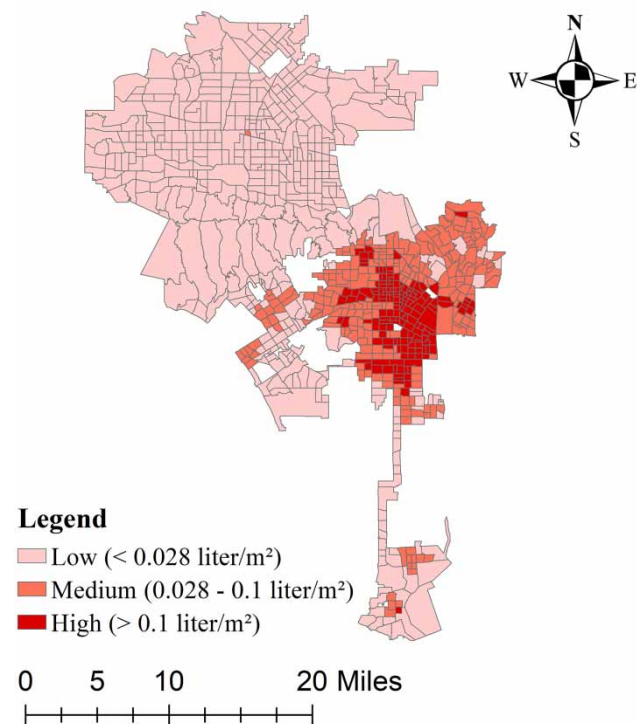


Figure 7 | Quantile maps for dependent variable (trash generation) in the final model.

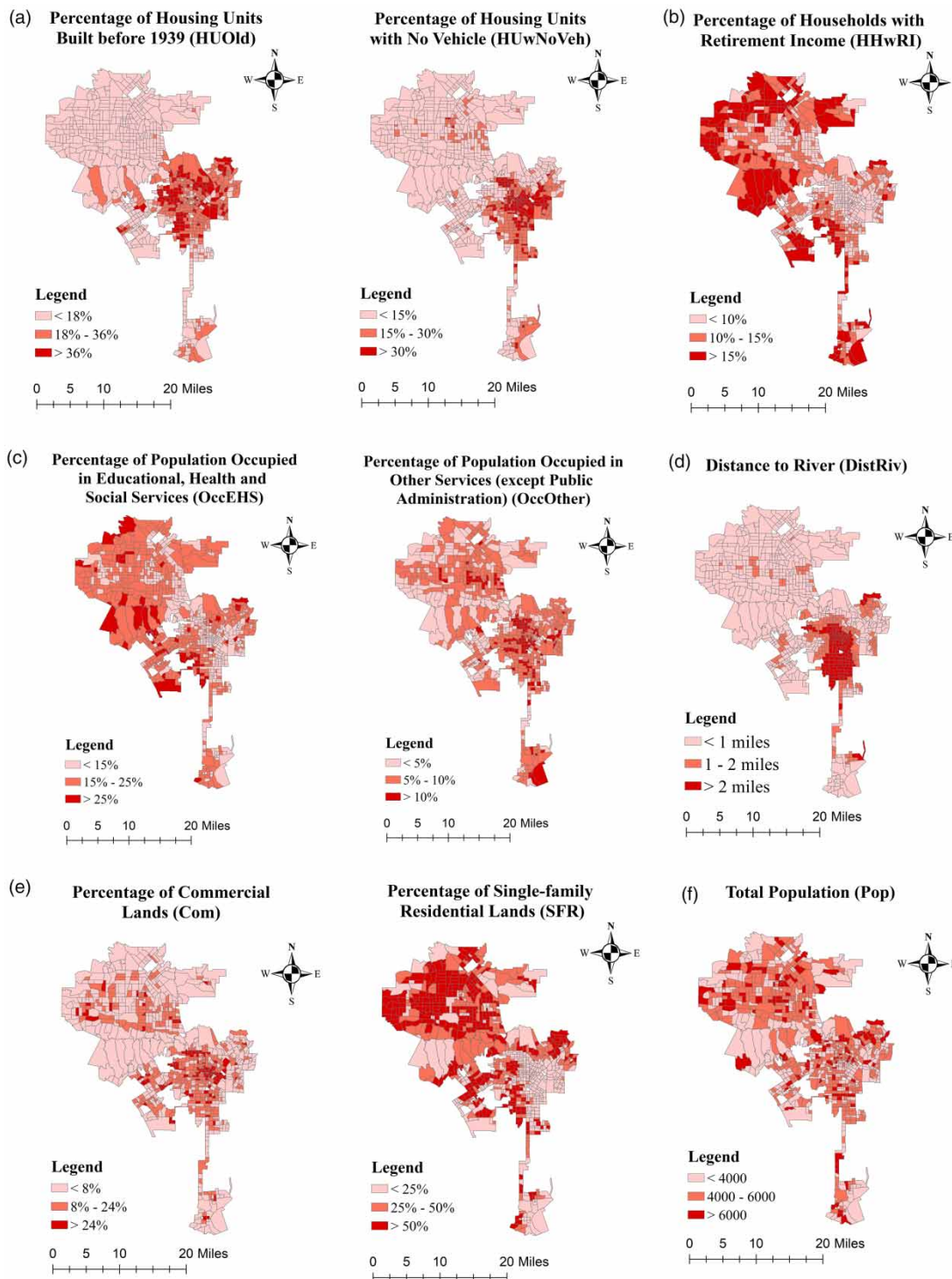


Figure 8 | Quantile maps for independent variables in the final model: (a) housing characteristics, (b) economic status, (c) occupational information, (d) environmental characteristics, (e) land use, (f) demographic characteristics.

