

Spatial analysis of graffiti in San Francisco

Veronika Megler^{a,*}, David Banis^b, Heejun Chang^b

^a Department of Computer Science, Portland State University, Oregon, USA

^b Department of Geography, Portland State University, Oregon, USA



A B S T R A C T

Keywords:

Graffiti
Youth
Spatial analysis
GWR
San Francisco

Graffiti, as a social phenomenon, has been with us since people first began painting on cave walls. Our responses to graffiti range from recognition of artistry, to interpreting it as a sign of urban decay and disorder. Major cities respond to tax-payers with this latter response by providing graffiti abatement programs, often at substantial cost; thus, understanding and mitigating the causes of graffiti has tangible value. Using spatial analysis, we explore the combined causes of graffiti creation and the subsequent reporting of graffiti for removal in San Francisco, CA, USA. Using a combination of census data and city data, we identify five factors that have significant correlation to graffiti reports, and use them to build a regression model. We show that graffiti is created in areas with high densities of young males, and that commercial zones have the highest rate of graffiti reports. We show that a Geographically Weighted Regression model of these five factors explains over two-thirds of the variation in graffiti reports in San Francisco. Further, our findings are consistent with the dual hypotheses of graffiti as a form of communication or advertising aimed at a target market of other young males, along with the broken window thesis of graffiti interpreted as a sign of social disorder.

© 2014 Elsevier Ltd. All rights reserved.

Introduction

Graffiti has been with us for as long as human history; anthropologists speculate that some prehistoric cave paintings are graffiti, with the majority, according to statistical analyses, drawn by teenage males (Carey, 2006; Guthrie, 2005). According to the Oxford Dictionary, the term “graffiti” is defined as “writing or drawings scribbled, scratched or sprayed illicitly on a wall or other surface in a public place,” and certainly cave paintings fit this model. By 1868, the graffiti in Pompeii was recognized as a record of social, political and domestic life, though with a focus on what was politely described, with some disappointment, as “ordinary people” (Baird & Taylor, 2010).

More than 150 years later, our response to graffiti is multifaceted. There is recognition of its status as art (Lachmann, 1988); graffiti as a reflection of societal customs and attitudes and their change over time (Stocker, Dutcher, Hargrove, & Cook, 1972); as a method to attain notice or fame (Halsey & Young, 2002, 2006; Lachmann, 1988); as a form of political statement (Ferrell, 1995) or as territorial markers (Ley & Cybriwsky, 1974). More recently,

there has been a recognition – or a creation of – a distinction between graffiti and street art or community art (a form of legitimized graffiti) (Mcauliffe, 2012), and the latter's claimed role in urban renewal (Schuermans, Loopmans, & Vandenaabeele, 2012). Graffiti may even be viewed as integral to the urban character of a place and in so doing become a tourist attraction (Dovey, Wollan, & Woodcock, 2012). A single city may have a variety of responses – from tolerance to zero-tolerance – at a single time, reflecting the different stances of local government officials, and changing as new officials are elected (Mcauliffe, 2012). However, graffiti is commonly seen by local residents and municipalities as a nuisance and sign of criminality and danger (Cresswell, 1992b; Doran & Lees, 2005; Glazer, 1979; Hung, Ly, & Ngo, 2010; Mcauliffe, 2012). In this view, graffiti is perceived as a sign of public disorder, along with public intoxication, garbage and abandoned cars (Sampson & Raudenbush, 2004). The much-quoted “broken window” theory of urban decline (Kelling & Wilson, 1982) holds that a single, unfixed broken window (or graffito) can lead to breakdown of community. As such, city agencies are expected by a subset of citizens to actively control and remove graffiti. At the same time, active removal of graffiti is expensive, as is the constant surveillance required to detect graffiti. Thus, some cities including San Francisco now require removal of graffiti by property owners within an established time period (San Francisco Public Works Code Article 23, n.d.). City agencies in San Francisco spend more

* Corresponding author. Tel.: +1 503 208 6152.

E-mail addresses: vmegler@cs.pdx.edu, hgaggin@mgh.harvard.edu, dbanis@pdx.edu (V. Megler).

than \$20M annually for graffiti abatement (Gathright, 2010). In 2009, San Francisco initiated a program called Zero Graffiti for a Beautiful City, where they attempted to reduce graffiti through a combination of community involvement and law enforcement.

To facilitate reporting of graffiti, as well as other civic nuisances, systems have been developed for citizens to report these nuisances using smart phones or web-based applications (DeMeritt, 2011). The availability of volunteered geographic information (VGI) has increased rapidly in recent years and attracted the attention of numerous scholars (Budhathoki, Nedovic-Budic, & Bruce, 2010; Elwood, Goodchild, & Sui, 2012; Flanagan & Metzger, 2008; Goodchild, 2007) as a new form of geographic information with much potential. Cities are actively using such volunteered information in order to engage and respond to citizens (Johnson & Sieber, 2013). This approach has a number of advantages for managing graffiti as it minimizes the surveillance costs, and allows the city to focus on removal activities. In addition, it aligns city expenditures on graffiti removal from public property with the citizens most concerned about graffiti. However, in areas where citizens tolerate graffiti, presumably because they do not see it as a threat, the citizens may be less likely to report that graffiti to the city. Other areas where citizens feel empowered, such as high income areas presumed to see little graffiti, may still actively report it.

The reporting of graffiti is, however, not always consistent with the presumed reasons for creating graffiti. If graffiti artists are creating graffiti as a political statement to undermine the existing power structure, it might be expected that they make those statements in areas where the power structure is strongest – that is, high income residential and commercial areas, or in other areas of power, such as near police stations. If graffiti artists are creating graffiti to control public space, it might be expected that they make those statements in areas where the public space is most contested (Mcauliffe, 2012). On the other hand, if graffiti artists are establishing identities or communicating with other artists, predominantly young men, they may choose to do so in their own neighborhoods (Ferrell, 1995; Mcauliffe, 2012; Monto, Machalek, & Anderson, 2013). The relationship between graffiti reports and actual graffiti is a complex one; it reflects differential acceptability of graffiti by location (Cresswell, 1992a, 1992b; Haworth, Bruce, & Iveson, 2013; Shobe & Banis, 2014). We would expect these differences to be reflected as detectable spatial variation in the reporting of graffiti.

Our analyses take a mixed approach by using a quantitative/statistical approach to understanding phenomena that are most often analyzed qualitatively. The spatial distribution of graffiti is a social process for which little or no quantitative analysis has been performed. By performing quantitative analysis of spatial patterns of graffiti against available demographic data that reflects spatial variations, we seek to identify correlations that either help support existing hypotheses, or identify predictive factors that may lead to new research and new hypotheses. In particular, we may be able to see the geospatial relationship of graffiti reports to public spaces, residential or commercial zoning, and population factors such as income levels, gender and age densities.

This study, therefore, investigates the statistical support for a previously unexplored tension. Extensive interviews with graffiti artists describe the effort graffiti artists go to in selecting the locations for their graffiti (Halsey & Young, 2002) and the thrill the artists experience when their graffiti is noticed (Lachmann, 1988; Monto et al., 2013). These artists, today as in prehistory (Carey, 2006; Guthrie, 2005), tend to be young males (Austin, 2001; Castleman, 1982; Ferrell, 1995; Miller, 2002; Monto et al., 2013). We test support for the hypothesis that the most graffiti will be created where the graffiti artists' target market is located (Lachmann, 1988; Monto et al., 2013). In practical terms, this

translates to locations with a high percentage of resident young males, or locations in which young males are likely to congregate or travel through; this is similar to the relationship frequently reported between the burglary locations and offender residence, and between the relationship between major roads and burglar risk (Breetzke, 2012). We can identify locations with a high percentage of resident young males from census data. Given limited data on locations where young males may travel to or through, we assume that their travel patterns will mirror the general population, and so used commercial districts and arterials as proxies for these locations.

