Applied Geography 44 (2013) 143–153

Contents lists available at ScienceDirect

Applied Geography

journal homepage: www.elsevier.com/locate/apgeog

Exploiting autoregressive properties to develop prospective urban arson forecasts by target

Jeffrey P. Prestemon [a,1], David T. Butry [b,*], Douglas S. Thomas [b,2]

[a] Forest Economics and Policy Research, Southern Research Station, USDA Forest Service, Forestry Sciences Laboratory, PO Box 12254, Research Triangle Park, NC 27709, United States

[b] Applied Economics Office, Engineering Laboratory, National Institute of Standards and Technology, 100 Bureau Drive, Mailstop 8603, Gaithersburg, MD 20899-8603, United States

Keywords:
Hotspot
Intentional fire
Economics
Devil's night
Serial arson

A B S T R A C T

Municipal fire departments responded to approximately 53,000 intentionally-set fires annually from 2003 to 2007, according to National Fire Protection Association figures. A disproportionate amount of these fires occur in spatio-temporal clusters, making them predictable and, perhaps, preventable. The objective of this research is to evaluate how the aggregation of data across space and target types (residential, non-residential, vehicle, outdoor and other) affects daily arson forecast accuracy for several target types of arson, and the ability to leverage information quantifying the autoregressive nature of intentional firesetting. To do this, we estimate, for the city of Detroit, Michigan, competing statistical models that differ in their ability to recognize potential temporal autoregressivity in the daily count of arson fires. Spatial units vary from Census tracts, police precincts, to citywide. We find that (1) the out-of-sample performance of prospective hotspot models for arson cannot usefully exploit the autoregressive properties of arson at fine spatial scales, even though autoregression is significant in-sample, hinting at a possible bias-variance tradeoff; (2) aggregation of arson across reported targets can yield a model that differs from by-target models; (3) spatial aggregation of data tends to increase forecast accuracy of arson due partly to the ability to account for temporally dynamic firesetting; and (4) arson forecast models that recognize temporal autoregression can be used to forecast daily arson fire activity at the Citywide scale in Detroit. These results suggest a tradeoff between the collection of high resolution spatial data and the use of more sophisticated modeling techniques that explicitly account for temporal correlation.

Published by Elsevier Ltd.

Introduction

Law enforcement organizations are increasingly using automated crime mapping tools that endeavor to produce short- and medium-term predictions of altered criminal activity (e.g., Bowers, Johnson, & Pease, 2004). These tools have been developed for alternative spatial and temporal scales but have typically not been subject to assessment of forecast accuracy, with few exceptions (e.g., Chainey, Tompson, & Uhlig, 2008; Gorr, 2009). A central challenge with such mapping is that hotspot identification requires abundant data, but at fine spatial or temporal scales such abundance is lacking, so coarse scale data are instead used; yet, it is not clear how this aggregation of data affects forecast results, or the ability to leverage information describing the dynamic process of crime (e.g., Bowers et al., 2004; Johnson, 2013; Lottier, 1938; Prestemon & Butry, 2005). Some modelers report advances in hotspot mapping (e.g., Cohen, Gorr, & Olligschlaeger, 2007); however, these efforts are focused on categories of relatively frequent more serious 'Part I' crimes (United States Department of Justice, 2004) such as robbery, aggravated assault, burglary, larceny, or motor vehicle theft, but rarely on less frequent Part I crimes such as murder (see Groff & McEwen, 2007), rape, and arson, or the many categories of less reliably reported less serious ('Part II') crimes (but see Kakamu, Polasek, & Wago, 2008).[3] Furthermore, while much

* Corresponding author. Tel.: +1 (301) 975 6136.
E-mail addresses: jprestemon@fs.fed.us (J.P. Prestemon), david.butry@nist.gov (D.T. Butry), douglas.thomas@nist.gov (D.S. Thomas).
[1] Tel.: +1 (919) 549 4033.
[2] Tel.: +1 (301) 975 4918.

[3] The Federal Bureau of Investigation's Uniform Crime Reports classify crimes into two categories—more serious, 'Part I', and less serious, 'Part II', crimes. Part I crimes are criminal homicide, forcible rape, robbery, aggravated assault, burglary, larceny-theft, motor vehicle theft, and arson. Part II crimes are other assault, forgery and counterfeiting, fraud, embezzlement, stolen property, vandalism, weapons, prostitution and commercial vice, sex offenses, drug abuse violations, gambling, offenses against the family and children, driving under the influences, liquor laws, drunkenness, disorderly conduct, vagrancy, all other state and local laws not included in Part I or II (traffic laws excluded), suspicion, curfew and loitering laws, and runaways.

0143-6228/$ – see front matter Published by Elsevier Ltd.
http://dx.doi.org/10.1016/j.apgeog.2013.07.015

effort has been devoted to developing static (backward-looking) crime mapping tools, few have been developed that are designed for forecasting—prospective hotspotting. The hotspot maps generated by the static tools are updated on a relatively frequent basis (e.g., weekly or monthly), but little is known about the value of more frequently updated or short time-step reliability. This is in spite of the recognition that such tools have potential high value in terms of tactical police response (e.g., Bowers et al., 2004) and in the planning of a built environment resilient to crime (e.g., Nelson, Bromley, & Thomas, 2001).

In addition to the lack of progress in developing prospective mapping tools for less common crimes at fine temporal scales, there is an under-appreciation of the negative effects of aggregation bias. Broad-scale analysis of crime may obscure patterns found at the micro-scale (Nelson et al., 2001), and the appropriateness of aggregating crimes by type may depend on spatial scale used (Andresen & Linning, 2012). For example, crime maps showing aggregates of Part I crimes do not always recognize subcategories within these crime categories—such as burglaries of commercial targets versus burglaries of residences. Lack of recognition of the within-category heterogeneity may lead to statistical biases and inconsistencies in the estimates of the model parameters embedded in the mapping tools—this is the Modifiable Unit Areal Problem (e.g., see Ratcliffe & McCullagh, 1999). It may also lead to the use of inappropriate (ineffective) mitigation strategies (e.g., Haworth, Bruce, & Iveson, 2013). For example, arson focused on residential structures may have a different amount of temporal, spatial, or spatio-temporal clustering or respond differently to law enforcement efforts compared to vehicular or outdoors arson (e.g., see Groff & McEwen, 2007).

The objective of this research is to evaluate how the aggregation of data across space and target types (residential, non-residential, vehicle, outdoor and other) affects daily arson forecast accuracy for several target types of arson, and the ability to leverage information quantifying the autoregressive nature of intentional firesetting. To do this, we estimate, for the city of Detroit, Michigan, competing statistical models that either recognize or do not recognize potential temporal autoregressivity in the arson counts. The spatial units that we study vary from Census tracts, police precincts, to citywide. We do not vary the temporal unit from daily, although the results of the modeling potentially can be used to design strategies that account for the regular variations in arson frequencies over time (e.g., those linked to days of the week or seasons of the year). Four specific target types for arson are modeled: residential structures, commercial structures, vehicles, and vegetation and outdoor targets (e.g., trash fires). Two aggregations are modeled separately and compared with the individual types: aggregation of structures (residential plus commercial) and aggregation of all arson (all structures plus vehicles plus outdoors and other).

A contribution of this research is that we find that temporal autoregressivity found for smaller spatial units is not beneficially exploited to improve forecast accuracy, whereas this temporal autoregressivity found at large spatial units is beneficially exploited to improve forecast accuracy, compared to alternative forecasting approaches. We conjecture that the lack of additional forecast accuracy provided by the autoregressive models for our smaller spatial units derives from the inappropriateness of the model specification for a count process occurring at low temporal frequency.

