GeoDa Center for Geospatial Analysis and Computation

Arizona State University

Sex Offenders And Residential Location: A Predictive Analytical Framework

Elizabeth A. Mack and Tony H. Grubesic

Working Paper Number 2010-03

Abstract

Despite the growing body of research dealing with sex offenders and the collateral consequences of legislation governing their post release movements, a complete understanding of the residential choices of registered sex offenders remains elusive. The purpose of this paper is to introduce a predictive analytical framework for determining which demographic and socioeconomic factors best forecast the residential choices of convicted sex offenders. Specifically, using a derived index of social disorganization (ISDOR) and a commercial geographic information system (GIS), we implement both linear statistical and non-linear data mining approaches to predict the presence of sex offenders in a community. The results of this analysis are encouraging, with nearly 75% of registered offender locations predicted correctly. The implications of these approaches for public policy are discussed.

Keywords: Sex Offenders; Social Disorganization; GIS; Data Mining; Residence Restrictions; Offender Tracking

Word Count: 6040

Introduction

Since the mid-1990s, a series of federal, state and local laws governing the post release movements of sex offenders have been enacted. Although the legislation is somewhat varied in scale and scope, attempts to manage sex offenders generally revolve around the establishment and implementation of registration laws, community notification laws, and residence restriction laws. The United States Department of Justice (USDOJ, 2008) suggests registration and community notification laws serve two important purposes. First, these policies help local law enforcement agencies track offenders' whereabouts upon release from correctional facilities. Second, they are intended to discourage registrants from perpetrating additional sex crimes via increased levels of visibility within the community. The most recent federal sex offender legislation in the United States, the *Adam Walsh Child Protection and Safety Act* (2006), requires all fifty states to implement and maintain registration systems for convicted offenders. At the state and local levels, offender residence restrictions are the most common type of management strategy. These laws are designed to minimize potential interaction between offenders and children and are typically implemented around sensitive facilities where children congregate, such as schools, bus stops and parks. Restriction distances typically range from 500 to 3,000 ft. in the United States. As of 2007, 27 states had implemented residency restriction laws.

Not surprisingly, the flurry of sex offender legislation passed during the past two decades is spurring inquiries regarding the impacts (both intentional and unintentional) of these laws. Research in this domain can be categorized into five specific areas (Mustaine et al., 2006, 179): 1) reviews of registry components and registrants' characteristics; 2) evaluations of impacts on recidivism; 3) monitoring of implementation of registries; 4) collateral consequences of registration for registrants; and, 5) assessments of registrants' residential locations. Despite the plethora of research in these areas, a complete understanding of the residential choices of registered sex offenders remains elusive, particularly as it relates to their neighborhood choices and geographic distribution.

An evaluation of sex offender residences is important for several reasons. First, it will provide descriptive information about the areas in which these individuals are permitted to reside. Second, it will identify communities that have a propensity to house greater numbers of these individuals and may therefore be at greater risk for crimes perpetrated by these individuals. Third, additional information about the potential challenges presented by the communities where these individuals reside may prove helpful to improving rehabilitation and reintegration efforts for sex offenders. Finally, it may assist law enforcement agencies with scarce human and budget resources to locate non-compliant offenders.

The purpose of this paper is to fill a notable gap in the literature regarding the evaluation of sex offenders' residential locations. A predictive-analytical framework based on social disorganization theory (Shaw and McKay, 1942; Sampson, 1985; Sampson and Groves, 1989) will be developed to help determine which demographic and socioeconomic factors best predict the residential locations of registered sex offenders. This paper will also develop an index of social disorganization (ISDOR) that describes disorder along a continuum for each unit of analysis (e.g. block groups) rather than in a binary (i.e. yes or no) context - the latter of which is insufficient for exploring the impacts of offender residence restrictions. Moreover, the ability to quantify varying levels of social disorganization at a fine geographic scale is important given the concerns associated with residence restrictions 'forcing' convicted offenders into socially disorganized areas and the negative consequences that this type of relocation may have on rehabilitation and reassimilation efforts (Levenson and Hern, 2007; Tewksbury, 2005; Levenson, 2008; Levenson et al., 2007).

The results of this analysis are expected to broadly contribute to the public debates regarding the efficacy of sex offender policies in the United States and abroad, by providing a robust and repeatable methodological framework for generating important descriptive information about the areas in which these individuals are both permitted and/or choose to reside. In addition, by identifying communities and neighborhoods having a propensity to house greater numbers of these individuals, issues of equity and risk may also be objectively explored.

