Original Article

WM&R

Waste Management & Research
2018, Vol. 36(11) 1049–1060

DOI: 10.1177/0734242X18804031
journals.sagepub.com/home/wmr
SAGE

Identifying the main physical and socioeconomic drivers of illegal landfills in the Canary Islands

L Quesada-Ruiz, V Rodriguez-Galiano and R Jordá-Borrell

Abstract

The management of disposed waste in illegal landfills (ILs) is a significant problem in contemporary societies due to respective hazards for the environment and human health. This paper presents a characterisation of ILs on the islands of La Palma (LP) and Gran Canaria (GC) based on multivariable statistical analysis. Inspection of numerous sites on both islands revealed a total of 153 and 286 ILs on LP and GC, respectively. A geospatial database was created composed of different potentially explanatory features of different typology (177): waste type, control and vigilance, socioeconomic, accessibility, distance to elements of interest, visibility and physical. The degree of association between the explanatory features and the occurrence of ILs was analysed with the support of exploratory statistics and the multivariable analysis techniques of principal component analysis (PCA) and binary logistic regression (LR). PCA explained 82.34% and 81.83% of total data variance in LP and GC, respectively, considering 7 and 6 components (Kaiser–Mayer–Olkin; LP: 0.715; GC: 0.711). The LR models for LP and GC had an overall accuracy of 93.5% and 92.5%. In LP and GC, 6 of 23 features and 9 of 21 features were, respectively, selected. The features most associated with the occurrence of ILs were: in LP, building density, distance to agricultural spaces and distance to green zones; in GC, the industrial activity indicator, density of ground use transition to artificial covers, density of greenhouses and distance to communication routes.

Keywords

Waste, illegal landfills, GIS, feature selection, logistic regression, principal component analysis, Canary Islands

Received 3rd April 2018, accepted 4th September 2018 by Associate Editor Dimitris Dermatas.

Introduction

The proliferation of illegal landfills (ILs) has a negative impact, especially for ecologically sensitive areas and human health, and is a major environmental problem for many countries. A report produced by the European Union Action to Fight Environmental Crime (EFFACE) indicated that the total number of ILs in the European Union (EU) is 12,628, with a total of 2,871,186 tons of waste (Watkins, 2015). The generation of ILs causes different impacts: deterioration of local landscapes (Matsumoto and Takeuchi, 2011), increased air pollution (Bridges et al., 2000), contamination of aquifers (Hafeez et al., 2016; Monteiro Santos et al., 2006) or increased risk for human health. ILs also have repercussions for local economies, reducing return from tourism operations and entailing high costs for repair and cleaning in affected areas, which can account for up to 30% of some local government budgets (Calò and Parise, 2009; Ichinose and Yamamoto, 2011; Jones, 2008; Notarnicola et al., 2004). To reduce the effects of ILs on the territory, restrictive laws have been introduced which regulate waste management and the penalties for landfill-related offences. The directive on waste and repealing certain Directives (European Commission, 2008) defines a landfill as being a waste elimination site where waste is into or on to land. Although that directive does not define an IL, the directive on environmental responsibility with respect to prevention and repair of environmental damage (European Commission, 2004) establishes that uncontrolled waste should be managed, as well as its collection, transport, recovery and elimination. It also demands measures to prevent and evaluate environmental damage in order to plan the respective remediation. On the other hand, the autonomous communities of the Spanish state altogether consider ILs to be land affected by waste deposits with no type of management or control for a period of more than two years and an area of more than 2000 m².

One of the main challenges of studying ILs is identification of the physical or socioeconomic features associated with their appearance (Apostol and Mihai, 2011; Biotto et al., 2009; Doak et al., 2007; Silvestri and Omri, 2008; Tasaki et al., 2004, 2007). A

Department of Physical Geography and Regional Geographical Analysis, University of Seville, Spain

Corresponding author:
Quesada-Ruiz L, Department of Physical Geography and Regional Geographical Analysis, University of Seville, Doña María De Padilla, S/N, Seville 41004, Spain.
Email: lquesada@us.es

better understanding of the specific features that influence the occurrence of ILs improves the waste management and reduces the impact linked to their appearance (Glanville and Chang, 2015; Matos et al., 2012). Those features may be different depending on the nature of the territories. Hence, for example, Apostol and Mihai (2011) identified the closeness of rural areas as being a major factor in the appearance of ILs in Romania, while Biotto et al. (2009) found that ILs in Italy's Veneto region are controlled by industrial activity and closeness to authorised landfills. Jordá-Borrel et al. (2014) concluded that in southern Spain socioeconomic features such as per capita income are more important; for Matos et al. (2012) in Slovenia it was the closeness of roads and for Tasaki et al. (2007) in Japan the topography. Silvestri and Omri (2008) in turn linked the occurrence of ILs to vegetation zones.

The characterisation of factors affecting the appearance of ILs normally entails consideration of numerous features which may potentially be involved in this process, giving way to a space with complex features. In this regard, a first analysis of the basic statistics may bring us closer to the relationship of the features most associated with IL occurrence (Apostol and Mihai, 2011; Tasaki et al., 2004). Also, by using more advanced statistical techniques such as principal component analysis (PCA), the degree of association between features can be discerned. The use of logistic regression (LR) can allow the establishment of causality relationships between the independent features and IL occurrence (Lucendo-Monedero et al., 2015) and determine the predictive, confounding or modifying role of the effect of each feature, distinguishing direct causality relationships and relationships resulting from chance. In that regard, the predictive features are precursors of IL occurrence, maintaining a real relation to it. Confounding features are those that apparently predict IL occurrence due to their statistical link to the phenomenon but which do not support a real relation to same. That may be due to their high correlation with other independent features. Finally, the features that modify the effect of IL occurrence are those whose interaction with other features may increase or decrease the respective effect. Hence, ascertainment of those feature types is vital for understanding the processes leading to the appearance of ILs in a given territory.

The features associated with IL appearance in the ultra-peripheral region of the Canary Islands have not been studied to date. This region has different characteristics which make waste management difficult and might be associated with the proliferation of ILs: (i) high population densities, with the islands of La Palma (LP) and Gran Canaria (GC) having 122.16 inhabitants/km² and 543.45 inhabitants/km², respectively; (ii) mass tourism (Hernández Rodríguez, 2006), receiving 167,838 and 4,223,679 tourists in 2016 in LP and GC, respectively (http://estadisticas.tourspain.es); (iii) significant urban growth linked to the early twenty-first century real estate boom (Fernández-Tabales and Cruz, 2013); (iv) intense farming activity which can be a major source of waste, especially plastic waste (Hernández Torres, 2003); (v) limited number of waste management and treatment infrastructures; and (vi) the lack of citizen awareness.

This study aims to identify the features associated with the appearance of ILs and to establish causal relationships between the presence and absence of ILs on LP and GC. This paper thus presents a characterisation and analysis of ILs on both islands based on exploratory analysis and multivariable analysis with the aim of understanding the features that drive the proliferation of ILs in order to improve damage repair and prevention policies. The specific objectives are: (i) to characterise the spatial patterns of coincidence between IL occurrence and other socioeconomic and physiographic features; (ii) to assess the degree of association between those features and IL occurrence; and finally (iii) to identify causality relationships between the explanatory features and IL occurrence, as well as to reduce the feature space, selecting the most important features and eliminating those that can be confounding or modifying.