The remainder of this paper is organized as follows. First, we describe the arson crime data generating processes for alternative targets, tying these to crime theory. Second, we outline empirical predictive models that may be useful for forecasting arson by target. Third, we apply the estimated predictive models and describe and compare their performance across spatial scales and targets. The paper concludes with recommendations for further research and development of forward-looking crime hotspotting tools that could be useful for law enforcement and fire agencies.

Methods

The Poisson type (count) model specifications we outline in this modeling effort are based on Rational Choice theory (Cornish & Clarke, 1986) as well as elements of wildfire theory. Rational Choice theory derives from an economic framework, such as that outlined by Becker (1968), and behavioral economics, such as that generally described by Wilson and Kelling (1982)—the Broken Windows hypothesis, which has received some support in the empirical literature (e.g., Frazier, Bagchi-Sen, & Knight, 2013; Keizer, Lindenberg, & Steg, 2008; Sampson & Raudenbush, 2004). In the Becker (1968) approach, the prospective criminal compares the benefits of crime commission with the personal costs of crime commission. Among the costs are those associated with being caught, arrested, and receiving a penalty (e.g., fines or jail time) that results in an income loss or other (e.g., psychological) loss for the firesetter. Costs can be direct, associated with each act of crime commission, including the opportunity costs of committing the crime compared to another activity (e.g., leisure, wage earning) and with the material costs for carrying out the crime (e.g., fire starting materials). Costs may also be indirect, connected to the information gathering costs associated with achieving crime success. When considered from an empirical perspective, an economics of crime model could be estimated using actual data on crime commission and predictor variables that include measures of wealth, labor market conditions (e.g., Gould, Weinberg, & Mustard, 2002), income, poverty, law enforcement efforts, and arrest rates (Di Tella & Schargrodsky, 2004).

Following from Rational Choice theory, arson is expected to be clustered in time and space due to increased and/or decreased time dependent costs/benefits (Becker, 1968; Cornish & Clarke, 1986). For example, typical work hours increase the cost of committing arson during a weekday for most individuals; therefore, it would be expected that fewer arson incidents would occur during these time periods, as is illustrated in Fig. 1. Arson is observed to occur in three types of cluster patterns: temporal, spatial, and autoregressive. These clusters occur because costs/benefits change in time and space for large portions of the population, and providing a paradigm to identify variables that can be successfully used to predict arson occurrence. Temporal clusters, for instance, include daily, weekly, and seasonal trends in arson as well as events such as holidays and sporting events. Understanding temporal identifiers

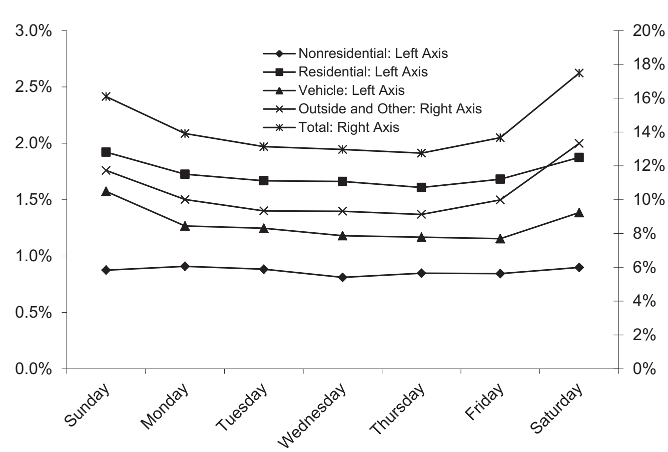

Fig. 1. Percent of total national arson by type by day of week, National Fire Incident Reporting System 2002–2006.

can be useful for predicting the timing of arson occurrence. Spatial clusters include target- and place-driven arson, such as abandoned structures and vehicles, or buildings that have designated purposes, such as schools or churches. Autoregressive clusters include temporally connected incidents such as serial arson and copycat arson.

There are a number of theories that support the use of variables related to the three cluster types as predictors of crime and arson. From the Broken Windows hypothesis, for example, it is suggested that petty crimes may be predictors of more serious crimes, such as arson. That is, petty crimes—a measure of disorder—may identify areas where spatial arson clusters occur. Another perspective is that a prospective criminal weighs the probability of arrest for committing a crime based on recent rates of the same or other kinds of crimes. The Broken Windows hypothesis relating to arson has been supported by recent research findings exploring arson fire patterns in Michigan (Thomas, Butry, & Prestemon, 2011). The Becker (1968) and Wilson and Kelling (1982) ideas can be supplemented from several theories in criminology that have been used to specify or statistically model criminal activity at micro (offender) to macro (population) scales. These include Strain theory (e.g., Cernkovich, Giordano, & Rudolph, 2000), Social Learning theory (e.g., Burgess & Akers, 1966), Social Control theory (Hirschi, 1969; Wiatrowski, Griswold, & Roberts, 1981), and Opportunity theory.

Some of the theories mentioned above support the use of temporal clusters as predictors of arson. Available evidence indicates that these clusters occur at very large as well as relatively small spatial scales. For example, the days surrounding New Year's Day, Halloween, and Independence Day have significant increases in the total national count of arson incidents. Another example, Devil's Night, known as an evening of pranks during the night before Halloween. In Detroit, this night is marked by rampant intentional fire setting throughout the city (Centers for Disease Control, 1997).

The third type of cluster, autoregressive, fits within Rational Choice theory and can be captured in a model using lagged crime occurrences. Such lagged—repeat or near-repeat crimes—have been analyzed empirically by, among others, Short, Brantingham, Bertozzi, and Tita (2010), who base the idea on observed behavior of offenders, tending to commit crimes close to where they reside. Johnson (2013) recently found empirical support for near-repeat behavior using individual offender data for burglaries in Bournemouth and Poole, U.K. Crime has been long known to be clustered in space and time (e.g., Lottier, 1938). Bowers et al. (2004), in recognition of such clustering, suggest explicitly parameterizing autoregressive crime processes in the design of forecasting tools. Butry and Prestemon (2005) and Prestemon and Butry (2005) validated this suggestion in their modeling of daily arson wildfire events in Florida. Although Butry and Prestemon (2005) examined the fine-scale spatio-temporal patterns or autoregressivity of arson wildfires, an untested hypothesis is whether the firesetting processes for each of the main categories of arson found in urban areas also demonstrates spatio-temporal as well as temporal autoregressivity.

Prospective hotspot modeling of crime is relatively new, so there is a limited amount of literature evaluating alternative forecasting approaches (Chainey et al. 2008; Gorr & Harries, 2003). The forecast model building process for crime involves first deciding whether to adopt theories in criminology (e.g., Routine Activities (Cohen & Felson, 1979), Rational Choice (Cornish & Clarke, 1986)), with the potential for constructing elaborate statistical models that recognize the hypothesized drivers of criminal activity at the individual, community, or societal scale. One alternative to adopting these theories is to specify ad hoc but also possibly parsimonious models that have fewer data input requirements and run a lower risk of over-fitting in model estimation. Another alternative is to employ a somewhat "naïve" approach using either temporal lags for each spatial unit to forecast future criminal activity in that unit (see Gorr, Olligschlaeger, & Thompson, 2003) or Geographic Information Systems methods, where recent criminal activity is used as to predict future activity in a spatial domain. Regardless, Gorr et al. (2003, p. 579) conclude that, "…practically any model-based forecast approach is vastly more accurate than current police practices."

Whether statistical, "naïve," or forms of nonparametric models are estimated (e.g., decision trees [e.g., Breiman, Friedman, Olshen, & Stone, 1984]), the modeling framework must recognize the form of the data generation process. Rare crimes at a chosen spatial and temporal scale can be modeled with binary choice models (e.g., logit, probit, decision tree models), while somewhat more frequent criminal activities, whose data generation process faces occasional zero truncation, can be modeled with count models (e.g., the Poisson and its variants). Occurrences of common or aggregate crime categories, or the occurrences of specific crime categories at large spatial and temporal scales, may reveal a data generation process that is essentially continuous and never facing zero-truncation, allowing for least squares model specifications. In the case of arson modeling, this relatively infrequent crime at most spatial and temporal scales calls for a count modeling process.