On the other hand, many researchers (Cresswell, 1992b; Glazer, 1979; Hung et al., 2010; Kelling & Wilson, 1982; Sampson & Raudenbush, 2004), hypothesize that another segment of the population will see graffiti as a sign of social disorder in well-to-do neighborhoods, to be discouraged; that is, parents of young males, and commercial property owners concerned that the appearance of local criminality and danger (as represented by graffiti) will impact their business. We test statistical support for this hypothesis that areas with a high proportion of young males or percentage of commercial property will be correlated with the highest number of graffiti reports.

Study area

The city of San Francisco is located at the tip of the San Francisco peninsula in California. It is the center of a larger Bay Area region of 7.5 million people that includes cities such as San Jose and Oakland. As of 2010 by the (U.S. Census Bureau, 2010), the city had a population of around 805,000 people, and a mainland area of approximately 11 km², giving it a density of around 73,180 people per km². Fig. 1 shows the study area with some key neighborhoods noted. The densely-populated city mixes numerous commercial zones along arterials with residential zoning, and has a small but increasing amount of mixed-use commercial-residential zones along arterials. Downtown and the financial district, with their cluster of commercial properties, are found in the northeast, and industrial lands are primarily located along the eastern waterfront; these areas tend to be low-lying, with little elevation change. The geographic center of the city includes gentrified neighborhoods such as Haight-Ashbury and Hayes Valley, located at higher elevations on steeper slopes, as well as the diverse and vibrant but rapidly gentrifying Mission District. The southeast part of the city is home to predominantly Hispanic and Asian neighborhoods, and both the northwest and southwest are dominated by park lands.

Data

The study period encompassed San Francisco's calendar 2009 and 2010. Table 1 lists the datasets used for this study. In this study, we used voluntary geographic information (VGI) collected by San Francisco's Department of Public Works reporting system during calendar years 2009 and 2010, just after the Zero Graffiti for a Beautiful City program began. Individuals are often in the best position to, and are much more likely to, provide current information about local conditions, and have been found to be more likely to report using digital systems (Feick & Roche, 2013; Flanagan & Metzger, 2008; Gira, Bédard, & Roche, 2010; Miller, 2010). The graffiti reports are unprompted responses reflecting visceral reactions to graffiti in the local environment, rather than artificial responses such as those solicited in periodic surveys. There are several reasons why this dataset might be skewed; for example, although the reporting system accepts both online and telephoned reports, it might be underrepresent those without access to either of these tools. However, the length (two years), coverage (the whole

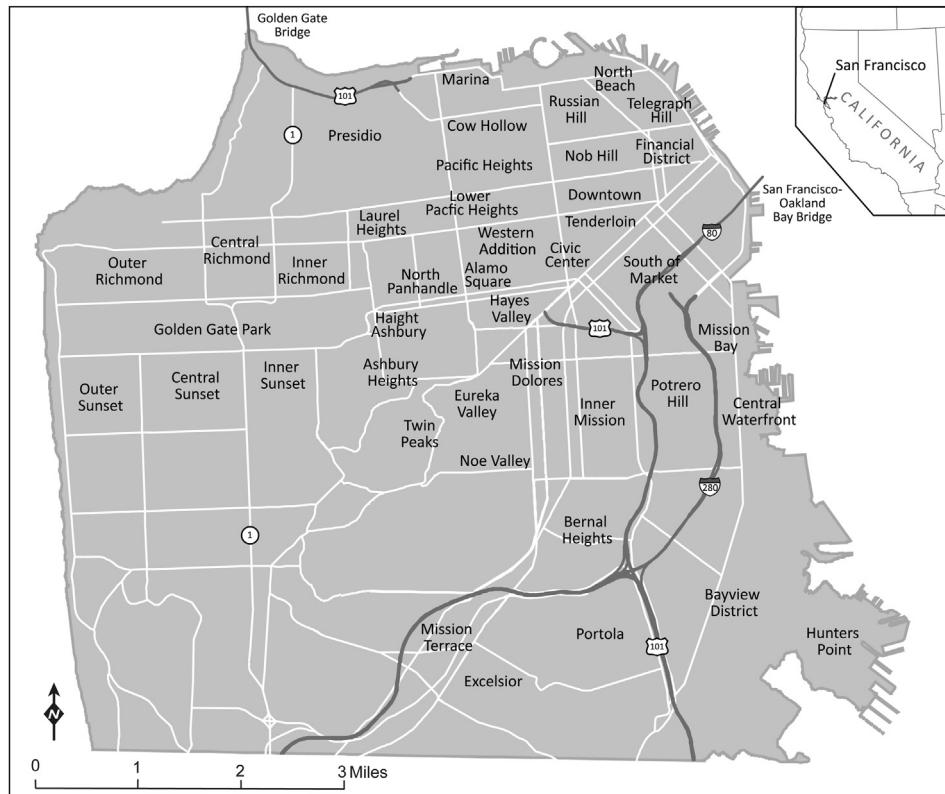


Fig. 1. Study area.

of San Francisco) and large size of this dataset (over 59,000 reports) offsets much of this bias (Griira et al., 2010; Johnson & Sieber, 2013; Miller, 2010). The people reporting the graffiti for removal are, by definition, people who are concerned about the existence of graffiti, and thus are of particular interest to the city (who uses the reports to focus their abatement efforts) and to our study (as likely adherents of the broken window theory).

Therefore, the focus of this study is on *reported* graffiti (herein referred to as “graffiti”), that is, the subset of graffiti that exists and

is reported to the city for removal. It is, however, a record only of graffiti that exists and was reported to the city; the data cannot be interpreted as absolute graffiti densities throughout the city (even if the pattern may be similar), but is instead a combination of the existence of both graffiti and citizens who perceive graffiti as a threat. In fact, by using Google Street View, we were able to confirm dense graffiti in industrial areas in which no graffiti reports were made.

Haworth et al. use similar VGI, reports of graffiti for removal, to identify hotspots in Sydney. They then used field surveys of the hotspots to identify different types and qualities of graffiti, such as form, style and complexity, and studied the relationship to council variations in graffiti tolerance (Haworth et al., 2013).

The graffiti data is a complete listing of the graffiti service requests upon which the city acted. The service requests included an address and geospatial location, and a textual request type. The request type sometimes provided information such as whether the graffiti was offensive or not, and whether it was on a commercial or residential building or on street articles such as a pole or a signal box. The address information fell into two main categories. A subset of requests provided the exact street address. The remainder of the requests had their address masked to the center of the closest street intersection. Exact locations of these graffiti are not available and must be estimated; the actual location may be ½ block in any direction from the reported location.

In 2009, there were 26,858 service requests, with 32,206 in 2010. A total of 3124 requests were for post-abatement inspections; we removed these from the study data. No temporal pattern could be seen in the data, and as we were more interested in the factors associated with graffiti location than the temporal changes, we aggregated the available data for the entire study period during further analysis.

Table 1
Source of data used in study.

	Source	Units ^a
Dependent variable		
Graffiti removal service requests	SF City	Number
Physical variables		
Elevation	SF City	Feet above sea level
Distance to police stations	SF City	Feet
Distance to major arterials	SF City	Feet
Demographic variables		
Density of males 15–24 years of age	2010 U.S. Census	% of residents
Income		
>\$100K	2010 U.S. Census	% of residents
\$75K–\$100K	2010 U.S. Census	% of residents
\$50K–\$75K	2010 U.S. Census	% of residents
\$25K–\$50K	2010 U.S. Census	% of residents
<\$25K	2010 U.S. Census	% of residents
City environment: Zoning		
Residential %	SF City, Zoning, aggregated	% of cell
Commercial %	SF City, Zoning, aggregated	% of cell
Public %	SF City, Zoning, aggregated	% of cell
Industrial %	SF City, Zoning, aggregated	% of cell

^a After transformation.