Methodology

Data

The IL localisation process consisted of three stages: (i) identification of potential ILs by means of photo-interpretation of orthophotos with spatial resolution of 0.5 m for the years 2012 and 2015; (ii) field inspection of 215 (LP) and 387 (GC) potential sites; (iii) screening of illegal ILs with deposits less than two years old, ultimately obtaining 153 (LP) and 286 (GC) IL sites (Figure 1). For each IL site, information referring to waste type, degree of accessibility, fencing, access control and existence of dissuasive measures was compiled. To apply LR and for the purpose of generating an IL occurrence probability model, the sample was complemented by including sites not affected by the existence of ILs, following the methodology described by Carranza et al. (2008). To that end, a sampling was applied, meeting the following conditions: (i) randomised sites based on point analysis of patterns (Boots and Getis, 1988; Diggle, 1983), (ii) dissimilitude in multivariable information with the IL sites, (iii) distances of more than 1594 m to IL sites based on the analysis of closer distances between ILs, and (iv) equal number of negative and positive occurrence zones (e.g. Breslow and Cain, 1988; Schill et al., 1993). The positive and negative occurrence zones were coded as 1s and 0s, respectively, resulting in a total of 306 (LP) and 572 (GC) cases.

Following Alexakis and Sarris (2013), Doak et al. (2007), Biotto et al. (2009) and Tasaki et al. (2004, 2007), a series of spatial features of different typology was considered: socioeconomic features, such as per capita income, population and economic, industrial and tourism activity indicators; management-related, such as waste type, degree of access, accessibility, security and control; and finally terrain features, such as elevation and slope. Based on this initial group of features, a subgroup of derived features was obtained by applying different GIS analysis procedures (Akbari et al., 2008; Demesouka et al., 2014; Kontos et al., 2005; Şener and Karag, 2011; Şener et al., 2010; Uyan, 2014). New features were thus extracted (Table 1) based on interpolation of the socioeconomic information broken down by population centres for the rest of the

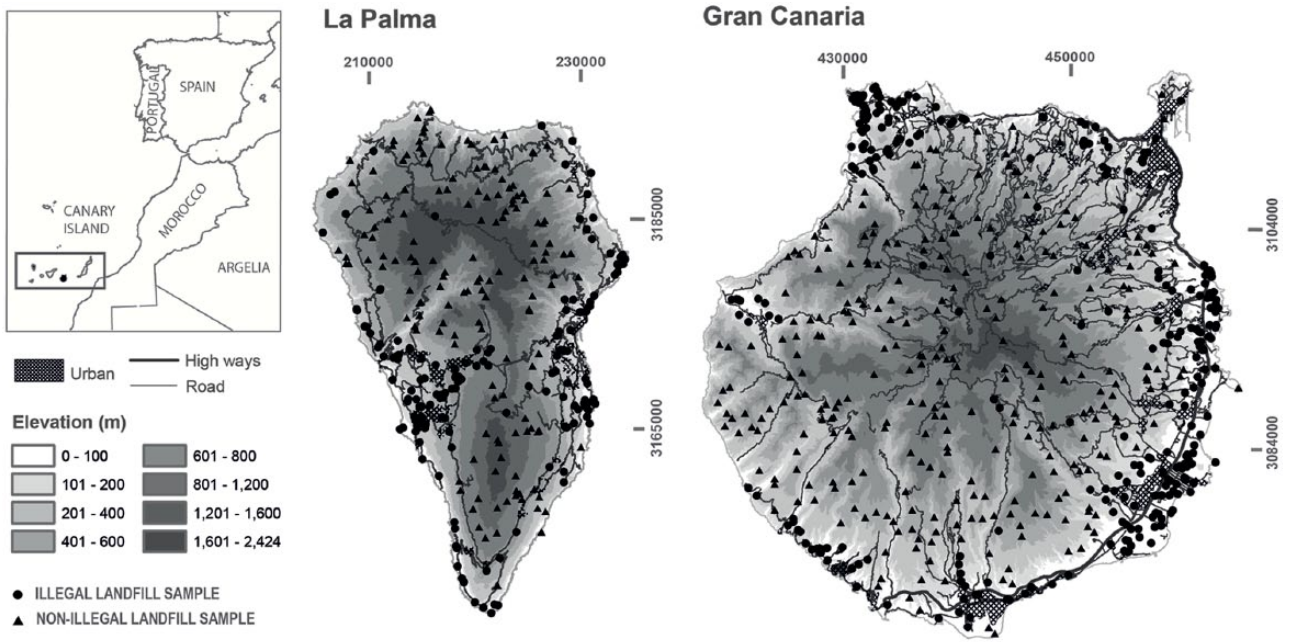

Figure 1. Study area.

territory, application of Euclidean distance (ED) criteria between the location of the ILs and the features of interest (Biotto et al., 2009; Jordá-Borrell et al., 2014; Tasaki et al., 2007), calculation of densities of elements of interest by applying kernel functions (Silverman, 1986) and other search functions based on distance to a given radius (250 m, 500 m, 1500 m) and the extraction of features associated with ground occupation, considering both the calculation of densities and the distance to a given ground use. Finally, the normalised difference vegetation index (NDVI) was calculated using a SPOT-5 image dated 31 August 2014. Each feature was standardised, rastered and resampled at a spatial resolution of 10 m. The values for all the aforementioned features were extracted for the positive and negative IL occurrence locations.

Multivariable statistical analysis: PCA and LR

The PCA and LR were done in the SPSS 24.0 software environment. PCA was solely applied to the cases of positive occurrence. It was used to determine the association between different features. The multivariable normality of the features and their interrelationship were verified by means of the Kaiser–Meyer–Olkin measure of sampling adequacy (LP: 0.715; GC: 0.711) and Barlett's sphericity test (LP: 2922.87; GC: 9013.14). The ratio of 10 cases per feature chosen for the PCA was respected. The factors were rotated using the quartimax method (Hair et al., 2010). A minimum presence of two features per factor was assured along with correlations equal to or above 0.40 for each feature.

LR was used to evaluate the features presumably associated with IL proliferation and to learn their role: predictive, confounding or modifying. The predictive role was evaluated based on magnitude and the sign of the ratios between the features and IL occurrence. The confounding role was assessed by evaluating the strength of association between the independent feature and IL occurrence when the latter is introduced in the LR equation. The modifying features of the effect of IL occurrence were recognised by evaluating the interaction of one feature with another after introduction of the first. Statistically significant features at 95% confidence level were considered. LR was also applied to the negative occurrence zones and positive occurrence zones. The parsimony principle was assured, reducing the features space based on the forward step inclusion method. The number of independent features needed to achieve the best discrimination between positive and negative occurrence zones was thus reduced. Construction of the binary LR model for each island was achieved based on the linear combination of the features. The method was evaluated based on the percentage of variance explained by the subgroup of features: R squared of Cox & Snell (LP: 0.629; GC: 0.602) and R squared of Nagelkerke (LP: 0.845; GC: 0.802). Also, the degree of precision of the IL prediction was evaluated by means of the confusion matrix and the overall accuracy.

Results and discussion

Characterisation of waste from the ILs of LP and GC

As can be seen in Figure 2, waste from construction and demolition is an environmental issue that has not been addressed in both islands, which could have negative repercussions for their tourism image (Sebenik, 1994; Smrekar, 2011; US Environmental Protection Agency, 1998). Indeed, such waste was the typology with the greatest presence in both LP (43%) (Figure 2(a)) and GC (52%) (Figure 2(b)), followed by waste derived from mining and extraction activity, which respectively accounted for 28% and 21%

Table 1. Selected features.