Empirical models

Crime forecasting models in this analysis are of two primary categories: multivariate parametric and naïve. The latter category is used to benchmark the multivariate models. The multivariate models are divided into two primary subcategories: static and autoregressive. These are both specified as Poisson type count models. The autoregressive model derives from work by Brandt and Williams (2001)—the Poisson autoregressive model of order p, or PAR(p). The naïve models are the random walk and a constant.

The PAR(p) model is (Brandt & Williams, 2001):

$$\Pr\left[y_{j,t} = n \middle| m_{j,t}\right] = \frac{m_{j,t}^{n} e^{-m_{j,t}}}{n!} \tag{1}$$

where $y_{j,t}$ is a random variable representing the count of arson fires in location j in time t, n is the observed count, $m_{j,t}$ is the expected count (conditional mean) of arson fires in location j in time t (a function, to be described later), and e is the exponential operator. Next, let $m_{j,t} = E[y_{j,t}|Y_{j,t-1}]$ be defined as the conditional mean of a linear AR(p) process in location j in time t; that is, conditional on counts in periods (t−1,…, t−p) and a set of hypothesized additional explanatory factors: The expected count can therefore be described as:

$$E\left[y_{j,t} \middle| Y_{j,t-1}\right] = \sum_{i=1}^{p} \rho_{j,i} y_{j,t-i} + \left(1 - \sum_{i=1}^{p} \rho_{j,i}\right) \exp\left(\mathbf{x}'_{j,t} \boldsymbol{\beta}_j\right) \tag{2}$$

where $\mathbf{x}_{j,t}$ is a vector of independent variables (including a constant) for location j, $\boldsymbol{\beta}_j$ is a vector of associated parameters for location j, and the $\rho_{j,i}$'s are the autoregressive parameters for location j. (The static model maintains the same form, as described above, except the autoregressive parameters, $\rho_{j,i}$'s, are restricted to zero.) There is evidence that arson fires in outdoors settings, at least, follow an autoregressive patterns (Butry & Prestemon, 2005; Prestemon & Butry, 2005). The autoregressive terms shown in Eq. (2) would capture either the serial or copycat or unexplained time-varying factors affecting the numbers of fires set, and it recognizes not just the long-standing knowledge of space-time crime clusters but also more recent work focusing on arson. As some analysts have shown (e.g., Mohler, Short, Brantingham, Schoenberg, & Tita, 2011), such patterns for some property crimes have been productively harnessed to improve forecast accuracy.

Variables contained in **x** in Eq. (2) were selected based on Becker (1968) and existing research (Tables 1 and 2). It is apparent that Becker's (1968) expected utility model strictly applies to the decision of a single individual in a single point in time. In a single location over a long time span, where counts of arson fires aggregated, a large segment—many individuals—in the population is faced multiple times with the choice. Across a population, then, variables contained in a statistical model such as indicated in Eq. (1) and Eq. (2), need to account for how variables affecting the decisions of multiple individuals change over the time span considered by the statistical analysis. A second consideration in the variables to include in **x** is in the interest of developing a forward-looking predictive model. Generally, this means that variables used to predict arson fires in day t need to be known with certainty before day t. In practice, we select variables available on day $t-1$ or earlier or include variables that vary in a predictable manner (such as those measuring seasonal variation in direct costs and opportunity costs).

To capture the time-varying direct costs of firesetting, regressors in **x** include a one-day lag of the day's total precipitation, minimum relative humidity, and maximum temperature, all of which are expected to measure the difficulty of igniting fires—weather affects how easily and quickly an outdoor fire can be started. Single-day lags of weather may not precisely measure either the specific conditions of all points within a spatial unit nor be perfectly predictive of the next day's firesetting conditions. Another way to capture this direct cost is through monthly dummy variables, which measure average fire setting conditions.

Opportunity costs of firesetting are partially measured by dummy variables indexing days of leisure, when such costs are expected to be lower: Saturday, Sunday, and non-weekend holidays (the effect of weekend holidays are captured by the Saturday and Sunday dummy variables). Additional variables include the employment rate (a more vigorous labor market is linked to higher opportunity costs), median household income (positively related to opportunity costs), the median value of owner-occupied dwellings (also positively related to opportunity costs), and the population of foreign-born residents (who may perceive higher opportunity costs due to more serious impacts from an arrest and conviction, such as deportation).

Needed are variables indexing factors linked to the probability of being caught and arrested for a crime, and connected to the ideas advanced by Wilson and Kelling (1982) related to urban decay and neighborhood crime vulnerability. For these, we include a seven-day total (day $t-1$ to $t-7$) lag of counts of Part I crimes recorded anywhere in the city; a single dummy variable that accounts for Halloween ("Devil's Night"), equal to 1 for October 25-November 7 and 0 all other days of the year (the date range is broader than what is typically known as "Devil's Night; " however, the lead-up and tail-off of Halloween arson occurs over a longer time period); the median rent of vacant units, which is expected to index absentee landlord efforts to reduce arson occurrence in their properties; and the percent of population in poverty.

Finally in terms of temporally static socioeconomic variables, we include variables intended to explain variation in the aggregate numbers of prospective arsonists in a location: the number of youths ages 5–17, who may be expected to be particularly active in event-related fire setting, such as around Halloween; and the older youth population (ages 16–19) not in school, which also captures the lower opportunity costs perceived by members of this population segment.

To evaluate the effects of spatial aggregation on forecast accuracy and statistical inference, models are estimated at three approximate spatial scales: Census tract, police precinct (Highland Park and Hamtramck aggregated together into a "Precinct 14"), and citywide. Spatio-temporal relationships were explored in both the tract and the precinct models. For tracts, spatio-temporal lags were for up to three days (i.e., the count of arson in day $t-1$, $t-2$, and $t-3$) and for three sets of neighbors (i.e., the count of arson in the adjacent neighboring tracts, the count of arson in the tracts neighboring the adjacent tracts, and the count of arson in the tracts neighboring the second ring of tracts). At the precinct spatial unit, a single spatial lag was built only for contiguous precincts—only the

Table 1
Regressors in the citywide Poisson and Poisson autoregressive count models.

Variable	Data source	Notes
Saturday dummy		
Sunday dummy		
Holiday dummy (non-weekend)		
Devil's Night (15-day period)		October 25–31
Monthly dummies (Jan–Nov)		
Time trend		
Maximum temperature (t−1)	National Weather Service	Degrees F
Maximum relative humidity (t−1)	National Weather Service	Percent
Precipitation (t−1)	National Weather Service	inches × 10
Assault count (t−1)	National Incident Based Reporting System	
Robbery count (t−1)	National Incident Based Reporting System	
Larceny count (t−1)	National Incident Based Reporting System	
Motor vehicle theft count (t−1)	National Incident Based Reporting System	
Structure fire count (t−1)	National Fire Incident System Database	Only lagged for other fire target models
Vehicle fire count (t−1)	National Fire Incident System Database	Only lagged for other fire target models
Residential structure fire count (t−1)	National Fire Incident System Database	Only lagged for other fire target models
Nonresidential structure fire count (t−1)	National Fire Incident System Database	Only lagged for other fire target models
Outdoors and other fire count (t−1)	National Fire Incident System Database	Only lagged for other fire target models
Structure fire count (t−2)	National Fire Incident System Database	Only lagged for other fire target models
Vehicle fire count (t−2)	National Fire Incident System Database	Only lagged for other fire target models
Residential structure fire count (t−2)	National Fire Incident System Database	Only lagged for other fire target models
Nonresidential structure fire count (t−2)	National Fire Incident System Database	Only lagged for other fire target models
Outdoors and other fire count (t−2)	National Fire Incident System Database	Only lagged for other fire target models
Neighbor aggregate arson count (t−1)	National Fire Incident System Database	Neighboring Census tracts or police precinct, depending on spatial scale
Neighbor aggregate arson count (t−2)	National Fire Incident System Database	Neighboring Census tracts or police precinct, depending on spatial scale

Table 2
Average values of variables used in the PAR(p) and Poisson models for three model years. (Note: City values are yearly totals except weather, which are daily averages; Precinct values are yearly precinct averages (i.e., averaged over precincts); Tract values are yearly tract averages (i.e., averaged over tracts)).