Methods

In this section we describe the methods used in our GIS analysis, followed by our spatial statistical analysis.

GIS analysis

We use GIS and spatial analysis to look for correlations between graffiti reports and several social factors identified in qualitative research, to see if the social hypotheses can be supported or falsified by a GIS model. The first step was to move to a common areal unit for the data. After removing the post-abatement inspections, 55,940 graffiti reports remained; of these, 29,408 had specific addresses, with the remaining 26,532 listed as the closest street intersection. Where the exact address was given, we used that location as given. Given the ½ block uncertainty in the physical location of much of the reported graffiti, we created a grid with individual cells measuring 381 m (1250 feet) per side. This size ensures that each cell contains several city blocks (often 6 or 7), while keeping the cell size small relative to the city area and in order to identify local patterns. Graffiti locations that did not have exact addresses were buffered with a 76.2 m (250 ft.) buffer, and the buffers intersected with the grid. The percentage of the area in each grid cell was taken as an estimate of the probability of the graffiti having occurred in that cell; the total of the buffer areas assigned to each cell was divided by the size of the buffer, to give a final count of the graffiti in each cell. Obviously, if the entire buffer for a graffiti report falls in one cell, then 100% of the buffer area will be within that cell. After transforming in this way, the total graffiti points allocated to grid cells was 29,405, with a loss of 3 due to calculation and rounding errors. Fig. 2a shows the resulting density of graffiti reports across the city; throughout the study period graffiti is most prevalent in the North-East commercial district and surrounding residential areas.

We grouped 41 different land-use zonings into four broad categories: residential (46.6% of San Francisco area), public lands (31.6%), industrial (10.9%), and commercial (9.4%). An additional 1.5% is classified as “water”; no graffiti points intersected with “water”. Mixed residential-commercial consists of less than 2% of the city, and was grouped with the residential category. Fig. 3a illustrates the distribution of zoning in San Francisco. The grid was intersected with these four zoning categories, and a percentage of area calculated for each zoning category for each cell.

We assigned zonings to reported graffiti according to the following rules. If the description of the graffiti included the zone

type (“Graffiti on commercial building”), we assigned the graffiti to that zone. For the remaining graffiti, where the graffiti location in the latitude/longitude columns intersected with a specific zoning, we assigned the graffiti to that zoning. The remaining graffiti were those that had been anonymized to the closest street corner. For these, we buffered the graffiti location to a ½ block radius and calculated a probability of the graffiti being in each zone based on the percentage of area within the buffer assigned to each zone category. An analysis of zoning area versus graffiti reports showed a much higher-than-expected proportion of graffiti reports in commercial areas, and a much lower-than-expected proportion in industrial areas, with highly significant results (Megler, Snodgrass, & Blackmore, 2011). The number of graffiti in each grid cell by zone type was then summed.

Elevation, shown in Fig. 3b, was calculated for each grid cell based on the proportion of the area at each elevation. We also calculated the elevation difference based on the difference between the minimum and maximum elevations as a proxy for the steepness of the grid cell (see Table 2).

Census data, such as percentage of young males (15–24) and income brackets, was summed for each census tract where necessary and converted to a density for each census tract. The census tracts were intersected with the grid, and new densities calculated based on the percentage of area with each intersecting census tract’s density.

Major arterials and police station locations are shown in Fig. 3a. For these distance-based measures, we calculated the Euclidean distance from each graffiti reported location to the nearest of these features. We calculated an average distance of the graffiti for each cell. We thereby take into account grid cells where, for example, all graffiti is reported on one side of the cell, and do not bias the results by using distance to the center of the cell or some other point. This approach has the most impact for grid cells with an arterial along one edge, where the majority of the graffiti is reported on the arterial. We recognize that there is potential error in this calculation caused by the uncertainty in some graffiti locations, but as this error is evenly distributed in each direction, it should not unduly bias our results.

Spatial statistical analysis

We explored the relationship between graffiti reports and each individual variable using scatterplots and QQ plots. Our variable distributions are not normal, as confirmed by the Jarque–Bera test for each variable. To measure the correlation between independent

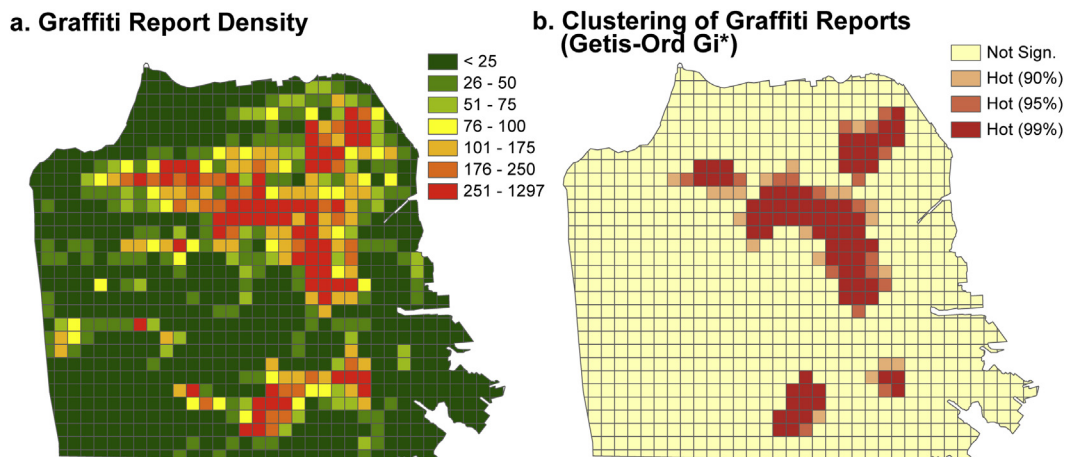


Fig. 2. a. Graffiti Report Density b. Clustering of Reported Graffiti Density (Getis–Ord G_i^*).