Short name	Long name	Units	Short name	Long name	Units
C_TYPA	Cadastral plot type	Unitless	E_PATH	ED to path	m
D_ARCC	Impervious cover transition density (1990–2012)	km^{-2}	E_PZ35	ED to pit (35 m kernel)	m
D_BUIL	Building density	km^{-2}	E_RAVI	ED to cliffs	m
D_CC00	Cover transition density between (2000-2006)	km^{-2}	E_REEQ	ED to leisure equipment	m
D_IELE	Element of interest density	km^{-2}	E_RELI	ED to religious infrastructures	m
D_GRHO	Greenhouse density	km^{-2}	E_ROAD	ED to roads	m
D_ROAD	Road density	km^{-2}	E_RUSE	ED to rural settlement	m
D_RUIN	Ruins density	km^{-2}	E_SEGI	ED to security infrastructures	m
E_COAS	ED to coast	m	E_SPOI	ED to sport infrastructures	m
E_COME	ED to commercial areas	m	E_TELI	ED to telecommunication infrastructures	m
E_CULI	ED to cultural infrastructures	m	E_TOHO	ED to tourist housing	m
E_EDEI	ED to education infrastructures	m	E_URAR	ED to urban areas	m
E_ENEI	ED to energy infrastructures	m	E_WAYS	ED to paths	m
E_EXAR	ED to extractive areas	m	P_MDTG	Altitude	Unitless
E_GRZO	ED to green zones	m	P_NDVI	NDVI index	%
E_HEIN	ED to health infrastructures	m	P _SLPE	Slope	km^{-2}
E_HIGH	ED to highways	m	H_PODE	Population density	€/m^2
E_HOST	ED to hostel	m	H_IBIR	Rural property tax	€/inhabitant
E_IELE	ED to features of interest	m	H_INPC	Income per capita	%
E_INAR	ED to industrial areas	m	I_COMA	Wholesale trade index	%
E_MONU	ED to monuments	m	I_ECAC	Economic activity index	%
E_MURC	ED to municipal recycling centre	m	I_TURI	Tourist index	%
E_NACL	ED to nature classroom	m	I_INDU	Industrial index	%
E_PRAR	ED to protected areas	m			

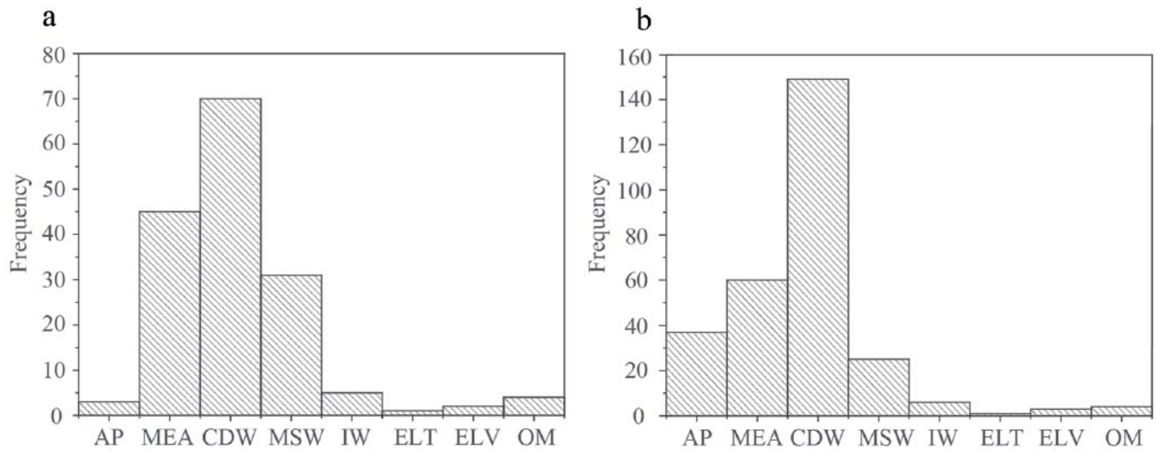

Figure 2. Types of waste in La Palma (a) and in Gran Canaria (b). Agricultural plastic (AP), mining and extraction activity (MEA), construction and demolition waste (CDW), municipal solid waste (MSW), industrial waste (IW), end-of-life vehicles (ELV), end-of-life tyres (ELT), organic matter (OM).

of total located ILs. Municipal solid waste was higher on LP (18%) than GC (10%), probably due to the less developed waste management systems on LP and the distance to the municipal solid waste treatment centres. Predictably, the lack of greenhouse crops on LP meant that plastic waste ILs were not generated; however, they did occur on GC (13%), with its extensive greenhouse occupation. Low industrial activity on both islands meant less production of industrial waste and ILs of that typology. It is noteworthy that ILs with organic material and end-of-life tyres and vehicles were sporadic and apparently did not pose a problem. On the other hand, 65% and 71% of the IL sites notably presented some type of secondary waste (LP: Figure 3(a); GC: Figure 4(a)), mostly associated with construction and demolition (LP: 32%; GC: 30%), earth moving (LP: 26%; GC: 12%) and municipal solid waste (LP: 16; GC: 14%).

Some 80% and 81% of the ILs presented no control or vigilance element in their surrounding area (LP: Figure 3(b); GC:

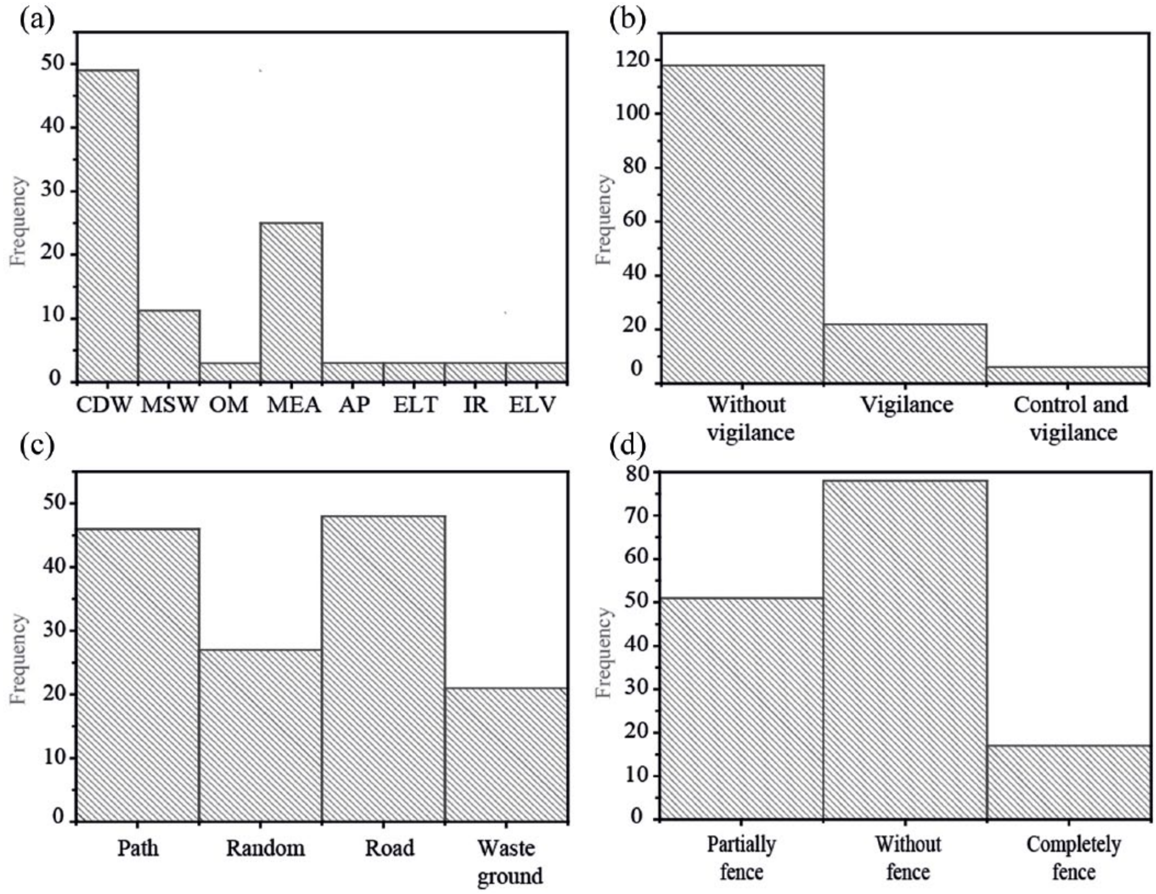

Figure 3. Histograms for La Palma. (a) Secondary type of waste. (b) Surveillance and security. (c) Type of access. (d) Fenced enclosures. Agricultural plastic (AP), construction and demolition waste (CDW), mining and extraction activity (MEA), municipal solid waste (MSW), industrial waste (IW), end-of-life vehicles (ELV), end-of-life tyres (ELT), organic matter (OM).

Figure 4(b)). Furthermore, no type of dissuasive measures were found in more than 95% of the ILs on both islands. That may be directly related to the low protection of the plots surroundings and the lack of dissuasive measures for the occurrence of ILs. Access to the IL sites (LP: Figure 3(c); GC: Figure 4(c)) from roads represented solely 33% and 34%, requiring more attention and with random, unpaved path and across-open-ground accesses standing out altogether. Likewise, there was limited fencing of plots in 90% and 78% of the sites visited. (LP: Figure 3(d); GC: Figure 4(d)).

Figure 5 shows the distribution by quartiles of the values of the IL sample for each island and each feature selected. In the subgroup of distance to ground use features (Figure 5(a)) it was seen that the ILs on both islands were present mainly in rural environments, as 100% of the ILs were situated less than 0.2 km and 0.6 km from agricultural land. 50% of the cases of located ILs were less than 0.8 km from an urban centre, concentrating in sites close to the outskirts of urban centres. Natural protected areas and green zones were a lesser occurrence, probably due to the respective terrain, low accessibility and vigilance. However, 25% of the ILs were found within or relatively close to such spaces. Despite the low industrial activity on both islands, 25% of the ILs were situated less than 1 km from land used by industry, which may be due to the specialisation of industrial poles in the construction or manufacturing sectors or their localisation in relatively non-cohesive spaces and with empty lots. Some 75% of the ILs on GC were situated less than 500 m from covered areas with high densities of ground use transitions for the period between 2000 and 2006. Possibly the highest occurrence of ILs may be situated in areas with more urban growth. Also, the rural environments with high greenhouse densities (between 15 and 40 greenhouses in a radius of 0.5 km) accounted for the location of 50% of the ILs on GC. Hence, the agricultural spaces producing large amounts of plastic could behave as potential IL generators. Conversely, the spaces with high density of elements of tourism and cultural interest maintained no relationship with IL occurrence, which may be due to higher people flow, control and vigilance (Figure 5(b)).

The spaces close to ILs on both islands presented a high density of communication routes in their surrounding area. All the ILs were found within 0.2 km from roads and 75% within 0.4 km from a major secondary road (Figure 5(c)). IL occurrence may consequently increase in places with more connection to the road network. Also, 25% of the ILs on GC were less than 0.6 km from the national motorway. The margins of this axis of communication may therefore attract IL occurrence by presenting both high industrial and construction activity, and empty lots and abandoned farmland with little vigilance and easy access. Some 25% of the ILs were situated less than 0.3 km from paths and unpaved

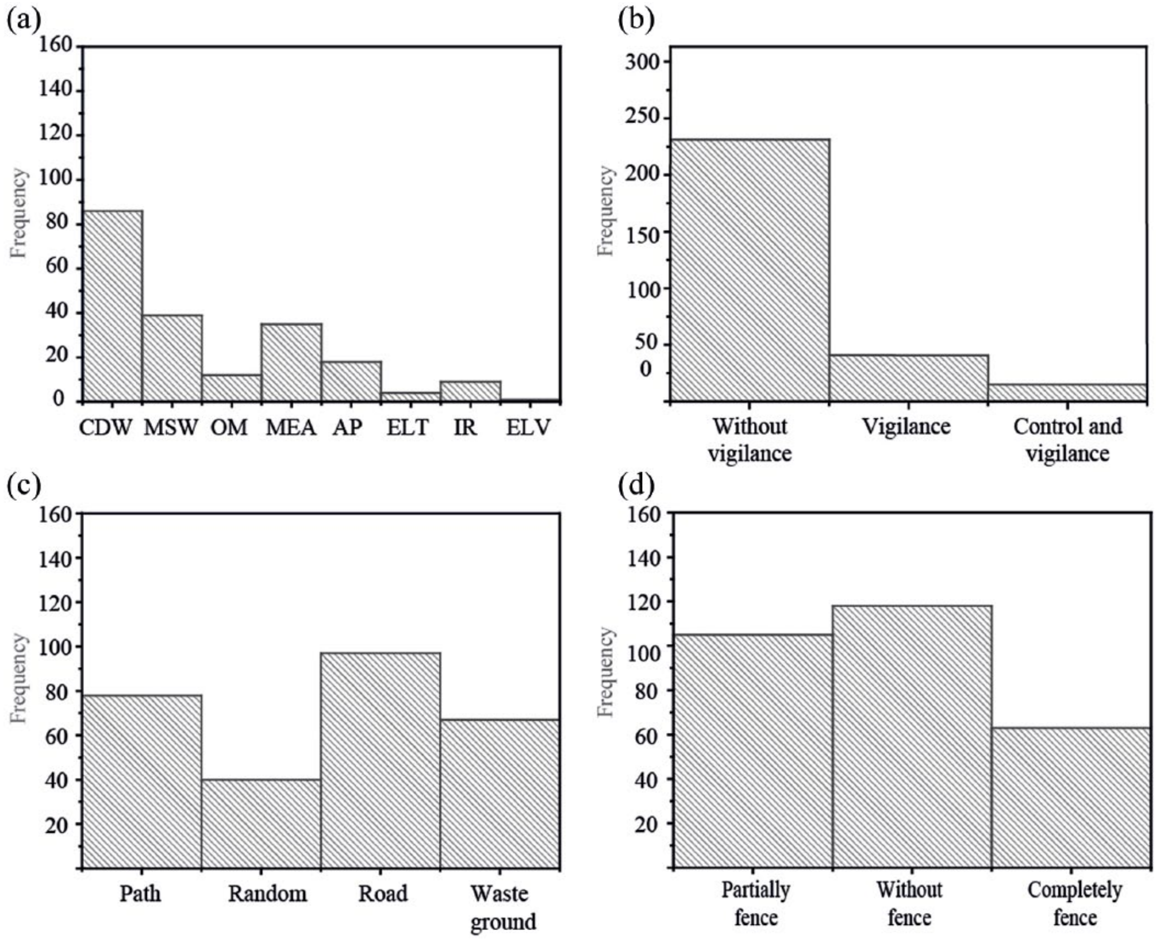

Figure 4. Histograms for Gran Canaria. (a) Secondary type of waste. (b) Surveillance and security. (c) Type of access. (d) Fenced enclosures. Agricultural plastic (AP), construction and demolition waste (AP), mining and extraction activity (MEA), municipal solid waste (MSW), industrial waste (IW), end-of-life vehicles (ELV), end-of-life tyres (ELT), organic matter (OM).

trails, which may be explained by the high number of that road category on LP.