Spatial unit and variable	1995	1996	1998
City			
Arson			
All types	5545	5411	5095
Structure	3420	3035	2772
Residential	1315	1248	1280
Non-residential	2105	1787	1492
Vehicle	1791	2100	1999
Outdoor & other	334	276	321
Crime			
Assault	4654	8551	11945
Robbery	224	612	957
Larceny	7868	13127	18700
Vehicle theft	1798	3645	4875
Weather			
Maximum temperature	59	56.9	62.1
Maximum relative humidity	87	89.5	89.7
Precipitation	8.3	7.7	9.4
Precinct			
Arson			
All types	396	387	364
Structure	244	217	198
Residential	94	89	91
Non-residential	150	128	107
Vehicle	128	150	143
Outdoor & other	24	20	23
Socio-economic			
Population ages 5–17	16289	16289	16289
Median rent of vacant units	256	256	256
Population in poverty	69441	69441	69441
Employment rate	0.4757	0.4757	0.4757
Median household income in 1999	27685	27685	27685
Median value of owner-occupied dwellings	59929	59929	59929
Population not schooled (ages 16–19)	1118	1118	1118
Population foreign born	3945	3945	3945
Tract			
Arson			
All types	15	14	12
Socio-economic			
Population ages 5–17	693	693	693
Median rent of vacant units	251	251	251
Population in poverty	2955	2955	2955
Employment rate	0.4916	0.4916	0.4916
Median household income in 1999	29008	29008	29008
Median value of owner-occupied dwellings	58752	58752	58752
Population not schooled (ages 16–19)	48	48	48
Population foreign born	168	168	168

count of arson from neighboring adjacent precincts was included. Boundary truncation did occur in the spatio-temporal lag construction, but we expect that the effects of these boundary truncations were greatest for the precinct models (e.g., neighboring precincts fell outside the city of Detroit, and no data existed). For citywide models, no spatio-temporal lags of arson ignitions were included.

Data sources for the modeling are shown in Table 1, and average values for the data covering the years of statistical estimation are shown in Table 2. The source for the arson fires reported in the city of Detroit and embedded jurisdictions of Hamtramck and Highland Park (Fig. 2) is the National Fire Incident Reporting System (NFIRS), a data collection system directed by the United States Fire Administration (USFA). Detroit was selected because of the frequency of daily arson occurrence. Each arson incident was geocoded, using the reported address, to a Census tract. Further, the time period this analysis was for daily arson fire

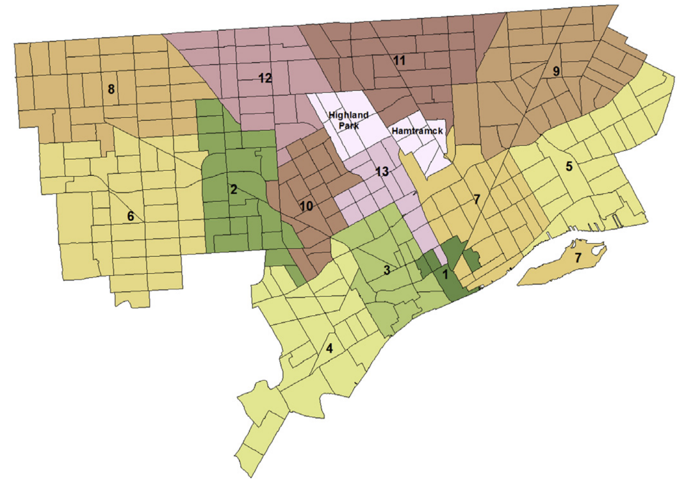

Fig. 2. Police precinct (by color) and Census tract boundaries (black lines) for Detroit, MI.

incidents occurring in 1995, 1996, and 1998. (Data for 1997 were not reported.)

National Incident-Based Reporting System (NIBRS) data on Part I crimes were available only from 1995 and later (Inter-University Consortium for Political and Social Research, 2009). The remaining socio-economic variables were collected from the Inter-University Consortium for Political and Social Research (2006).

Model estimation details

Because of our interest in evaluating the forecast accuracy of alternative models, models are estimated in-sample for the years 1995 and 1996; estimated model parameters are then used to predict arson occurrences each day in 1998. In other words, the training datasets are twice as large as the model validation datasets. The limit to three years of data is because of the lack of data after 1998 for some predictor variables.

It was not possible to estimate exactly the same model at each spatial scale because some variables had temporal but no spatial variation while others had spatial but no temporal variation. Variables with temporal variation but no spatial variation were included in models specified for all three spatial scales: seasonality measures, day and holiday dummies, weather variables, and citywide crime incidents (NIBRS has a limited spatial scale). Variables with spatial but no temporal variation could not be included in a citywide model for arson because they would have been perfectly collinear with the intercept term, although they could be included in models for precincts and tracts, which were pooled across precincts and tracts.

In modeling at the level of the Census tract and the police precinct, we faced additional constraints of model identification related to a lack of temporal variation. Due to collinearity, models could be estimated with dummy variables indexing mean levels of arson fires for each tract or precinct, or they could be estimated socioeconomic variables that varied across tracts or precincts (but not over time). Therefore, two classes of tract-level and precinct-level estimates were made.

As stated, we are also interested in the effects of aggregation across types of arson. For the tract models, then, the primary objectives were to test for fine-scale spatio-temporal relationships and temporal autocorrelation and to evaluate model forecast accuracy. For precinct and citywide models, we are able to evaluate the differences in model forecast accuracy by target.

Table 3
Root mean squared error statistics, averaged over tracts, for out-of-sample forecasts, by model specification.

Model	PAR(p)	Poisson	Constant	Random walk
Tract Dummy & Spatio-Temporal Arson Version	0.2617	0.2266	0.2297	0.3185
Tract Dummy, Spatio-Temporal Arson, & Citywide Crime Lags	0.2607	0.2266	0.2295	0.3186
Socioeconomic Variables	0.2638	0.2266	0.2297	0.3186

Results

Results for tracts

Detroit, Highland Park, and Hamtramck contain 329 Census tracts. Model solutions using GAUSS required compromising model solution time with inference capabilities. Daily observations on arson counts for 1995 and 1996 totaled 731. Up to 7 days of lagged counts were allowed in the models, meaning that conditioning on the first seven days of the first year (1995) left available observations per tract at 724. In the interest of minimizing solution time, we estimated several models with each model including ten tracts (i.e., 7240 observations), but because model initial conditions required a nonzero first observation this often meant that the number of observations was slightly less than 7240. Although 33 groups of ten tracts could be created (the 33rd group with only nine, however), we found that a minimal number of nonzero arson counts, about 100 in the 1995–1996 time period for the pooled ten tracts, were needed to ensure model convergence. This convergence limit in the count modeling also precluded the estimation of tract spatial unit separate arson models by target; instead, all tract spatial unit models were estimated for aggregate arson (all targets).

In summary, we estimated three general sets of model specifications: (i) those with tract dummy variable shifters, month dummies, day dummies (Saturday, Sunday, Holiday, and Devil's Night), and spatio-temporal lags (up to three days); (ii) those identical to (i) but also including lagged crimes of other kinds; and (iii) those that drop the tract dummies, replacing them with socioeconomic variables, and dropping the spatio-temporal lags but including the month and day dummies. In this last specification, the spatio-temporal lags were excluded to facilitate model solution and limit the number of parameters estimated.