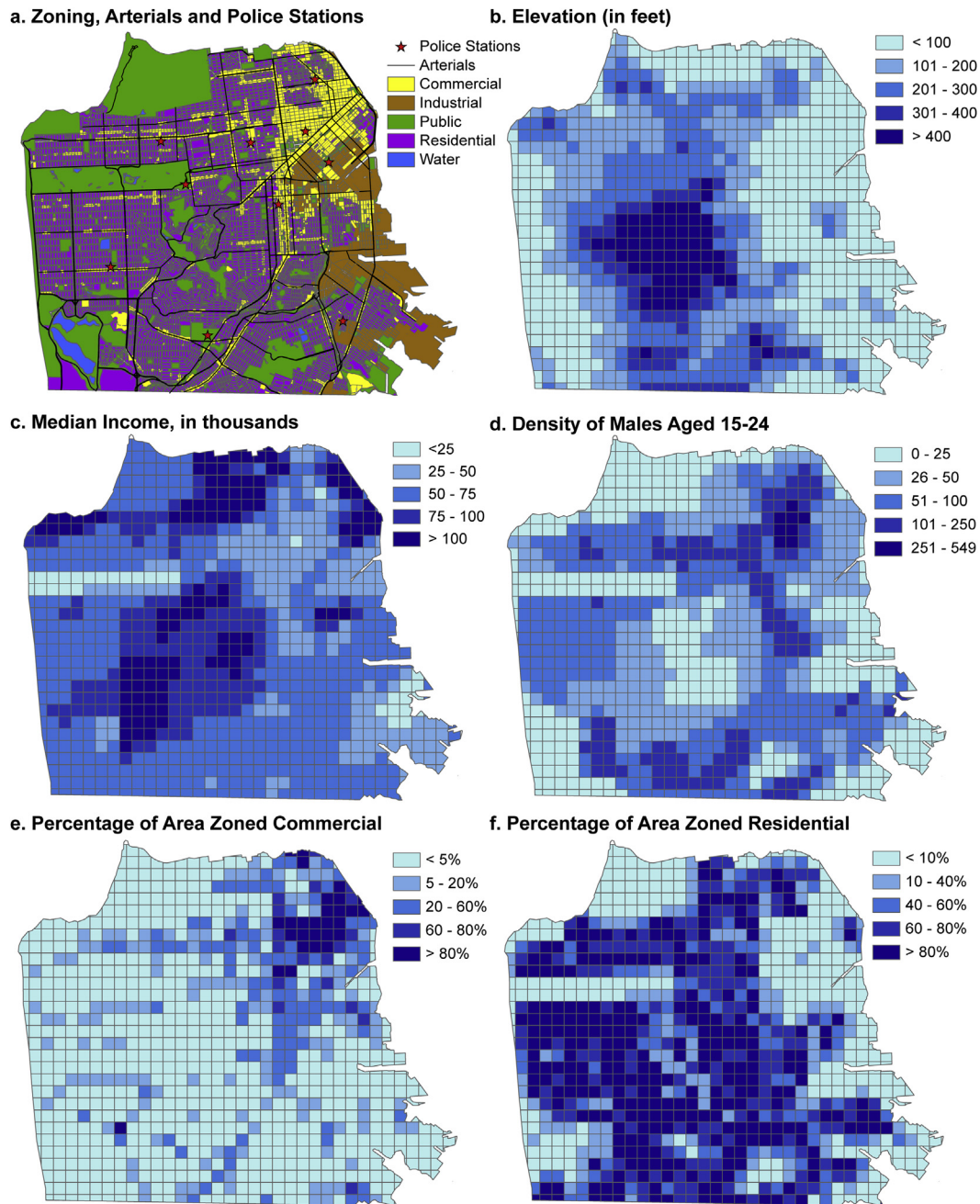


Fig. 3. Data. a. Zoning, Arterials and Police Stations; b. Mean Elevation per Grid Cell (in feet); c. Median Income, in Thousands; d. Density of males aged 15–24; e. Percentage of Area Zoned Commercial; f. Percentage of Area Zoned Residential.

variables and graffiti reports individually, we performed correlation analysis for each variable, and reviewed the Pearson's correlation coefficient (results shown in Table 2). We did not normalize and standardize the data, since our intended regression and analysis techniques calculate coefficients that adjust for very different units of scale (as can be seen in our results below).

As inspection of the QQ plot showed the log of graffiti counts for non-0 grid cells was very close to normal, we used the log of graffiti counts as our dependent variable. Grid cells with 0 reported graffiti were primarily in park or industrial areas (see Fig. 2a).

We performed Ordinary Least Squares (OLS) regression (using SPSS) to analyze the relationships between our dependent variable and various subsets of our explanatory variables. We performed a

variety of forward and backward regressions, using different combinations of variables suggested by our hypothesis. We also tested several combinations of variables not supporting our hypothesis (specifically the income variables), to confirm that we would not get better results from these alternatives. In each case, we assessed the quality of our model by reviewing the log likelihood (where higher numbers identify a better model), Akaike Information and Schwarz Criteria (where lower numbers are better), in addition to the overall and adjusted r^2 achieved (with higher numbers being better). The results of each test were used to identify further models, in an iterative process. We used Moran's I to measure the degree of spatial autocorrelation in our response and independent variables and the residuals of OLS and GWR models.

Table 2
Correlation coefficient between density of graffiti reports and explanatory variables.

Explanatory variable	<i>r</i>	Significance
Density of males 15–24 years of age	0.476	***
Elevation	0.032	
Elevation difference	0.070	**
Distance to police stations	0.066	**
Distance to major arterials	0.056	*
Income		
Income greater than \$100K	0.143	***
Income between \$75K–\$100K	0.063	*
Income between \$50K–\$75K	0.000	
Income between \$25K–\$50K	0.191	***
Income less than \$25K	0.058	*
Zoning		
Percent zoning of residential	0.117	***
Percent zoning of commercial	0.378	***
Percent zoning of public	0.227	***
Percent zoning of industrial	0.229	***

*: Significant at the 10% level.

**: Significant at the 5% level.

***: Significant at the 1% level.

OLS regression assumes that the values of the coefficients do not change across the study area. However, we believed that our variables would exhibit spatial non-stationarity, that is, that the strength of each variable's effects varies across the study area. For example, it is likely that different school districts have different programs that may affect the behavior of young males (at least those still in school), while changing social demographics across the study region may also modify behavior. We use Geographically Weighted Regression (GWR), as we expect it to produce more informative results regarding the variation of parameters across the study area than other approaches (Fotheringham, Charlton, & Brunson, 1998). The GWR technique is an extension of ordinary regression, wherein regression is performed at each location using a “local” view of the surrounding data locations. Given that graffiti report data are not evenly distributed over space, we used an adaptive Gaussian kernel, so the bandwidth reduces as feature density increases.

Results

We first describe the spatial patterns of graffiti in Section 5.1. We then describe our Ordinary Least Squares model, in Section 5.2, followed by our Geographically Weighted Regression model in Section 5.3.

Spatial patterns of graffiti

Moran's *I* index for the raw graffiti densities shows significant clustering ($I = 0.65$, $z = 27.04$, $p = 0.00$). Fig. 2b shows the results of Getis-Ord G_i^* cluster analysis, with a large high–high cluster over downtown and several smaller outlying high–high clusters. No low–low or high–low clusters are present.

Preliminary analysis explored and eliminated several possible factors as explanatory variables, including many that we initially expected to have a high correlation: these included education levels, proximity to parks, population density, vandalism arrest patterns and known gang zones (Megler et al., 2011). This same report also did not find that the “watch streets” program, intended to raise citizen awareness, supervision and reporting in their community, had any correlation with graffiti reports; it is, however, unlikely that precisely offsetting trends exist (that is, for example, that a watch streets program in an area that would otherwise have higher graffiti precisely offsets the factors that would lead to higher graffiti).

The same analysis identified several factors (that is, zoning, percentage of young males, distance to arterials and distance to police stations) as appearing to exhibit spatial relationships to graffiti reports; those factors were brought forward to this study.

In addition, we explored elevation as a potential explanatory factor given the nature of property distribution of San Francisco's many hills; in his recent study of Tshwane, South Africa, Breetzke found a positive relationship between lower altitude and burglary risk, while no relationship was found between slope and burglary risk (Breetzke, 2012). The results of correlation analysis between graffiti reports and the individual variables are reported in Table 2. Based on these results, all but five variables (elevation; income between \$75k and 100k; income between \$50k and \$75k; income less than \$25k; and distance to arterials) were individually statistically significant at the 5% level. Although statistically significant, the *r* for the remaining two income-related variables (income greater than \$100k and income between \$25k and \$50k) were extremely low; thus, no income category was meaningfully correlated with the density of graffiti reports.

In contrast, the density of young males explained a meaningful proportion of the variance ($r = 0.48$, $p < 0.001$) in graffiti reports. The second-highest *r* is associated with commercial zoning ($r = 0.38$, $p < 0.001$). This combination seemed to support our hypothesis that graffiti reports were highly correlated with where young males either live, or hang out.