The features for distance to infrastructures of interest (Figure 5(d)) showed that 75% of the ILs on LP were situated less than 1 km from tourism lodgings, while only 25% of the ILs on GC were located less than 1.1 km from them; that disparity may be due to the higher concentration of tourism lodging close to urban centres on the island of LP versus GC. Some 25% of the ILs on LP and GC were found less than 0.7 km from a sports infrastructure; that would be due to the fact that such infrastructures are situated on the outskirts of urban centres.

The physiographic features (Figure 5(e)) showed the relationship between potential IL occurrence and distance to ravines. Some 50% of the ILs on LP and GC were found less than 0.2 km from a ravine. Ravines might thus be attractive because they are hidden and lack vigilance measures in the surrounding area or because they are close to or crossed by communication routes. Some 50% of the ILs on both islands were located less than 1.4 km from the coastline, largely explained by the terrain and the higher number of urban settlements.

Places with better socioeconomic indicators are more favourable for the appearance of ILs (Figure 5(f)), with the influence of those features heightened on GC. For both islands it can be seen that more than 50% of ILs are situated in non-touristic areas, of less importance for the local economy and therefore with fewer preventive and/or persuasive measures. ILs were notably more frequent in spaces with more economic, industrial and retail trade activity, which may be explained by higher waste generation and the presence of urban settlements in such areas.

As a previous step before PCA and LR, ES enabled an approach to the behaviour of the features and the existence of spatial patterns between certain values of the independent features and IL occurrence. Its application nevertheless did not enable evaluation of the existence of statistically significant associations between the features and the presence or absence of ILs.

Principal component analysis

The PCA allowed identifying the associations between explanatory features and the appearance of ILs (Jordá-Borrell et al., 2014; Tasaki et al., 2007). The PCA explained 82.34% of the total variance in LP (Table 2). Seven components were obtained; the first three accumulated 52.22% of the data variance. The first, second and third components thus explained 28.36%, 14.22% and 10.64%, respectively. From the fourth component, the variance explained was notably lower, 9.77% (fourth), 7.87% (fifth),

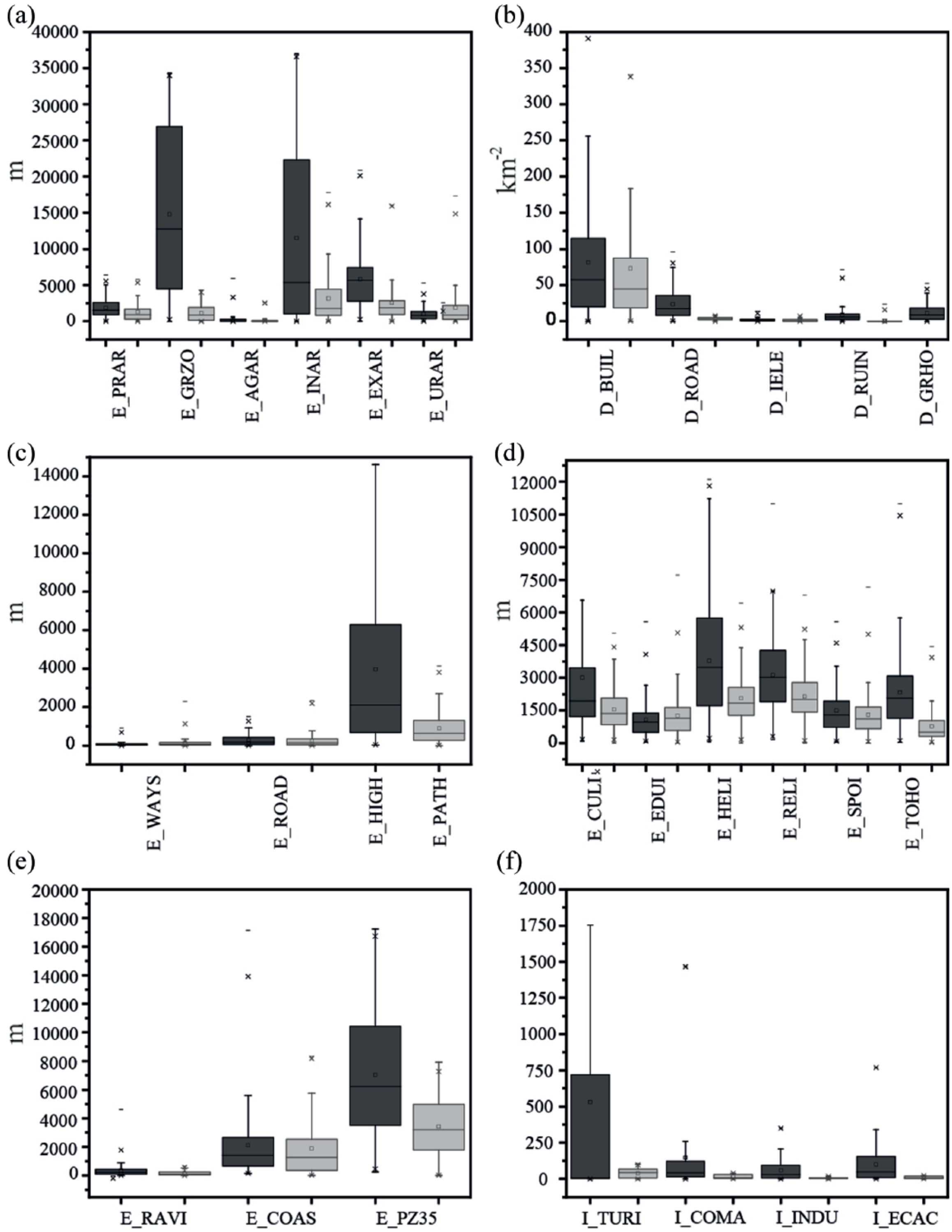

Figure 5. Boxplots. (a) Distance to land use. (b) Density of elements of interest. (c) Distance to road networks. (d) Distance to infrastructures of interest. (e) Distance to physiographic elements. (f) Socioeconomic indexes. GC: dark grey; LP: grey.

6.48% (sixth) and 5.00% (seventh). In the case of GC, the PCA explained 81.83% of the total variance (Table 3) and six components were obtained; the first three accumulated 52.82% of the data variance, with the first, second and third components explaining 32.55%, 13.16% and 12.83%, respectively. From the fourth component, the variance explained was much lower – 11.51% (fourth), 6.53 % (fifth) and 5.29% (sixth). In the case of LP, the first component grouped features that were associated with population. It showed a positive relationship for the density of constructions, population density, density of elements of interest, density of communication routes and the industrial indicator. Each of those characteristics would increase its value when the value of one of them increases. That may directly connect higher population and urban and social dynamism to the occurrence of

Table 2. Principal component analysis model for La Palma.