Table 3 reports out-of-sample forecast error, as measured by the root mean squared error (RMSE), averaged over tracts, comparing the forecast goodness of fit for all of calendar year 1998 using the parameters estimated using data for 1995 and 1996. Lower values of the RMSE indicate better forecast accuracy. We also evaluated out-of-sample error as measured by the mean absolute error and median absolute error, but there was not an appreciable change in the rankings of the alternative statistical forecast models. These added results are therefore not included in Table 3. The table reports the RMSE for the PAR(p) models; a competing Poisson of the same specification except excluding temporal autoregressive terms (but including the spatio-temporal terms), which is used to evaluate the predictive value of including recent arson history contained in the PAR(p) forecasts; and two parsimonious models: a constant rate model (taking the average count within each tract each day for 1995 and 1996 as the forecast count for tract on each day of 1998), and a random walk (the arson count in day $t-1$ in the tract is the forecast of the count for day t). The forecasts were made one day ahead (t), and iterated using data known on day $t-1$.

The Poisson out-performed the PAR(p), constant rate, and random walk forecast models in every case, according to the RMSE of the forecast out-of-sample. (Also, the constant rate model out-performed the PAR(p) in every case.) These results indicate that a Poisson model is the preferred tool for forecasting among the four forecast tools examined. Across model types, the one including tract dummies appears to forecast very slightly better than other model types. Further, it appears temporal clustering lasts for days, and not much longer. Of course, daily forecasts increase operational challenges, as they require data from the previous day. Results of model parameter significances and signs for aggregate arson (not shown) indicate that the most consistent statistically significant relationships between arson and potential explanatory variables are the temporal and spatio-temporal lagged components, validating statistically the presence of space-time clusters of arson in Detroit.

Results for police precincts

Precinct hotspot models were estimated for aggregate arson as pooled models, in the same general way as in the tracts approach, and they were also estimated by target but with a limited set of predictors. For aggregate arson, we estimated five general models: (i) a simple model that included only daily and monthly dummies, a time trend, weather variables, and police precinct dummies; (ii) a model of the same form as (i) but replacing the precinct dummies with socioeconomic variables; (iii) a model of the same form as (i) but replacing the precinct dummies with spatio-temporal lags of counts of arson in neighboring precincts; (iv) a model of the same form as (i) but dropping precinct dummies; and (v) a model the same form as (ii) but also including lagged counts of some crime variables. All five model types included AR terms when estimating a PAR(p) model. As in the tract models, these were estimated using data for 1995 and 1996 and then used to forecast arson in 1998. Across the 14 precincts, total observations available were 10107.

Table 4 presents summary results of the five types of precinct models, both in-sample (1995–1996) and out-of-sample (1998) forecast errors (RMSE), for the PAR(4) whose parameter estimates are reported in the upper part of the table and for an identically

Table 4
Precinct PAR(p) model statistics, pooled models.

	Precinct dummy model (i)	Socioeconomic variables model (ii)	Spatio-temporal lag model (iii)	Seasonal/autoregressive model (iv)	Socioeconomic/Crime variable model (v)
Log-Likelihood	−13191.47	−13288.24	−13491.12	−13497.56	−13284.43
Observations	10107	10107	10107	10107	10107
RMSE In-Sample	1.2184	1.2646	1.2490	1.2513	1.2528
RMSE Out-of Sample	1.1483	1.1677	1.1793	1.1776	1.1552
RMSE Poisson In-Sample	1.1158	1.1482	1.3226	1.3239	1.1571
RMSE Poisson Out-of-Sample	1.0927	1.0986	1.2319	1.2319	1.0965
RMSE Constant In-Sample	1.3202	1.3423	1.3423	1.3423	1.3423
RMSE Constant Out-of-Sample	1.2447	1.2447	1.2447	1.2447	1.2447
RMSE Random Walk In-Sample	1.5359	1.5662	1.5662	1.5662	1.5662
RMSE Random Walk Out-of-Sample	1.5032	1.5032	1.5032	1.5032	1.5032

Table 5
Precinct PAR(p) model statistics of arson by target, time series-cross sectional models with cross-sectional intercept shifters.

	All structures	Vehicles	Outdoors and other	Residential	Non-residential
Log-Likelihood	−10482.90	−7981.09	−2300.45	−6580.07	−7797.43
Observations	10107	10005	9989	10107	9869
RMSE In-Sample	1.0064	0.7172	0.2853	0.5930	0.7528
RMSE Out-of Sample	0.9018	0.7486	0.2823	0.6024	0.6662
RMSE Poisson In-Sample	0.8592	0.6414	0.2526	0.5155	0.6748
RMSE Poisson Out-of-Sample	0.8006	0.6885	0.2507	0.5269	0.6019
RMSE Constant In-Sample	0.9648	0.6935	0.2578	0.5324	0.7458
RMSE Constant Out-of-Sample	0.8703	0.7258	0.2540	0.5410	0.6427
RMSE Random Walk In-Sample	1.1859	0.8965	0.3560	0.7256	0.9265
RMSE Random Walk Out-of-Sample	1.0947	0.9510	0.3139	0.7393	0.8146

specified non-autoregressive count model (the Poisson), a constant rate model, and a random walk model. The results show that the PAR(4) models have somewhat different in-sample goodness-of-fit, ranging from 1.218 (model i) to 1.265 (model ii). Out-of-sample fitness was slightly better, with an RMSE range between 1.148 and 1.179. Comparing the out-of-sample RMSE, the Poisson and PAR(4) models outperformed the constant and random walk models for all five models. The Poisson version of specification (i) performed the best with an RMSE of 1.093. The Poisson out-performed the PAR(p) model for specifications (i), (ii), and (v) while PAR(p) performed better in specifications (iii) and (iv). This better out-of-sample performance by the Poisson in these three cases may be linked to its lower bias. By design, the Poisson model has zero in-sample bias, minimizing the variance while constraining prediction errors to sum zero. The PAR(p) model, on the other hand, is not designed to be unbiased, rather to simply minimize variance. So while a competing PAR(p) model might have a lower variance in-sample due to inclusion of autoregressive terms, its higher potential bias leads to a worse forecast compared to the Poisson, particularly if the autoregressive parameter estimates are biased upward, a classic bias-variance tradeoff brought about by over-fitting.

In the interest of evaluating whether aggregation across arson targets leads to lower overall model out-of-sample forecast accuracy, we estimated a version of model (i) for each of the arson targets. In-sample and out-of-sample forecast performance (for 1998) in these by-target arson PAR(p) models cannot be compared across targets, but they can be compared within targets across model types. Generally, we see that the Poisson out-performs all competing models (see Table 5). In fact, the PAR(p) models are all out-performed by the constant rate models. In summary, it appears that the PAR(p) models, which account for the autoregressive patterns in arson fires, cannot be exploited to improve forecast accuracy for arson at the precinct spatial unit, either in aggregate or by specific target.

A last look at forecast performance can focus on how a by-target total arson forecast generated by summing the forecasts across target models compares to a total arson forecast generated by an aggregate arson model. The RMSE out-of-sample calculated by summing across the forecasts produced by the four disaggregated target precinct models reported in Table 5 is 1.30. The RMSE out-of-sample for an aggregate arson model (model (i)) at the precinct spatial unit for the PAR(p) model is 1.15.[4] In summary, forecasts of aggregate arson by precinct using a PAR(p) specification are best done with an aggregate arson model rather than by target type.

[4] The 1998 data generally reveal lower overall counts of arson, which tends to reduce model forecast variance across all model types.