Ordinary least squares regression

Given the frequent expectation of a correlation between low income areas and a high density of graffiti, we explored income as a predictor in addition to our main hypothesized influences. We expect higher income areas to have less graffiti but a higher tendency to report graffiti, while lower income areas may have higher graffiti but be less likely to report it. While our data does not allow us to validate these assumptions, it is unlikely that these counter-vailing influences balance sufficiently to completely remove any correlation that actually exists. As expected, most income variables were eliminated early in the backwards regression model. The exception was “income below \$25k”, which was eliminated in the second-last step; eliminating this variable did not change the adjusted r^2 (0.277), but raised the *F*-statistic from 41.1 to 55.7 (with $AICc = 2877.09$; Schwarz criterion = 2904.66). However, despite the insignificance of the adjusted r^2 for distance to police stations and arterials, these variables were retained in the last step, giving the best-performing OLS model with the fewest variables.

The final OLS model has a multi-collinearity condition number of 6.90, implying that multi-collinearity is not an issue. The model has a Durbin–Watson statistic of 0.970, implying that there is positive correlation between the residuals; however, at that value the test is inconclusive. Thus, the *F*-statistic is inflated, over-estimating the overall significance. However, with a model *F*-statistic of 47.6, the result is significant. In addition, the Jarque–Bera statistic is significant, showing that our residuals are not normally distributed; since the error distribution is not normal, the Durbin–Watson test may not be the best test to apply here. Given the significance of the Jarque–Bera statistic, there is likely skew or kurtosis in our residuals. Moran's *I* of 0.51 with $p = 0$ confirms this.

With these notes, we report here the resulting OLS regression equation:

$$\begin{aligned} \text{Log}(\text{Graffiti reports}) = & 1.13 + 1.3*\text{RESPCT} + 2.9*\text{COMPCT} \\ & + 0.01*\text{M15DENS} + 0.000066*\text{ARTDIST} \\ & - 0.00044*\text{PDDIST} \end{aligned}$$

Euclidean distance to police station has negative coefficients, which is the opposite of what we expect. However, as noted, we expect the variables in this equation to exhibit non-stationarity, and combined with the non-normality of the residuals, we do not regard the equation produced by OLS as a good predictor of our dependent variable.

In every model, elevation difference (our proxy for steepness) was identified as having relatively high correlation with our dependent variable, but actual elevation was the first variable eliminated; the sociological process underlying this correlation is not clear. A hypothesized relationship between income and elevation (higher incomes living at higher elevations) was tested and found not significant. We tested whether the real influence is the magnitude of change in with flatter areas being more attractive to

graffiti artists, elevation, using the maximum elevation difference as a proxy, but this variable was also shown not to be significant in most model runs.

Geographically weighted regression

For geographically weighted regression, we used the same five independent variables (commercial zoning percent, residential zoning percent, density of young males, distance to arterials and distance to police stations). Using these same variables, GWR explained over two-thirds of the variation in our dependent variable, with an r^2 of 0.81. Our adjusted r^2 of 0.74 was also much higher than with OLS; AICc of 2267.86 was substantially lower than in our OLS model, showing that this is a stronger model.

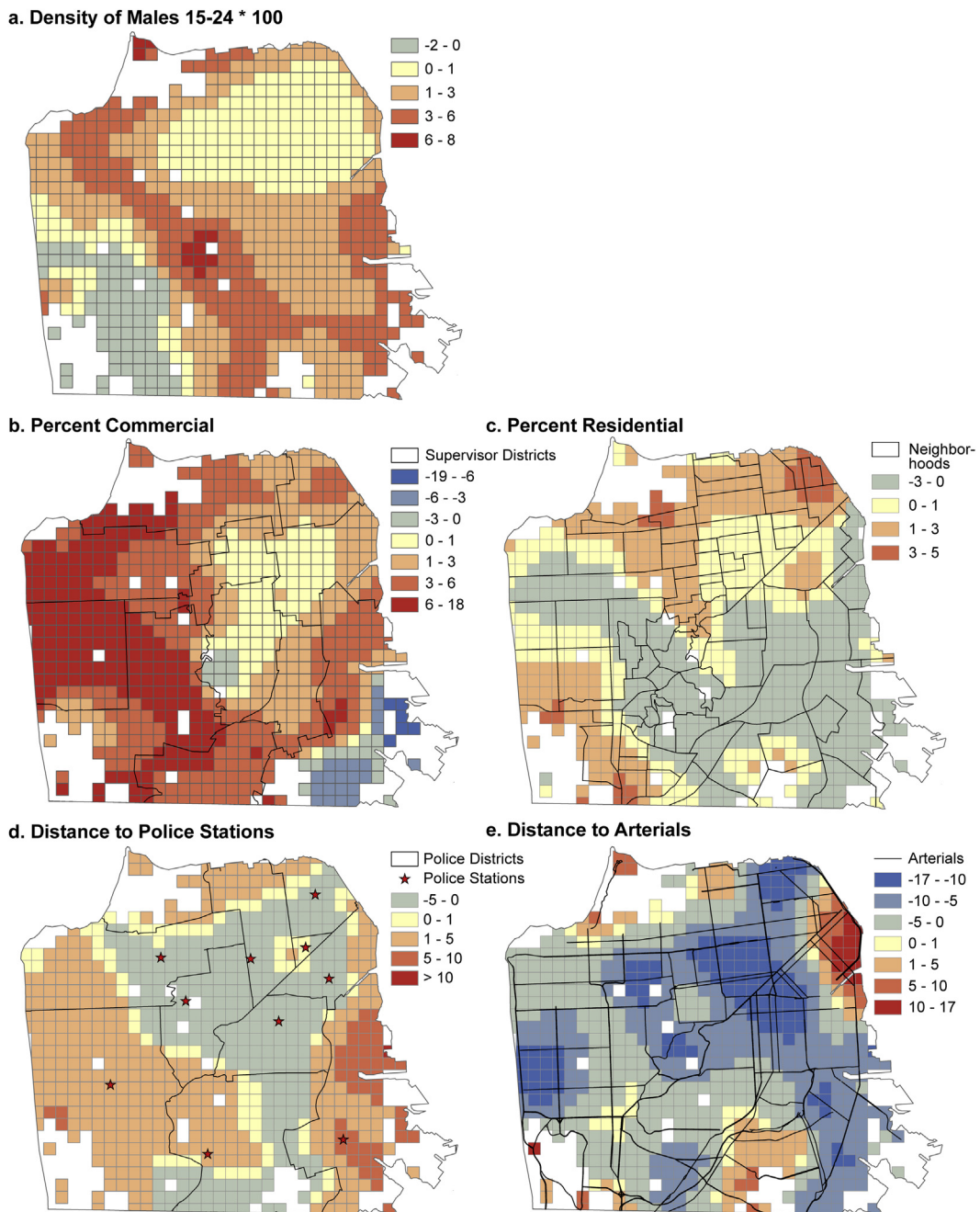


Fig. 4. Coefficient Maps of GWR Variables. a. Males, 15–24, multiplied by 100; b. Commercial Percent; c. Residential Percent; d. Distance to Police Stations * 10,000; e. Distance to Arterials * 10,000.

Table 3
Global (OLS) and local (GWR) coefficients of the empirical regression model.

	GWR					OLS
	Minimum	Lower quartile	Median	Upper quartile	Maximum	
Intercept	-4.714	-0.051	0.770	2.545	5.640	1.13
Young males	-0.018	0.005	0.167	0.319	0.075	0.01
% Commercial	-19.054	1.545	3.732	6.101	17.964	2.93
% Residential	-2.394	-0.551	0.264	1.199	4.417	1.32
Distance to police stations	-0.0017	-0.0008	-0.0004	-0.00009	0.0016	-0.00044
Distance to arterials	-0.0005	-0.0008	0.0001	0.0003	0.001	0.000066

Fig. 4 maps the coefficients for our resulting variables across the city.