Components (1st level of rank)	Initial values	Weight %	Features (2nd level of rank)	Weight (correlation)	Rotation sums of squared loadings
				% Variance	% Accumulated
C1 Settlement	6.807	28.36	D_BUIL	0.88	28.36
			H_PODE	0.82	
			D_IELE	0.75	
			D_ROAD	0.63	
			I_INDU	0.62	
C2 Distance to cultural areas and green zones	3.412	14.22	E_PATH	0.93	42.58
			E_IELE	0.93	
			E_GRZO	0.69	
C3 Socioeconomic	2.554	10.64	I_ECAC	0.92	53.22
			H_INPC	0.84	
			I_ETME	0.79	
C4 Distance to element of interest	2.346	9.77	E_EXAR	0.88	62.99
			E_URAR	0.83	
			E_INAR	0.72	
			E_ROAD	0.69	
C5 Distance to rural areas	1.889	7.87	E_AGAR	0.90	70.87
			E_RUSE	0.88	
			E_SPOI	0.79	
C6 Distance to coast	1.556	6.48	E_COAS	0.87	77.35
			P_MDTG	0.81	
			H_IBIR	−0.75	
C7 Cover change density	1.199	5.00	D_ARCC	0.95	82.34
			E_PRAR	0.90	

Table 3. Principal component analysis model for Gran Canaria.

Components (1st level of rank)	Initial values	Weight %	Features (2nd level of rank)	Weight (correlation)	Rotation sums of squared loadings
				% Variance	% Accumulated
C1 Distance to element of interest	7.49	32.55	E_COME	−0.94	32.55
				E_GRZO	0.88
				E_PZ35	−0.81
				E_HIGH	−0.69
C2 Distance to industrial areas	3,02	13.12	E_ENEQ	−0.94	45.67
				E_INAR	−0.93
				E_REEQ	−0.74
				E_TELI	−0.66
				E_MURC	−0.61
C3 Socioeconomic	2.95	12.83	I_ECAC	0.96	58.50
				I_COMA	0.91
				I_INDU	0.88
				H_INPC	−0.64
				H_IBIR	−0.59
C4 Distance to coast	2.64	11.51	P_MDTG	−0.93	70.01
				E_COAS	−0.90
C5 Density of cover change	1.50	6.53	D_CC00	0.85	76.54
				D_ARCC	0.79
				H_PODE	0.63
C6 Accessibility to roads	1.21	5.29	E_ROAD	−0.78	81.83
			D_ROAD	0.76	

ILs. The second component encompassed features associated with the distance to elements of interest such as: the network of paths, relevant cultural elements and green zones. This second component might have a direct relationship to potential IL occurrence and greater distance to those elements. More proximity to those elements of interest may thus discourage IL occurrence. The third component grouped socioeconomic features such as: economic activity indicator, per capita income and extractive industrial activity. More socioeconomic activity would thus mean more IL occurrence. That may accordingly be explained by the higher consumption and waste generation that occurs in areas with higher spending power and socioeconomic activity. The fourth component associated features of distance to elements of interest: aggregate extraction areas, urban centres, industrial areas and distance to roads. That may associate IL occurrence to the proximity of those elements. The fifth component likewise grouped features of distance to: agricultural areas, rural settlements and sports infrastructures. It was nevertheless observed that the first two features with more weight are those that define this component. That may link IL occurrence to the proximity of agricultural spaces and peri-urban settlements.

The sixth component concentrated the features: distance to the coast, altitude and property taxes (IBI). Based on this component, IL occurrence could be interpreted as being higher with less distance to the coast, lower altitude and higher IBI rate. This relationship of features may be because spaces at lower levels and close to the coast presented higher tax pressure due to the greater cadastral value compared to interior areas. Finally, the seventh component grouped the features of: density of changed uses between 1990 and 2012 and distance to protected natural areas. Both may have a direct relationship with ILs, whereby the increase of one of them may imply a higher occurrence of ILs.

In the case of GC, the first two components grouped features that maintain a connection to the distance to elements of interest. The second component was formed by the features of distance to large shopping centres, subsidence areas, motorways and green zones. The higher distance to areas and the lower distance to large shopping centres, subsidence areas and the motorway may be directly related to IL occurrence. The third component was formed by the features of distances, to industrial areas, telecommunication infrastructures, energy infrastructures, recreation infrastructures and distance to recycling centres.

The third component grouped socioeconomic features: economic activity, retail activity, industrial activity, per capita income and taxes on goods and property. This may imply that ILs are more likely to occur where there is greater economic, commercial, retail and industrial activity. Conversely, ILs may occur less in spaces with less per capita income and lower tax rates for goods and rural property, such as spaces farther from the coast.

The fourth component grouped the features of altitude and proximity to the coast. It could be interpreted that areas with lower altitude closer to the coast are likely to have higher occurrence of ILs. The fifth component was formed by the features of densities of: population, transitions to artificial covers and change of covers in the period between the years 2000 and 2006. This may associate IL occurrence to the increase of urban surface and population density. Finally, the sixth component grouped the features of: distance to secondary roads and density of communication routes. Spaces with high communication route density and less distance to roads may thus enable higher IL occurrence due to their easier accessibility.

Table 4. Confusion matrix of La Palma.

Observed		Predicted		
		Illegal landfill		Correct percentage
		NO	YES	
Illegal landfill	No	142	7	95.4
	Yes	13	140	91.5
Overall accuracy				93.5

Table 5. Confusion matrix of Gran Canaria.

Observed		Predicted		
		Illegal landfill		Correct percentage
		NO	YES	
Illegal landfill	No	266	20	93.0
	Yes	23	263	92.0
Overall accuracy				92.5

Both PCAs identified population and distance to elements of interest (i.e. industrial areas, motorway) as being decisive factors for IL occurrence, as did Biotto et al. (2009). However, the density and closeness to secondary communication routes were not as important as for Biotto et al. (2009), Jordá-Borrell et al. (2014) and Matos et al. (2012), probably due to the high accessibility of both islands. Otherwise, the socioeconomic factors were not as important as in other works (Jordá-Borrell et al., 2014; Matsumoto and Takeuchi, 2011), probably owing to the study's scale and the breakdown of the official socioeconomic information in local administrative units (NUTS3). Likewise, the altitude feature assumed a great deal of importance due to the high number of ILs situated in coastal spaces, as opposed to the mainland studies of Biotto et al. (2009), Gorsevski et al. (2012), Jordá-Borrell et al. (2014), Matsumoto and Takeuchi (2011) and Tasaki et al. (2007).

Logistic regression

The LR models of LP and GC had an overall accuracy of 93.5% (Table 4) and 92.5% (Table 5), respectively. For LP, six of 23 features were selected (Table 6) and for GC nine of 21 features (Table 7). The lower reduction of the model's features for GC was due to consideration of the economic indicators. That implies that the process of generating ILs is different on both islands. On LP the predictive role of the building density and distance

Table 6. Features in the LR equation, La Palma.

	B	Standard error	Wald	Sig.
D_BUIL	0.237	0.051	21.334	0.000
E_PRAR	−0.048	0.016	9.117	0.003
E_AGAR	−0.192	0.042	21.101	0.000
E_GRZO	0.170	0.016	32.601	0.000
E_COST P_MDTG	0.001	0.000	5.394	0.020
Constant	−1.567	0.538	3.879	0.004

Table 7. Features in the LR equation, Gran Canaria.