The final spatial unit of our models, citywide, offers the only remaining opportunity for the PAR(p) to outperform static Poisson, constant rate, or random walk models in the forecasting effort.

Results of citywide models

Citywide models are estimated in a manner similar to the precinct spatial unit models. Model structures are the same as those tested for the precincts and tracts—alternately including and excluding lagged citywide crime, including month dummies and day dummies to capture regular cyclical patterns, but dropping the essentially static socioeconomic variables and possessing no location dummies (i.e., precincts, tracts). The autoregressive structure was also varied.

Model in-sample prediction and out-of-sample forecast performance, as measured by the RMSE, shows that the PAR(p) models out-performed not only the constant rate and random walk models but also the static Poisson model. This occurred across all model structures (see Table 6).

A final set of models was developed with estimates up to a PAR(7) model without a Saturday dummy, maximum temperature in day $t-1$, and crime variables, as these variables were not found to be significant. One set of models was estimated with by-type lagged counts of arson for each arson type and one set of models was estimated without these lags. As seen in the out-of-sample RMSE estimates in Tables 7 and 8, the best performing model for structure fires is the PAR(p) model that includes target lags. For vehicle and non-residential structure fires, the PAR(p) that does not include target lags performs best. For outdoor and other fires, the Poisson that does not include target lags is the best performing, while for residential structure fires it is the Poisson that includes target lags. Although PAR(p) models outperformed competing Poisson models, this was only true for PAR(p) models of all structures, vehicles, and nonresidential structures. The additional by-type lags worked to improve out-of-sample forecast performance for total structure fires and residential structure fires for both the Poisson and the PAR(p) models. In summary, we can conclude that autoregressive components included in the citywide model do enhance forecast performance out-of-sample relative to static models but that including type lags worsens out-of-sample performance for some arson types while improving it for others. Apparent statistical correlations found with in-sample model estimates were data-dependent, not helping to explain future arson occurrences. This last result is evidence of an effect of over-fitting leading to a bias-variance tradeoff; in this case, employing statistically significant type lags worsened forecasts for some arson types.

Forecasts for by-target citywide arson forecast models can be compared in their performance with a model estimated for citywide aggregate arson. We calculated the RMSE out-of-sample for a PAR(p) and a Poisson model specification using aggregate arson and one using the sum of the forecasts for vehicles, residential

Table 6
Citywide PAR(p) model statistics, aggregate arson.

	PAR(6), no lagged crime	PAR(6), with lagged crime	PAR(1), no lagged crime	PAR(1), parsimonious
Log-Likelihood	−2143.85	−2139.57	−2176.78	−2142.18
Observations	725	725	725	725
RMSE In-Sample	4.7862	4.7718	4.7876	4.7879
RMSE Out-of Sample	4.3757	4.4239	4.3813	4.3728
RMSE Poisson In-Sample	4.8096	4.7949	4.8178	4.8180
RMSE Poisson Out-of-Sample	4.5997	4.6204	4.5143	4.5141
RMSE Constant In-Sample	5.7583	5.7583	5.7583	5.7583
RMSE Constant Out-of-Sample	4.8505	4.8505	4.8505	4.8505
RMSE Random Walk In-Sample	6.4381	6.4381	6.4381	6.4381
RMSE Random Walk Out-of-Sample	5.7795	5.7795	5.7795	5.7795

structures, nonresidential structures, and outdoors and other, as seen in the last column of Table 9. The PAR(p) model outperformed the Poisson for both of the citywide models, and the summed target model outperformed the aggregate arson model; thus, at the citywide spatial unit, a summed target model slightly outperforms an aggregate model.

A further examination on the impact of spatial aggregation can be done by comparing summed precinct models to citywide models, which is also presented in Table 9. The table presents a PAR(p) version and Poisson version by target type, including one summed across all targets and one for aggregate arson. The citywide PAR(p) model outperformed the Poisson models and the PAR(p) summed precinct model for all but the outdoors and other fire model and the residential fire model. The Poisson summed precinct model performed better than the others for outdoors and other fires while the Poisson citywide model performed better for residential fires. Thus, in all but one case, the citywide model performed better than the summed precinct model. In terms of total arson forecasts, the summed target model outperformed the aggregate arson model for both the PAR(p) and Poisson citywide models. The best performing total arson forecast overall was the citywide PAR(p) summed across targets.

Discussion

Development of prospective hotspotting models requires assessment of out-of-sample performance, the conditions in which any hotspotting tool used by a fire or law enforcement organization would be expected to operate. Model developers have rarely reported out-of-sample forecast accuracy for any type of crime expected to be a part of a law enforcement tool. From another perspective, little attention has apparently been paid to the implications of aggregating crime subtypes within Part I or Part II crimes that are the forecast units of import. A comparison of disaggregated models with the aggregated models can provide a clue as to the effects of such aggregation. Thirdly, crime hotspot tools are often modeled at fine spatial scales, which law enforcement organizations could find more useful for tactical responses to crime outbreaks. However, forecast performance at such fine spatial scales has not been widely reported, as far as we can tell. Finally, very little effort has been put into developing prospective hotspot tools for arson, even though arson in outdoor settings has demonstrated autoregressive properties that would seem to make it possible to develop forecasts with increased accuracy.

Our model results have several implications relevant to the possible limitations of previous forecasting efforts. First, we find that the out-of-sample performance of prospective hotspot models for arson in Detroit cannot usefully exploit the autoregressive properties of arson at fine spatial scales, even though autoregression is significant in-sample. For example, the tract level Poisson model, which excluded temporal autoregressive terms, outperformed the PAR(p) model (see Table 3). Similar results occurred at the precinct level where the Poisson model, outperformed the PAR(p) model for specifications (i), (ii), and (v), while the PAR(p) performed better in specifications (iii) and (iv) (see Table 4). This lack of benefit provided by explicit recognition of temporal autoregressivity at the tract or precinct spatial scales is evidence of a classic bias-variance tradeoff, at least in the context of our chosen Poisson autoregressive model, the PAR(p). Here, we find that models estimated with temporal autoregressive terms may have lower variance over the in-sample data but higher variance out-of-sample, compared to a competing static Poisson. Since the PAR(p) model outperforms all the other aggregate arson models at the citywide spatial unit (see Table 6), we conclude that the autoregressive terms do reduce out-of-sample forecast errors with larger spatial units. This leads us to conclude that prospective arson hotspot tools for cities may be best employed to develop citywide "alerts," leaving the finer spatial targeting decisions to police staff, who have the greatest familiarity with the phenomenon of fine-scale arson patterns within their cities. Nevertheless, tactical responses by local police to these alerts could still be aided by additional graphical representations. For example, Fig. 3 illustrates

Table 7
Citywide PAR(p) parsimonious model statistics, by target and structure aggregate, without target lags.

	Structure fires	Vehicle fires	Outdoors and other fires	Residential structure fires	Non-residential structure fires
Log-Likelihood	−1888.27	−1681.81	−868.06	−1515.73	−1678.18
Observations	725	725	721	725	725
RMSE In-Sample	3.4086	2.6240	1.0299	2.1696	2.6410
RMSE Out-of Sample	3.0615	2.6659	1.1679	1.9938	2.3235
RMSE Poisson In-Sample	3.4310	2.5569	0.9347	1.9389	2.6557
RMSE Poisson Out-of-Sample	3.1205	2.7648	1.1350	1.9134	2.3592
RMSE Constant In-Sample	3.9077	2.8313	1.0067	1.9830	3.1225
RMSE Constant Out-of-Sample	3.4495	2.8179	1.1424	1.9457	2.6786
RMSE Random Walk In-Sample	4.6778	3.5263	1.3193	2.7653	3.5931
RMSE Random Walk Out-of-Sample	4.1827	3.7469	1.5394	2.6320	3.1319

Table 8
Citywide PAR(p) parsimonious model statistics, by target and structure aggregate, including target lags.