variables were suitable for specific areas while others were suitable for the entire city, depending on their p -values. HUwNoVeh as an alternative to poverty was more related

to high trash generating areas than other areas with statistical significance. The spatial pattern of SFR was positively related with the high trash generation regime, while COM

Table 2 | The resulting spatial lag model with regimes for trash generation areas

Variable	Coefficient		p-value	
	High trash	Other	High trash	Other
Intercept	3.93114924	−0.12794866	<0.01	>0.05
HUOld	−0.04741788	0.02268323	<0.01	<0.001
HUwNoVeh	0.11709919	0.01807994	<0.001	>0.05
DistRiv	−0.09202984	−0.06095189	<0.05	<0.05
HHwRI	0.09163937	−0.03699983	>0.05	<0.05
SFR	0.05922656	−0.00099329	<0.001	>0.05
COM	0.00484178	0.03429776	>0.05	<0.001
OccEHS	−0.00735334	0.02113703	>0.05	>0.05
OccOther	0.06910985	0.01191855	>0.05	>0.05
POP	−0.18473513	0.02974533	>0.05	>0.05
lag.trash	0.75955		<0.001	

was positively related with other areas with statistical significance. HUOld and DistRiv were statistically significant indicators both within and outside the regime. As for spatial lag, some of the variables became less significant because the spatial lag was correlated with them. This means an indicator in the unit also had influence on its neighbors. The indicators of occupation (OccOther and OccEHS) and total population were not significant in the model, but they were necessary to explain part of the residual autocorrelation.

The comparison of the three models (Table 3) shows that the spatial lag model with regimes had lower AIC than the OLS regression and the spatial regime models. The Moran's I statistics became insignificant in the spatial lag model with regime. These statistics confirm that the spatial lag regression with regimes was the best model to explain trash generation without spatial autocorrelation.

The findings of this study show that human-environment interaction is a crucial factor in trash generation collected in catch basins. Besides developing efficient trash capturing systems, more attention should be paid to public behavior to improve the trash problem in urban runoff, which can affect the quality of receiving waters. Raising public

awareness to keep roads clean through advertising and education, especially in the low-income areas, will help to reduce the amount of trash to achieve the zero trash TMDL. Neighbors influence each other in littering habits and the environment of the neighborhood affects public attitude according to the 'Broken Window Theory'. High and dense population would cause more trash generation. Developed lands with impervious surfaces would contribute to runoff accumulation, carrying trash abandoned on the roads to the catch basins. Proper land use planning and reducing imperviousness will prevent runoff from being the vehicle for taking trash to the waters.

CONCLUSIONS

This study examined the effects of human activity on trash generation to achieve a zero trash TMDL and improve the quality of urban runoff. The spatial models using GIS provided a useful tool to identify major variables related to trash generation and to explain the human impacts on trash generation. The results of this study present the following conclusions:

- (1) Trash generation was influenced by the economic status of the population. Poverty-related variables were crucial factors in trash generation.
- (2) The occupation of residents was related to littering habits. The areas with educational, health and social services generated less trash, probably due to their awareness of environmental problems and their influence on the surrounding people.
- (3) Environmental factors affected trash accumulation. The places near rivers or dominated by single-family residential lands were less likely to fill the catch basins with trash abandoned on the roads, while commercial lands tended to generate more trash.
- (4) The influences of each indicator were not specific to each unit but were also placed on the neighborhood.

This study provided trash generation from a socio-economic perspective. It also considers the spatial effects in the analysis. The approach and the methods in this study can be applied to other areas to examine the major drivers for trash generation for its proper management. The findings from this study can be used for trash prediction. An extended study can be conducted with updated data as the trash data are not routinely collected. Future work can be focused on the time series analysis of trash generation.

Table 3 | Comparison of OLS, spatial regime model and spatial lag model with regimes

	OLS	Spatial regime	Spatial lag with regimes
AIC	4848.645	4305.2	3590.5
Moran's I	0.4491 (p-value < 0.001)	0.5186 (p-value < 0.001)	0.0308 (p-value > 0.05)

ACKNOWLEDGEMENT

This study was conducted in the support of UCLA-Cross-disciplinary Scholars in Science and Technology (CSST) program.