Table 3 summarizes the characteristics of the coefficients for GWR, and compares them to the coefficients for OLS. The medians from GWR can be seen to be relatively similar to those for OLS but the minimums and maximums vary substantially, lending support to the spatial non-stationarity of the variables.

Visual inspection of the GWR coefficient maps in Fig. 4 shows that each variable exhibits non-stationarity; the patterns are also more homogenous than the patterns exhibited in the data, implying a geographic process is present rather a direct relationship at each location. In addition, each variable shows a different spatial pattern of effects. These variations support the use of GWR in modeling our dependent variable.

Visual inspection shows that graffiti hotspots coincide with high percentages of young resident males; however, the coefficients in these areas are low. This is to be expected since we would not expect every young male to be tagging nor for every tag to be reported, it is reasonable to expect that several additional resident young males are needed for each increase in graffiti report.

Police stations tend to be located on arterials; however, distance to police stations and distance to arterials show different patterns, and the two were not found to be collinear. The area where the police stations are closest to each other show uniformly low coefficients, perhaps related to their smaller coverage and presumably denser travel patterns by officers of the law.

Commercial zoning had the greatest impact in areas with medium density of young males and low percentage of commercial properties. Residential zoning varies strongly across the city, with almost zero contribution throughout much of the city but a high contribution in specific neighborhoods, such as North Beach and Telegraph Hill.

Fig. 5 shows the residuals for OLS and for GWR. The GWR residuals show less clustering than OLS; this is supported by the Moran's I statistic, which is 0.352 for OLS ($z = 27.45$, $p = 0$) versus 0.042 for GWR ($z = 3.358$, $p = 0.0008$). In addition, the GWR residuals do not show heteroskedasticity. Fig. 6 compares the (log) actual data (Fig. 6a) with the (log) predicted values for graffiti reports, based on the GWR model (Fig. 6b); visually, the pattern of clustering is quite similar to that of the (log) actual data. Fig. 7 shows the actual observed and predicted values; while the overall pattern is very similar and hotspots are accurately predicted, there are some differences in the absolute observed versus predicted values.

The GWR model explains over two-thirds of the variation in graffiti reports in the city of San Francisco. The maps of predicted and observed graffiti reports under the GWR show very similar patterns across the city. Our model overestimates graffiti report densities in high-report areas of the city, while overestimating densities in low- or close-to-zero report areas – although in some low-report areas, we know there to be higher actual incidences of graffiti, such as in industrial areas or parks.

Discussion

Our GWR model explains over two-thirds of the variation in graffiti reports in the city of San Francisco. The maps of predicted and observed graffiti reports under the GWR show very similar patterns across the city. The model accurately predicts the major neighborhoods with a high number of graffiti reports: the centrally-located Mission District, the Hayes Valley-Haight Ashbury-Richmond corridor, the Civic Center area, and Excelsior in the south. Areas with low numbers of graffiti reports were also accurately predicted, such as the industrial areas along the east edge of

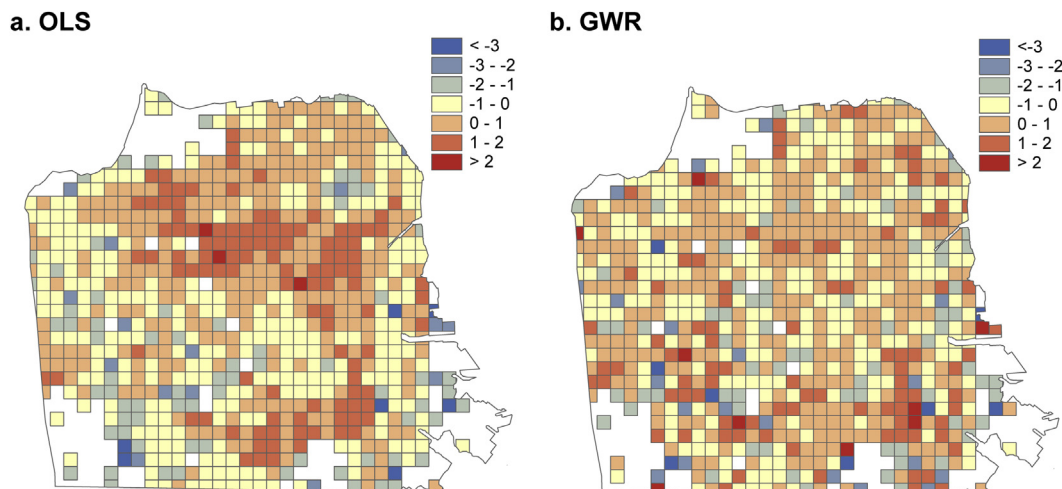


Fig. 5. Standardized Residuals (in Std. Dev.): a. OLS, b. GWR.

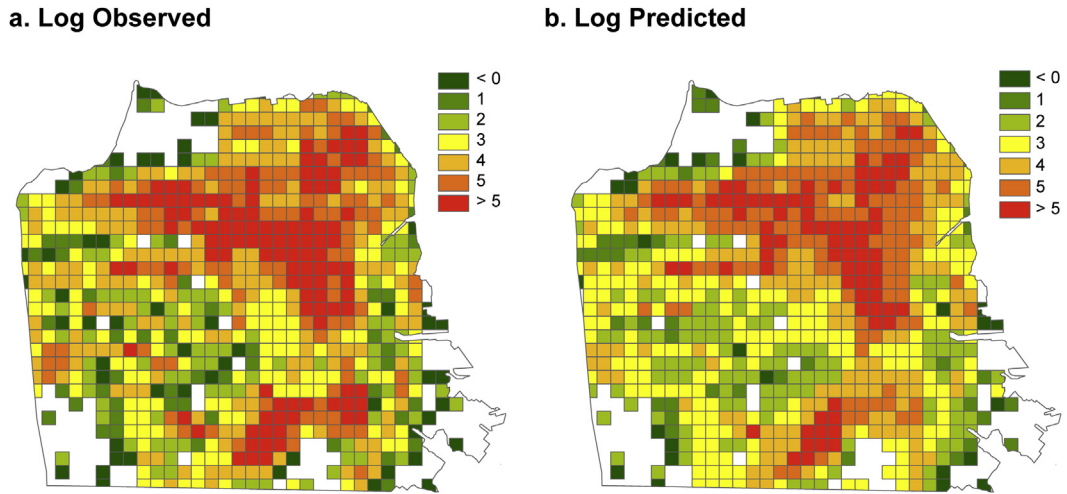


Fig. 6. Graffiti Report Density: a. Log Observed, b. Log Predicted from GWR.

the city, the outer southwest, and high-elevation residential areas in the Twin Peaks area in the central part of the city. The model under-predicts graffiti reports in the residential areas in the north and Portola in the south, a neighborhood with an active anti-graffiti program. Our model overestimates graffiti report densities in low-graffiti areas of the city, while underestimating densities in the highest report locations. This high variability of graffiti reports across the city makes the use of the GWR model necessary.