	Structure fires	Vehicle fires	Outdoors and other fires	Residential structure fires	Non-residential structure fires
Log-Likelihood	−1888.31	−1675.02	−865.07	−1517.24	−1678.40
Observations	725	725	721	725	725
RMSE In-Sample	3.4187	2.6017	1.0299	2.1849	2.6374
RMSE Out-of Sample	3.0352	2.7432	1.1750	1.9910	2.3509
RMSE Poisson In-Sample	3.4175	2.5428	0.9355	1.9340	2.6371
RMSE Poisson Out-of-Sample	3.1031	2.8049	1.1404	1.9024	2.3704
RMSE Constant In-Sample	3.9077	2.8313	1.0067	1.9830	3.1225
RMSE Constant Out-of-Sample	3.4495	2.8179	1.1424	1.9457	2.6786
RMSE Random Walk In-Sample	4.6778	3.5263	1.3193	2.7653	3.5931
RMSE Random Walk Out-of-Sample	4.1827	3.7469	1.5394	2.6320	3.1319

Table 9
Out-of-Sample RMSE for Citywide Forecast: by Target, Spatial Scale, and Model Type (weather only models).

	All structures	Vehicles	Outdoors and other	Residential	Non-residential	Summed across targets	Aggregate arson
PAR(p) models							
From summed precinct models	5.3319	3.5734	1.2995	2.8697	3.4385	8.3795	5.6755
From citywide models	2.9872	2.5932	1.0980	1.9942	2.2659	4.1668	4.2179
Poisson models							
From summed precinct models	3.0687	2.7146	1.0621	1.8912	2.3166	4.4318	4.4201
From citywide models	3.0253	2.6642	1.0659	1.8886	2.2837	4.3372	4.3466

locations of spatio-temporal arson hotspots can be defined based on historical data. In this illustration, based on Getis-Ord Gi* statistics of significant spatio-temporal clustering (Getis & Ord, 1992), using the space as well as the time distance between arson fires, statistically significant historical hotspots can be used to identify where to redirect law enforcement efforts at fine spatial scales when a citywide temporal forecast model indicates that higher than average rates of arson are predicted for the next day. Such a clustering representation can be seen as one way to overcome the difficulties we encountered in our attempts to identify an accurate count data forecast model that could have produced an analogous spatio-temporal hotspot map.

Second, it is clear that aggregation of arson across reported targets can yield a model that differs from by-target models. The differences lead to different qualities of forecasts, and this can be traced to the different temporal autoregressive patterns for each type. As seen in Fig. 4, the impulse response varies significantly for each target type. An impulse response function provides a forecast of the deviation in expected arson rates (i.e., a change from 100%) from an unexpected increase (shock) in daily arson activity. One fire shock lasts longer and has a larger impact on residential and outdoor and other fires than it does on structure, vehicle, or non-residential structure fires. The different patterns do not provide improved aggregate arson forecasts at the precinct levels; however, they do provide slightly improved citywide forecasts, as documented in Table 9. In short, law enforcement organizations interested in developing arson forecasts for specific types could do so to develop citywide alerts by arson target. These citywide forecasts could be summed to generate citywide alerts for arson in aggregate, which would be better than an alert developed from the aggregate arson figures.

Finally, spatial aggregation tends to increase forecast accuracy for Detroit arson. For all but one model in Table 9, the citywide model outperformed the summed precinct model. This was true for four of the five target types along with the summed target model and aggregate model, suggesting that there is a clear tradeoff between fine spatial scale and forecast accuracy. This further supports the idea that prospective arson hotspot tools for cities may be best employed to develop citywide "alerts."

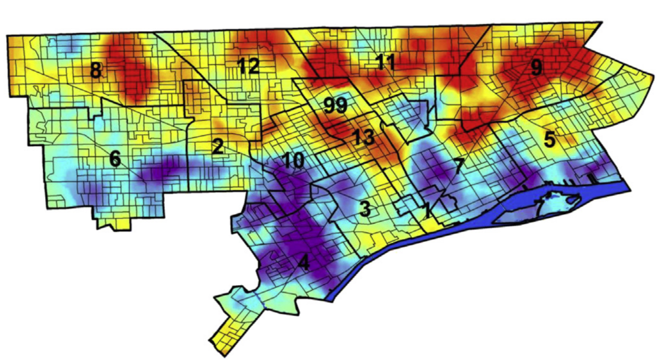

Fig. 3. Spatio-Temporal Arson Clusters in Detroit over 1995, 1996, and 1998, with areas of positive spatio-temporal correlation shown in red and negative correlation in purple. (Note: Police precincts identified numerically, 1–13, with 99 corresponding to Highland Park and Hamtramck.)

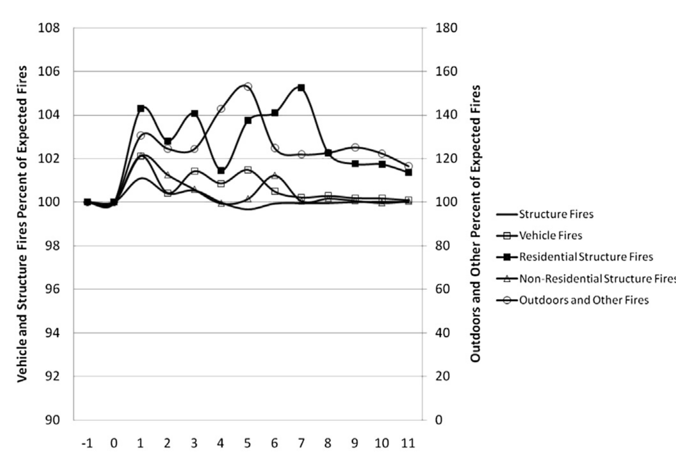

Fig. 4. Impulse response functions for arson by target—percent of fires compared to expected, by day.

In a number of instances, prospective arson hotspot models developed using Poisson Autoregressive models outperform less sophisticated non-autoregressive and even somewhat naive forecast algorithms at fine spatial scales (see Tables 3–5). We have indicated that part of the reason for this worse performance of the PAR(p) at some spatial scales is due to a classic bias-variance tradeoff, brought about by over-fitting involving the autoregressive terms in the model estimation stage. Additionally, we offer the conjecture that, at least for Detroit, the over-fitting results from modeling a count process with a low expected value at fine spatial and temporal scales. Operationally, in our data, much of the additional forecast error produced by PAR(p) models at tract and precinct spatial units is traced to over-predictions (forecasts) following single arson events; at fine scales, outbreaks were over-inferred. Perhaps a better performing autoregressive count model would more accurately contend with the high frequency of zeros and ones. While this could be a "zero-inflated" autoregressive Poisson (adjusted to accommodate over-occurrence of ones as well), potentially fruitful future research could develop models of crime or arson outbreaks, which focus especially on modeling occurrences of multiple incidents in short time periods. Alternatively, perhaps the more sophisticated autoregressive models are better at forecasting outbreaks rather than regular seasonal and spatial patterns of arson, similar to the distinction made between forecasting "exceptions" and "ordinary conditions" described in Gorr (2009). Finally, it's unclear how much of the results are specific to Detroit, which has experienced rapid abandonment in recent years, and how much are generalizable to other cities.

Conclusion

This research evaluated how the aggregation of data across space and target types affects forecast accuracy of arson incidents in the city of Detroit, Michigan, and whether statistical models using different data aggregates can take advantage of information on recent arson activity. Four specific target types for arson were modeled: residential structures, commercial structures, vehicles, and vegetation and outdoor targets. Two aggregations were modeled separately and compared with the individual types: aggregation of structures (residential plus commercial) and aggregation of all arson (all structures plus vehicles plus outdoors and other).