REFERENCES

- Alam, M. Z., Anwar, A. H. M. F. & Heitz, A. 2017 Stormwater solids removal characteristics of a catch basin insert using geotextile. *Science of The Total Environment* **618**, 1054–1063.
- Anselin, L. 2005 *Spatial Regression Analysis in R A Workbook*. University of Illinois, Urbana, IL, USA.
- Bandara, N. J. G. J., Hettiaratchi, J. P. A., Wirasinghe, S. C. & Pilapiiya, S. 2007 Relation of waste generation and composition to socio-economic factors: a case study. *Environmental Monitoring and Assessment* **135**, 31–39.
- California Regional Water Quality Control Board 2007 *Trash Total Maximum Daily Loads for the Los Angeles River Watershed*. Available from: https://www.waterboards.ca.gov/losangeles/board_decisions/basin_plan_amendments/technical_documents/2007-012/09_0723/L.%20A.%20River%20Trash%20TMDL_Final%20%20Staff%20Report_August%209,%202007.pdf.
- City of Los Angeles 2002 *High Trash-Generation Areas and Control Measures*. Available from: http://www.lastormwater.org/wp-content/files_mf/trash_gen_study.pdf (accessed 10 July 2017)
- Cliff, A. D. & Ord, J. K. 1982 Spatial processes: models & applications. *Quarterly Review of Biology* **57** (2), 236.
- Conley, G., Beck, N., Riihimäki, C. A. & Hoke, C. 2019 Improving urban trash reduction tracking with spatially distributed Bayesian uncertainty estimates. *Computers, Environment and Urban Systems* **77**, 101344.
- Ertur, C., Gallo, J. L. & Baumont, C. 2006 The European regional convergence process, 1980–1995: do spatial regimes and spatial dependence matter? *Econometrics* **29** (1), 3–34.
- Favot, M. & Grassetti, L. 2017 E-waste collection in Italy: results from an exploratory analysis. *Waste Management* **37**, 222–231.
- Furnival, G. M. 1971 All possible regressions with less computation. *Technometrics* **13** (2), 403–408.
- Ghermandi, A. 2016 Analysis of intensity and spatial patterns of public use in natural treatment systems using geotagged photos from social media. *Water Research* **105**, 297–304.
- Kang, J.-H., Debats, S. R. & Stenstrom, M. K. 2009 Stormwater management using street sweeping. *Journal of Environmental Engineering ASCE* **135** (7), 479–489.
- Kannangara, M., Dua, R., Ahmadi, L. & Bensebaa, F. 2018 Modeling and prediction of regional municipal solid waste generation and diversion in Canada using machine learning approaches. *Waste Management* **74**, 3–15.
- Keser, S., Duzgun, S. & Aksoy, A. 2012 Application of spatial and non-spatial data analysis in determination of the factors that impact municipal solid waste generation rates in Turkey. *Waste Management* **32** (3), 359–371.
- Kim, L.-H., Kayhanian, M. & Stenstrom, M. K. 2004 Event mean concentrations of litter from highways during storms. *Science of the Total Environment* **330** (1–3), 101–113.
- Kostarelos, K., Khan, E., Callipo, N., Velasquez, J. & Graves, D. 2011 Field study of catch basin inserts for the removal of pollutants from urban runoff. *Water Resources Management* **25**, 1205–1217.
- Lau, S. L., Khan, E. & Stenstrom, M. K. 2001 Catch basin inserts to reduce pollution from stormwater. *Water Science & Technology* **44**, 23–34.
- Lebersorger, S. & Beigl, P. 2011 Municipal solid waste generation in municipalities: quantifying impacts of household structure, commercial waste and domestic fuel. *Waste Management* **31** (9–10), 1907–1915.
- Liu, A., Hong, N., Zhu, P. & Guan, Y. 2019 Characterizing petroleum hydrocarbons deposited on road surfaces in urban environments. *Science of the Total Environment* **653**, 589–596.
- McCuen, R. H., Sharkey, M. & Walker, A. 2014 Trash deposition in rivers and streams: guidelines for prediction of counts, volumes, and masses. *Journal of Hydrologic Engineering* **19**, 1–10.
- Miron, J. 1984 *Spatial Autocorrelation in Regression Analysis: A Beginner's Guide*. Springer, Dordrecht, The Netherlands.
- Moore, S. L., G. D., Carreon, M., Weisberg, S. B. & Leecaster, M. K. 2001 Composition and distribution of beach debris in Orange County, California. *Marine Pollution Bulletin* **42** (3), 241–245.
- Núñez, P., García, A., Mazarrasa, I., Juanes, J. A., Abascal, A. J., Méndez, F. & Medina, R. 2019 A methodology to assess the probability of marine litter accumulation in estuaries. *Marine Pollution Bulletin* **144**, 309–324.
- Ramachandra, T. V., Bharath, H. A., Kulkarni, G. & Han, S. S. 2018 Municipal solid waste: generation, composition and GHG emissions in Bangalore, India. *Renewable and Sustainable Energy Reviews* **82**, 1122–1136.
- Roales-Nieto, J. G. 1988 A behavioral community program for litter control. *Journal of Community Psychology* **16** (2), 107–118.
- Tobler, W. 1970 A computer movie simulating urban growth in the Detroit region. *Economic Geography* **46**, 234–240.
- US Senate 2002 *Federal Water Pollution Control Act*. Available from: <https://www.epa.gov/sites/production/files/2017-08/documents/federal-water-pollution-control-act-508full.pdf>.