Commercial zoning had the greatest impact in areas with low percentages of commercial properties, as well as in areas with generally low numbers of graffiti reports. The coefficient for percent commercial increases as residential percentage increases, indicating that neighborhoods with isolated commercial areas (the majority of the city) have a high proportion of graffiti in those commercial areas. Business owners in isolated commercial areas may perceive more of a feeling of threat from graffiti, which may increase the likelihood of reporting. Conversely, in dense commercial zones, percent commercial is less of a factor. Despite a uniform no-tolerance policy for graffiti across the city, it is likely that different business districts, or supervisor districts, may have different relationships to the presence and meaning of graffiti and different approaches to abatement; the effect of these differing relationships on graffiti incidence has been studied and shown in

Sydney (Haworth et al., 2013). High-end businesses in Haight-Ashbury may be vigilant in reporting graffiti, resulting in higher counts for those neighborhoods than one might find in other commercial zones. The rapid gentrification of areas such as the Mission District may invite graffiti as a social response (Shobe & Banis, 2014).

The impact of residential zoning varies strongly across the city, with almost zero effect throughout much of the city but a high contribution in specific neighborhoods, such as North Beach and Telegraph Hill. This may reflect demographic differences in tolerance in these areas. Residential zoning also made a high contribution in the SW section of the city with its much higher than average concentration of public spaces such as parks.

The coefficient for young male density increases in areas where the percent commercial zoning is low; the coefficient for percent residential decreases as young male density increases. This interplay between young male density and percent commercial/residential in the model suggests that at least a medium density of young males is needed for any substantial number of graffiti reports, but also that graffiti writers will travel to commercial areas. If one accepts the premise of graffiti as communication, then writers will not necessarily write where they live, but write in places where their writing will be seen. Graffiti reports also have a strong inverse

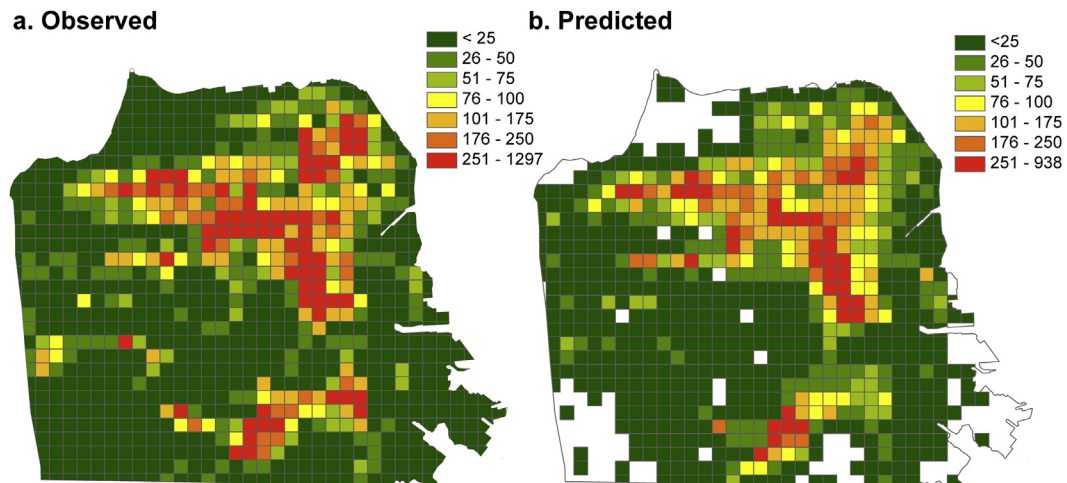


Fig. 7. Graffiti Report Density: a. Observed, b. Predicted from GWR.

relationship with distance to arterials, again indicating that graffiti is more likely to occur in areas where it will be seen.

Police stations tend to be located on arterials; however, distance to police stations and distance to arterials show different patterns, and the two were not found to be collinear. The relationship of graffiti reports to proximity to police stations may seem counter-intuitive at first with large areas of the city having a negative coefficient for distance to police stations. However, more graffiti reports closer to police stations in both commercial and residential areas may result simply from a greater awareness of graffiti in those areas. Coefficients for percent residential decrease as distance to police stations increases while the coefficients for male density increase as distance to police stations increases.

Although the five variables in our GWR model are strong predictors of graffiti reports, the social phenomenon we are investigating is complex, and there are certainly other factors that impact both the creation and reporting of graffiti. One such factor is tolerance of graffiti, and that is not directly measurable. In industrial areas, graffiti is just part of doing business, and reports are almost nonexistent. Income was not a statistically significant variable, but spatial patterns from this analysis suggest that it may still be important. Higher income areas may have lower densities of graffiti but may have a low tolerance for it and report almost all occurrences. Lower income or rapidly gentrifying neighborhoods may have higher densities of graffiti but report a lower proportion of them. The offsetting of these dual processes cannot be tested statistically given available data.

Elevation change was also not statistically significant but once again may still be important, acting mostly as a deterrent to access in some locations such as the steep winding streets of the central Twin Peaks area, and not in others such as Pacific Heights that are within the street grid and closer to commercial areas. On the other hand, Breetzke's finding of the relationship between elevation and burglary risk in Tshwane, South Africa (Breetzke, 2012) may, in fact, be a reflection of other underlying social factors not reflected in his study or relevant to local social processes that do not translate to other locations.

There is no known evidence that the density of, or factors driving the creation of, graffiti in San Francisco are different from (or, for that matter, the same as) those in other major cities in the United States.

Conclusions

Grffiti is a social process that manifests in a location; the specific locations chosen are affected by a variety of social, physical, and demographic factors. By analyzing the geospatial relationship of these locations to measurable factors, we offered a new insight on graffiti as a process than that determined through qualitative studies. As with all social phenomena, the causes are undeniably complex. However, we are encouraged by the level of accuracy of our model, and speculate that several simple variables interacting with each other may in fact drive the majority of this complex behavior — just as simple models of bird flight can predict complex flight patterns.

Our findings are consistent with the interpretation of graffiti as both a form of communication, aimed at a target market of other young males, along with the broken window thesis of graffiti interpreted as a sign of social disorder. These results are complicated by the city's requirement for reporting and abating graffiti, and also by the variable nature of volunteered geographic information used in the study. Since the graffiti reports used here are a volunteered sample, rather than a census of the occurrence of graffiti, what we are discussing is graffiti that people care about — in essence, adding recognition of and the impetus to report graffiti

occurrences to the social processes of graffiti production that we are studying. In addition, using VGI allows us to analyze a much larger geographic area than would be practical if relying on on-the-ground surveys of graffiti. The nature of this VGI may change with time as more (or fewer) people participate, and municipalities learn more about who is reporting and how the data can be used. Haworth et al. (2013) note that graffiti reports do not tell us anything about the qualities of graffiti removed. Opinions about and relationships to graffiti vary widely across space and can change as policies surrounding graffiti abatement change. A spatially-explicit model such as GWR is necessary to capture some of this complexity.

While our analysis focuses on San Francisco, these similar social processes and tensions are present in all major populated places. Performing a GIS-based analysis allows contributing factors to be identified and analyzed, and misconceptions highlighted, such as assumptions about the direct association of graffiti with low income areas. Although such an analysis can never explain such a complex social process as graffiti creation, it does contribute to a more informed discussion of causes of and appropriate responses to graffiti and indicates that solutions other than, or in addition to, abatement may be appropriate. Bringing the tension of these competing forces to light may provide cities with the opportunity to identify and experiment with alternate approaches to graffiti abatement, such as a proposal in Melbourne for varying zones of tolerance (Young, 2010) or Sydney's by-district approach (Haworth et al. 2013), and avoid ineffective monolithic zero-tolerance programs.