	B	Standard error	Wald	Sig.
D_ARCC	0.139	0.036	14.560	0.000
D_ROAD	0.063	0.018	12.638	0.000
E_ROAD	−0.182	0.053	11.887	0.001
E_COME	−0.066	0.018	12.632	0.000
E_GRZO	0.098	0.020	24.470	0.000
I_INDU	0.185	0.039	22.935	0.000
H_PODE	0.109	0.054	4.031	0.045
I_ECAC	−0.027	0.005	28.695	0.000
P_MDTG	−0.104	0.015	48.253	0.000
Constant	−2.052	0.706	8.443	0.004

to agricultural spaces features stood out, as they obtained the highest LR coefficient. There would thus be more IL occurrence in those areas with higher building density and less proximity to agricultural spaces. Likewise, for GC the relationship between IL occurrence and high greenhouse density was considered; for both cases it might be reasoned that there is higher occurrence in spaces close to eminently rural spaces, as in the study by Apostol and Miahi (2011). The high number of protected natural spaces on the island of LP means that the distance to protected natural spaces feature may have a confounding role in its relation to higher IL occurrence. Nevertheless, on both LP and GC the predictive role of the distance to green zones feature vis-à-vis less IL occurrence could be deduced, probably owing to the lower accessibility of those spaces. On the other hand, on LP the modifying role of the effect of interaction of the altitude and distance to the coast features stood out, because their inclusion improves the preciseness of the LR equation. The high aspect ratio of the island, as well as the higher number of settlements close to the coast and the higher accessibility may explain that interaction. Conversely, that interaction did not have a positive effect on the LR equation for GC, although the altitude feature did play a predictive role, probably because the ILs are situated in more accessible areas with higher population density located at lower levels. That notwithstanding, in the specific case of GC, the features with a notable predictive role and higher LR coefficients were: the industrial activity indicator, density of ground use transition to artificial covers and distance to roads. The areas with higher industrial activity had higher occurrence of ILs – an important feature due to the high localisation of ILs in industrial poles in eastern GC. On GC there would be a high ratio of IL occurrence and spaces with high densities of ground use transition to artificial covers between the years 1990 and 2012, a possible consequence of the effect of urban growth and the real estate boom in the first decade of the twenty-first century.

The binary LR equation was thus built with the estimated regression coefficients (*B*) for LP (equation (1)) and for GC (equation (2)), expressed as follows:

$$\begin{aligned} Logit(P) &= -1.567 + (0.237 * D\_BUIL) \\ &-(0.048 * E\_PRAR) - (0.192 * E\_AGAR) \\ &+ (0.170 * E\_GRZO) \\ &+ (0.001 * E\_COST * P\_MDTG) \end{aligned} \tag{1}$$

$$\begin{aligned} Logit(P) &= -2.052 + (0.139 * D\_ARCC) \\ &+ (0.063 * D\_ROAD) - (0.182 * E\_ROAD) - \\ &(0.066 * E\_COME) + (0.098 * E\_GRZO) \\ &+ (0.185 * I\_INDU) + (0.109 * H\_PODE) - \\ &(0.027 * I\_ECAC) - (0.104 * P\_MDTG) \end{aligned} \tag{2}$$

On both islands the modifying role of the effect of interactions between the features associated with density and distance to communication routes was assessed, as well as interactions between the socioeconomic indicators. However, their inclusion did not contribute to improved precision. In that regard, all the models used (PCA and LR) are linear methods. The use of novel methodologies based on machine learning such as regression and classification trees, random forest and support vector machines (Rodriguez-Galiano and Chica-Rivas, 2014) may improve the selection of features and identification of the causal relationships between them and IL occurrence, by establishing non-linear relationships and because they are non-parametric methods (Rodriguez-Galiano et al., 2018). This paper aims in turn to be a prologue of other future ones wherein the relationship between the generation of construction and demolition waste and the urban boom in the early twenty-first century is analysed in depth, and the oversight and sanction policies of local governments, extent of citizen awareness and involvement of the cost of access to waste treatment infrastructures are assessed. Other future lines of enquiry consist of assessing the risk ILs pose for tourism potential.

Conclusions

According to the ES, the typology of IL 'construction and demolition waste' is the most abundant one in LP and GC. No types of dissuasive measures were found in more than 95% of the ILs on both islands. ILs were present mainly in rural environments close to agricultural land. Some 50% of the ILs of LP and GC were found less than 0.2 km from a ravine and less than 1.4 km from the coast. Some 75% of the ILs on LP were situated less than 1 km from a tourist lodging. Some 75% of the ILs of GC were located less than 500 m from covers with high densities of ground use transitions for the period between 2000 and 2006.

For the case of LP the first three components of the PCA, corresponding to the features of population, distance from elements of interest such as path networks, relevant cultural elements and green zones, and socioeconomic features, explained 53.22% of the data variance. The latter components, with less explanatory importance, corresponded to the distance to rural areas, distance to the coast and density of ground use changes between 1990 and 2012. For GC the first three components of the PCA, corresponding to the features of distance to elements of interest such as industrial areas and telecommunication, energy and recreation infrastructures, distance to recycling centres, and socioeconomic ones, explained 52.82% of the data variance. The latter components, with less explanatory importance, corresponded to the distance to elements of interest, distance to rural areas, distance to the coast and density of cover transitions. The differences between the components of each island may be due to the lower demographic and building pressure, lower economic activity or greater extent of forested areas and protected natural areas on LP than on GC.

The LR analyses for each island simplified the feature space. On LP 6 of 23 features were selected and on GC 9 of 21 features. The LR models of LP and GC had an overall accuracy of 93.5% and 92.5%. The results show different drivers generating ILs for each island. On LP the most decisive features were building density, distance to agricultural spaces and distance to green zones, with more likelihood of IL occurrence in those areas with higher building density, less proximity to agricultural areas and farther from green zones. On GC, the most important features were: the industrial activity indicator, density of ground use transition to artificial covers, greenhouse density and distance to communication routes. Hence, the likelihood of IL occurrence would be higher when the value of the first three increases and the last one diminishes. The LR analysis on GC suggests that there is a high ratio of IL occurrence and spaces with high densities of ground use transition to artificial covers between the years 1990 and 2012, a possible consequence of the effect of urban growth and the real estate boom in the first decade of the twenty-first century. In future works it would be desirable to consider the involvement in IL generation of urban growth processes, manufacturing industry activities and greenhouse crops.

Acknowledgements

The authors are grateful for the financial support provided by the Department of Physical Geography and Regional Geographical Analysis of the University of Seville. The authors would also like to extend their gratitude to the Spanish Statistics Institute, Cartográfica de Canarias, S.A. (GRAFCAN) and Caixa Bank for proving the data.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The first author is a FPU Grant holder (MECD). Financial support was provided by the Department of Physical Geography and Regional Geographical Analysis of the University of Seville.

ORCID iD

Quesada-Ruiz L https://orcid.org/0000-0001-7886-5678

References

Akbari V, Rajabi MA, Chavoshi SH, et al. (2008) Landfill site selection by combining GIS and fuzzy multi criteria decision analysis, Bandar Abbas. *World Applied Science Journal* 3: 39–47.

Alexakis DD and Sarris A (2013) Integrated GIS and remote sensing analysis for landfill sitting in Western Crete, Greece. *Environmental Earth Sciences* 72: 467–482.

Apostol L and Mihai FC (2011) The process of closing down rural landfills. Case study: Neamț County. *Present Environment and Sustainable Development* 5: 167–174.

Biotto G, Silvestri S, Gobbo L, et al. (2009) GIS, multi-criteria and multi-factor spatial analysis for the probability assessment of the existence of illegal landfills GIS, multi-criteria and multi-factor spatial analysis for the probability assessment of the existence of illegal landfills. *International Journal of Geographical Information Science* 23: 1233–1244.

Boots BN and Getis A (1988) *Point Pattern Analysis*. Beverly Hills, CA: Sage Publications.