We find that (1) the out-of-sample performance of prospective hotspot models for arson in Detroit cannot usefully exploit the autoregressive properties of arson at fine spatial scales, even though autoregression is significant in-sample, which might be revealing a classic bias-variance tradeoff due to over-fitting of the temporal autoregressive components at fine spatial scales; (2) aggregation of arson across reported targets can yield a model that differs from by-target models; (3) spatial aggregation of data tends to increase forecast accuracy of arson, at least for the city of Detroit; and (4) arson forecast models that recognize temporal autoregression can be used to forecast daily arson fire activity at the Citywide scale in Detroit, which means that prospective hotspot mapping can utilize recent firesetting activity to predict future activity. We conjecture that the lack of additional forecast accuracy provided by the autoregressive models for our smaller spatial units derives from the inappropriateness of the model specification for a count process occurring at low temporal frequency, leading to autoregressive model biases.

The results suggest that additional tradeoffs also exist. One such tradeoff is between collecting high resolution spatial data and the use of more sophisticated modeling techniques accounting for temporal correlation, brought about by lower resolution spatial data. Another modeling trade-off has to do with the utility of resulting prospective hotspot maps. Models specified at high spatial aggregation may have lower overall forecast variance compared to those estimated at finer spatial scales. But the use of these more aggregated models is more limited, inferring the need to change aggregate law enforcement readiness, for example. Likewise, modeling arson with disaggregated target data can reveal divergent causal and correlative driving factors, but their forecast ability may be weaker compared to models that aggregate across classes of crime.

References

Andresen, M. A., & Linning, S. J. (2012). The (in)appropriateness of aggregating across crime types. *Applied Geography, 35*, 275–282.

Becker, G. S. (1968). Crime and punishment: an economic approach. *Journal of Political Economy, 76*, 169–217.

Bowers, K. J., Johnson, S. D., & Pease, K. (2004). Prospective hot-spotting: the future of crime mapping? *British Journal of Criminology, 44*(5), 641–658.

Brandt, P. T., & Williams, J. T. (2001). A linear Poisson autoregressive model: the Poisson AR(p) model. *Political Analysis, 9*, 164–184.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Belmont, CA: Wadsworth and Brooks.

Burgess, R., & Akers, R. L. (1966). A differential association-reinforcement theory of criminal behavior. *Social Problems, 14*, 363–383.

Butry, D. T., & Prestemon, J. P. (2005). Spatio-temporal wildland arson crime functions. In *Proceedings of the Annual Meeting of the American Agricultural Economics Association, 26–29 July 2005, Providence, Rhode Island*.

Center for Disease Control. (1997). Urban community intervention to prevent Halloween arson – Detroit, Michigan, 1985–1996. *MMWR Weekly, 46*(14), 299–304.

Cernkovich, S. A., Giordano, P. C., & Rudolph, J. L. (2000). Race, crime, and the American dream. *Journal of Research in Crime and Delinquency, 37*(2), 131–138.

Chainey, S., Tompson, L., & Uhlig, S. (2008). The utility of hotspot mapping for predicting spatial patterns of crime. *Security Journal, 21*, 4–28.

Cohen, L. E., & Felson, M. (1979). Social change and crime rate trends: a routine activity approach. *American Sociological Review, 44*, 588–605.

Cohen, J., Gorr, W. L., & Olligschlaeger, A. M. (2007). Leading indicators and spatial interactions: A crime - forecasting model for proactive police deployment. *Geographical Analysis, 39*(1), 105–127.

Cornish, D., & Clarke, R. V. (1986). *The reasoning criminal: Rational choice perspectives on offending*. New York: Springer-Verlag.

Di Tella, R., & Schargrodsky, E. (2004). Do police reduce crime? Estimates using the allocation of police forces after a terrorist attack. *American Economic Review, 94*, 115–133.

Frazier, A. E., Bagchi-Sen, S., & Knight, J. (2013). The spatio-temporal impacts of demolition land use policy and crime in a shrinking city. *Applied Geography, 41*, 55–64.

Getis, A., & Ord, J. K. (1992). The analysis of spatial association by use of distance statistics. *Geographical Analysis, 24*(3), 189–206.

Gorr, W. L. (2009). Forecast accuracy measures for exception reporting using receiver operating characteristic curves. *International Journal of Forecasting, 25*, 48–61.

Gorr, W., & Harries, R. (2003). Introduction to crime forecasting. *International Journal of Forecasting, 19*, 551–555.

Gorr, W., Olligschlaeger, A., & Thompson, Y. (2003). Short-term forecasting of crime. *International Journal of Forecasting, 19*, 579–594.

Gould, E. D., Weinberg, B. A., & Mustard, D. B. (2002). Crime rates and local labor market opportunities in the United States: 1979–1997. *The Review of Economics and Statistics, 84*, 45–61.

Groff, E. R., & McEwen, T. (2007). Integrating distance into mobility triangle typologies. *Social Science Computer Review, 25*, 210–238.

Haworth, B., Bruce, E., & Iveson, K. (2013). Spatio-temporal analysis of graffiti occurrence in an inner-city urban environment. *Applied Geography, 38*, 53–63.

Hirschi, T. (1969). *Causes of delinquency*. Berkeley: University of California Press.

Inter-University Consortium for Political and Social Research (ICPSR). (2006). *Census of population and housing*. http://www.icpsr.umich.edu/icpsrweb/ICPSR/ Accessed 02.12.08.

Inter-University Consortium for Political and Social Research (ICPSR). (2009). *National incident based reporting system (NIBRS)*. http://www.icpsr.umich.edu/icpsrweb/ICPSR/ Accessed 05.02.09.

Johnson, D. (2013). The space/time behaviour of dwelling burglars: finding near repeat patterns in serial offender data. *Applied Geography, 41*, 139–146.

Kakamu, K., Polasek, W., & Wago, H. (2008). Spatial interaction of crime incidents in Japan. *Mathematics and Computers in Simulation, 78*(2–3), 276–282.

Keizer, K., Lindenberg, S., & Steg, L. (2008). The spreading of disorder. *Science, 322*(5908), 1681–1685.

Lottier, S. (1938). Distribution of criminal offenses in sectional regions. *Journal of Criminal Law and Criminology, 29*(3), 329–344.

Mohler, G. O., Short, M. B., Brantingham, P. J., Schoenberg, F. P., & Tita, G. E. (2011). Self-exciting point process modeling of crime. *Journal of the American Statistical Association, 106*(493), 100–108.

Nelson, A. L., Bromley, R. D. F., & Thomas, C. J. (2001). Identifying micro-spatial and temporal patterns of violent crime and disorder in the British city centre. *Applied Geography, 21*, 249–274.

Prestemon, J. P., & Butry, D. T. (2005). Time to burn: modeling wildland arson as an autoregressive crime function. *American Journal of Agricultural Economics, 87*(3), 756–770.

Ratcliffe, J. H., & McCullagh, M. J. (1999). Hotbeds of crime and the search for spatial accuracy. *Journal of Geographical Systems, 1*(4), 385–398.

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.

Short, M. B., Brantingham, P. J., Bertozzi, A. L., & Tita, G. E. (2010). Dissipation and displacement of hotspots in reaction-diffusion models of crime. *Proceedings of the National Academy of Sciences, 107*(9), 3961–3965.

Thomas, D. S., Butry, D. T., & Prestemon, J. P. (2011). Enticing arsonists with broken windows and social disorder. *Fire Technology, 47*(1), 255–273.

United States Department of Justice. (2004). *Uniform crime reporting handbook*. Clarksburg, WV: Federal Bureau of Investigation.

Wiatrowski, M. D., Griswold, D. B., & Roberts, M. K. (1981). Social control theory and delinquency. *American Sociological Review, 46*(5), 525–541.

Wilson, J. Q., & Kelling, G. L. (1982). The police and neighborhood safety: broken windows. *Atlantic, 249*(3), 29–38.