References

- Austin, J. (2001). *Taking the train: How graffiti art became an urban crisis in New York City*. Columbia University Press. Retrieved from <http://books.google.com/books?hl=en&lr=&id=0amU10WApNgC&oi=fnd&pg=PA1&dq=Austin,+Joe.+2001.+Taking+the+Train:+How+Graffiti+Art+Became+an+Urban+Crisis+in+New+York+City&ots=11QhMETHGc&sig=1A6NKIMGI8uWyogf-BeQuqycWc>.
- Baird, J., & Taylor, C. (2010). *Ancient graffiti in context*. US: Taylor & Francis.
- Breetzke, G. D. (2012). The effect of altitude and slope on the spatial patterning of burglary. *Applied Geography*, 34, 66–75. <http://dx.doi.org/10.1016/j.apgeog.2011.10.017>.
- Budhathoki, N. R., Nedovic-Budic, Z., & Bruce, B. (2010). An interdisciplinary frame for understanding volunteered geographic information. *Geomatica*, 64(1), 11–26.
- Carey, B. (2006, February 14). *Ancient cave art full of teenage graffiti*. LiveScience.com. Retrieved from <http://www.livescience.com/7028-ancient-cave-art-full-teenage-graffiti.html>.
- Castleman, C. (1982). *Getting up: Subway graffiti in New York*. MIT Press. Retrieved from http://books.google.com/books?hl=en&lr=&id=FFSD-UOKI54C&oi=fnd&pg=PP10&dq=Castleman,+Craig.+1982.+Getting+Up:+Subway+Graffiti+in+New+York&ots=PAs04cmn6C&sig=MNs_cqA7Zfcw97Bog7mdZrYf430.
- Cresswell, T. J. (1992a). *In place/out of place: Geography, ideology and transgression* (Vol. 1). University of Wisconsin–Madison.
- Cresswell, T. J. (1992b). The crucial “where” of graffiti: a geographical analysis of reactions to graffiti in New York. *Environment and Planning D: Society and Space*, 10(3), 329–344. <http://dx.doi.org/10.1068/d100329>.
- DeMeritt, M. (2011, Winter). *Simplifying citizen reporting: Smartphone 311 app enhances government responsiveness*. ArcUser Online. Retrieved from <http://www.esri.com/news/arcuser/0111/citysourced.html>.
- Doran, B. J., & Lees, B. G. (2005). Investigating the spatiotemporal links between disorder, crime, and the fear of crime. *The Professional Geographer*, 57(1), 1–12.
- Dovey, K., Wollan, S., & Woodcock, I. (2012). Placing graffiti: creating and contesting character in inner-city Melbourne. *Journal of Urban Design*, 17(1), 21–41.
- Elwood, S., Goodchild, M. F., & Sui, D. Z. (2012). Researching volunteered geographic information: spatial data, geographic research, and new social practice. *Annals of the Association of American Geographers*, 102(3), 571–590.
- Feick, R., & Roche, S. (2013). Understanding the value of VGI. In *Crowdsourcing geographic knowledge* (pp. 15–29). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-94-007-4587-2_2.
- Ferrell, J. (1995). Urban graffiti: crime, control, and resistance. *Youth & Society*, 27(1), 73–92. <http://dx.doi.org/10.1177/0044118X95027001005>.
- Flanagin, A. J., & Metzger, M. J. (2008). The credibility of volunteered geographic information. *GeoJournal*, 72(3–4), 137–148.
- Fotheringham, A. S., Charlton, M. E., & Brunsdon, C. (1998). Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. *Environment and Planning A*, 30, 1905–1927.

- Gathright, K. (2010, August 3). *City's struggle against graffiti tries rewards, murals and profiling*. SF Public Press. Retrieved March 1, 2012, from <http://sfpublicpress.org/news/2010-08/citys-struggle-against-graffiti-tries-rewards-murals-and-profiling>.
- Glazer, N. (1979). On subway graffiti in New York. *The Public Interest*, 54, 3–11.
- Goodchild, M. F. (2007). Citizens as sensors: the world of volunteered geography. *GeoJournal*, 69(4), 211–221.
- Grira, J., Bédard, Y., & Roche, S. (2010). Spatial data uncertainty in the VGI world: going from consumer to producer. *Geomatica*, 64(1), 61–72.
- Guthrie, R. D. (2005). *The nature of paleolithic art*. University of Chicago Press.
- Halsey, M., & Young, A. (2002). The meanings of graffiti and municipal administration. *Australian & New Zealand Journal of Criminology*, 35(2), 165–186.
- Halsey, M., & Young, A. (2006). "Our desires are ungovernable" writing graffiti in urban space. *Theoretical Criminology*, 10(3), 275–306.
- Haworth, B., Bruce, E., & Iveson, K. (2013). Spatio-temporal analysis of graffiti occurrence in an inner-city urban environment. *Applied Geography*, 38, 53–63.
- Hung, J., Ly, K., & Ngo, V. (2010). *Graffiti in the urban everyday: Comparing graffiti occurrence with crime rates, land use, and socio-economic indicators in Vancouver, BC*. Retrieved from <http://eaves.ca/wp-content/uploads/2010/Vancouver%20Graffiti%20Analysis.pdf>.
- Johnson, P. A., & Sieber, R. E. (2013). Situating the adoption of VGI by government. In *Crowdsourcing geographic knowledge* (pp. 65–81). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-94-007-4587-2_5.
- Kelling, G. L., & Wilson, J. Q. (1982). Broken windows. *Atlantic Monthly*. <http://www.theatlantic.com/magazine/archive/1982/03/broken-windows/304465/>.
- Lachmann, R. (1988). Graffiti as career and ideology. *American Journal of Sociology*, 229–250.
- Ley, D., & Cybriwsky, R. (1974). Urban graffiti as territorial markers. *Annals of the Association of American Geographers*, 64(4), 491–505.
- Mcauliffe, C. (2012). Graffiti or street art? Negotiating the moral geographies of the creative city. *Journal of Urban Affairs*, 34(2), 189–206. <http://dx.doi.org/10.1111/j.1467-9906.2012.00610.x>.
- Megler, V. M., Snodgrass, J., & Blackmore, D. (2011). *Patterns of spatial distribution for reported graffiti in San Francisco*. Portland State University. Retrieved from http://web.pdx.edu/~jduh/courses/geog492w11/Projects/MeglerSnodgrassBlackmore_Graffiti%20v1.pdf.
- Miller, H. J. (2010). The data avalanche is here. Shouldn't we be digging? *Journal of Regional Science*, 50(1), 181–201.
- Miller, I. (2002). *Aerosol kingdom: subway painters of New York city*. University Press of Mississippi.
- Monto, M. A., Machalek, J., & Anderson, T. L. (2013). Boys doing art: the construction of outlaw masculinity in a Portland, Oregon, graffiti crew. *Journal of Contemporary Ethnography*, 42(3), 259–290.
- Sampson, R. J., & Raudenbush, S. W. (2004). Seeing disorder: neighborhood stigma and the social construction of "broken windows." *Social Psychology Quarterly*, 67(4), 319–342.
- San Francisco Public Works Code Article 23. Retrieved from http://www.nograffiti.com/ordinance/san_francisco.htm.
- Schuermans, N., Loopmans, M. P. J., & Vandenebeele, J. (2012). Public space, public art and public pedagogy. *Social & Cultural Geography*, 13(7), 675–682. <http://dx.doi.org/10.1080/14649365.2012.728007>.
- Shobe, H., & Banis, D. (2014). Zero graffiti for a beautiful city: the cultural politics of urban space in San Francisco. *Urban Geography*, 35(4).
- Stockner, T. L., Dutcher, L. W., Hargrove, S. M., & Cook, E. A. (1972). Social analysis of graffiti. *The Journal of American Folklore*, 85(338), 356–366.
- U.S. Census Bureau. (2010). *Census*. Retrieved from <http://www.datasf.org/story.php?title=2010-census-data-for-san-francisco>.
- Young, A. (2010). Negotiated consent or zero tolerance? Responding to graffiti and street art in Melbourne. *City*, 14(1–2), 99–114.