Breslow NE and Cain KC (1988) Logistic regression for two-stage case-control data. *Biometrika* 75: 11–20.

Bridges O, Bridges JW and Potter JF (2000) A generic comparison of the airborne risks to human health from landfill and incinerator disposal of municipal solid waste. *Environmentalist* 20: 325–334.

Calò F and Parise M (2009) Waste management and problems of groundwater pollution in karst environments in the context of a post-conflict scenario: the case of Mostar (Bosnia Herzegovina). *Habitat International* 33: 63–72.

Carranza EJM, Hale M and Faassen C (2008) Selection of coherent deposit-type locations and their application in data-driven mineral prospectivity mapping. *Ore Geology Reviews* 33: 536–558.

Demesouka OE, Vavatsikos AP and Anagnostopoulos KP (2014) GIS-based multicriteria municipal solid waste landfill suitability analysis: a review of the methodologies performed and criteria implemented. *Waste Management & Research* 32: 270–296.

Diggle PJ (1983) *Statistical Analysis of Spatial Point Patterns*. London: Academic Press.

Doak M, Khan S, Kelly G and Silvestri S (2007) *The Use of Remote Sensing to Map Illegal Landfills at the Border of Ireland/Northern Ireland*. Paper presented at the 11th International Waste Management and Landfill Symposium, Cagliari, Sardinia, 1–5 October 2007.

European Commission (2004) Directive 2004/35/CE of the European Parliament and of the Council of 21 April 2004 on environmental liability with regard to the prevention and remedying of environmental damage. *Regulation* 2003: 56–75.

European Commission (2008) Directive 2008/98/EC of the European Parliament and of the Council of 19 November 2008 on waste and repealing certain directives. *Official Journal of the European Union*: 3–30.

Fernández-Tabales A and Cruz E (2013) Análisis territorial del crecimiento y la crisis del sector de la construcción en España y la Comunidad Autónoma de Andalucía. *EURE (Santiago)* 39: 5–37.

Glanville K and Chang H-C (2015) Remote sensing analysis techniques and sensor requirements to support the mapping of illegal domestic waste disposal sites in Queensland, Australia. *Remote Sensing* 7: 13053–13069.

Gorsevski PV, Donevska KR, Mitrovski CD, et al. (2012) Integrating multi-criteria evaluation techniques with geographic information systems for landfill site selection: a case study using ordered weighted average. *Waste Management* 32: 287–296.

Hafeez S, Mahmood A, Syed JH, et al. (2016) Waste dumping sites as a potential source of POPs and associated health risks in perspective of current waste management practices in Lahore city, Pakistan. *Science of the Total Environment* 562: 953–961.

Hair JF, Black WC, Babin BJ, et al. (2010) *Multivariate Data Analysis*. 7th ed. London: Pearson, pp. 92–150.

Hernández Rodríguez GF (2006) Turismo sotenible en La Palma, un territorio reserva mundial de la biosfera. *Ambienta* 6267: 56–63.

Hernández Torres S (2003) *El espacio industrial en Canarias. El tejido empresarial y el territorio en las áreas insulares*. Las Palmas de Gran Canaria, Spain: Universidad de Las Palmas de Gran Canaria.

Ichinose D and Yamamoto M (2011) On the relationship between the provision of waste management service and illegal dumping. *Resource and Energy Economics* 33: 79–93.

Jones BJ (2008) The geography of open dumps in rural Appalachia. *Marshall Digital Scholar* 1–34.

Jordá-Borrell R, Ruiz-Rodríguez F and Lucendo-Monedero ÁL (2014) Factor analysis and geographic information system for determining probability areas of presence of illegal landfills. *Ecological Indicators* 37: 151–160.

Kontos TD, Komilis DP and Halvadakis CP (2005) Siting MSW landfills with a spatial multiple criteria analysis methodology. *Waste Management* 25: 818–832.

Lucendo-Monedero AL, Jordá-Borrell R and Ruiz-Rodríguez F (2015) Predictive model for areas with illegal landfills using logistic regression. *Journal of Environmental Planning and Management* 58: 1309–1326.

Matos J, Oštir K and Kranjc J (2012) Attractiveness of roads for illegal dumping with regard to regional differences in Slovenia. *Acta Geographica Slovenica* 52: 431–451.

Matsumoto S and Takeuchi K (2011) The effect of community characteristics on the frequency of illegal dumping. *Environmental Economics and Policy Studies* 13: 177–193.

Monteiro Santos FA, Mateus A, Figueiras J, et al. (2006) Mapping groundwater contamination around a landfill facility using the VLF-EM method – a case study. *Journal of Applied Geophysics* 60: 115–125.

Notarnicola C, Angiulli M and Giasi C (2004) Southern Italy illegal landfills detection based on spectral analysis of remotely sensed data and landcover maps. In: *Proceedings on remote sensing for environmental monitoring, GIS applications, and geology*, Barcelona, 13 February 2004, pp. 483–493. Bellingham, WA: SPIE.

Rodriguez-Galiano V and Chica-Rivas M (2014) Evaluation of different machine learning methods for land cover mapping of a Mediterranean area using multi-seasonal Landsat images and Digital Terrain Models. *International Journal of Digital Earth* 7: 492–509.

Rodriguez-Galiano VF, Luque-Espinar JA, Chica-Olmo M, et al. (2018) Feature selection approaches for predictive modelling of groundwater nitrate pollution: An evaluation of filters, embedded and wrapper methods. *Science of the Total Environment* 624: 661–672.

Schill W, Jockel KH, Drescher K, et al. (1993) Logistic analysis in case-control studies under validation sampling. *Biometrika* 80: 339–352.

Sebenik I (1994) Pokrajinske značilnosti manjših neurejenih odlagališč odpadkov v Sloveniji z zasno-vo akcijskega načrta ureditve in oceno tveganja onesnaženja podzemnih voda. *Geographica Slovenica* 26: 1.

Şener E and Karag R (2011) Solid waste disposal site selection with GIS and AHP methodology: a case study in Senirkent – Uluborlu (Isparta) Basin. *Environmental Monitoring and Assessment* 173: 533–554.

Şener Ş, Şener E, Nasc B, et al. (2010) Combining AHP with GIS for landfill site selection: a case study in the Lake Beyşehir catchment area. *Waste Management* 30: 2037–2046.

Silverman BW (1986) *Density Estimation for Statistics and Data Analysis*. London: Chapman and Hall.

Silvestri S and Omri M (2008) A method for the remote sensing identification of uncontrolled landfills: formulation and validation. *International Journal of Remote Sensing* 29: 975–989.

Smrekar A (2011) From environmental awareness in word to environmental awareness in deed: The case of Ljubljana. *Acta Geographica Slovenica* 51: 277–292.

Tasaki T, Kawahata T, Osako M, et al. (2007) A GIS-based zoning of illegal dumping potential for efficient surveillance. *Waste Management* 27: 256–267.

Tasaki T, Matsui Y, Kawahata T, et al. (2004) Analysis of geographic attributes and probabilities related to illegal dumping. *Journal of the Japan Society of Waste Management Experts* 15: 1–10.

US Environmental Protection Agency (1998) A framework for ecological risk assessment at the EPA. *Federal Register* 63: 26846–26924.

Uyan M (2014) MSW landfill site selection by combining AHP with GIS for Konya, Turkey. *Environmental Earth Sciences* 71: 1629–1639.

Watkins E (2015) *A case study on illegal localized pollution incidents in the EU*. A study compiled as part of the EFFACE Project. London: IE.