The remainder of this paper is organized as follows. In the next section, we review the hypothesized collateral consequences of sex offender laws in the United States, focusing on the impacts of residence restrictions and the resulting spatial distributions of sex offenders. Next, a brief review of social disorganization theory and its practical linkages to sex offender policies is presented. This is followed by the introduction of a multivariate index, rooted in social disorganization theory, for capturing the pertinent demographic and socio-economic determinants of sex offender residential choice. Components of the developed index are then utilized in both linear statistical models and nonlinear data mining approaches for predicting sex offender locations in Hamilton County, Ohio and Jefferson County, Kentucky. Empirical results are presented and we conclude the paper with a brief discussion of the results and their implications for public policy.

Sex offenders, communities and collateral consequences

Registration and Community Notification Laws

Sex offender registration and community notification laws are designed to increase community awareness so that the public may take proactive measures to protect their children from these individuals and to reduce the likelihood of recidivism (Edwards and Hensley, 2001). Where social impacts are concerned, there is a large body of literature regarding the collateral social consequences of community notification. While evidence concerning the impacts of offender registration and notification laws on offender recidivism is nearly non-existent, many studies suggest that the consequences associated with these laws may potentially encourage recidivism. For example, Levenson and Cotter (2005) argue that notification can exacerbate the stressors which trigger acts of sexual abuse, such as isolation, disempowerment, shame, depression, anxiety and the lack of social support. Qualitative studies also indicate that community notification can spur vigilantism, harassment, loss of employment and threats upon offenders (Zevitz et al., 2000; Levenson and Cotter 2005; Tewksbury, 2005). It is also suggested that the unstable housing and employment situations of offenders, a direct consequence of their known sex offender status, "can have a critical impact on the minimum essentials needed for reintegration of offenders within the community" (Zevitz et al., 2000, 375).

Residence restrictions

In addition to the issues associated with offender registration and community notification, a number of collateral consequences associated with residence restrictions have also been found. For example, Levenson and Cotter (2005, 169) suggest that the spatial arrangement of schools and parks *can* create a massive, overlapping restriction zone, "making it essentially impossible for sex offenders in some cities to find suitable housing". This is supported by recent empirical work which suggests that housing availability, not unexpectedly, is diminished when restriction zones are in place (Zandbergen and Hart, 2006; Chajewski and Calkins-Mercado, 2008; Barnes et al., 2008; Zgoba et al., 2009). Much of the empirical work on this topic however focuses on the availability of housing for offenders to the exclusion of other important research questions including the propensity for offenders to live in restricted areas (Grubesic et al., 2007; Barnes et al., 2008), the affordability of available housing (Grubesic et al., 2007), and the characteristics of areas where offenders are permitted to live.

Qualitative evaluations of neighborhood characteristics and sex offenders suggest that spatial restriction zones force offenders into socially disorganized areas (Levenson and Cotter, 2005) and that these environments may contribute to the perpetration of additional sexual offenses (Mustaine et al., 2006). Given the potential link between environmental characteristics and increased recidivism, a more rigorous *quantitative* evaluation of the restricted vs. non restricted areas is needed for understanding the wide-reaching impacts of these ordinances. Interestingly, the sole non-survey based quantitative study examining the characteristics of offender residential locations (Grubesic et al., 2008) found unrestricted block groups in Hamilton County, Ohio actually contained more favorable demographic and socio-economic characteristics than restricted block groups. This study also found that despite this dichotomy, offenders appeared to reside in restricted areas in greater numbers than unrestricted areas. However, before any firm conclusions can be drawn, it is clear that further inquiry into the residential environments of registered sex offenders is required.

Residence restrictions and social disorganization theory

Social disorganization theory (Shaw and McKay, 1942) provides an excellent framework for evaluating how geography, socio-economic status and demographic composition can play a role in criminogenic activity at the neighborhood level (Bursik, 1988; Sampson and Groves, 1989; Krivo and Peterson, 1996). In the context of sex offenders, there are concerns that a lack of social cohesion in a neighborhood effectively limits the informal social controls that keep them from committing sexual offenses (Burchfield and Mingus, 2008). In turn, this lack of social controls within a neighborhood environment is potentially conducive to recidivism (ibid). Furthermore, there are concerns that some offenders actually seek to establish a residence in socially disorganized neighborhoods for two purposes. First, there are general fears that convicted offenders target these areas because of the potentially larger availability of unsupervised victims (i.e. children) (Tewksbury and Mustaine, 2006).[1] Second, Burchfield and Mingus (2008: 359) note that many offenders are attempting to "disappear" in an effort to "avoid the shame and humiliation of being a registered sex offender." Thus, socially disorganized neighborhoods are certainly appealing in this context because they represent locations where the chances of being recognized are greatly reduced. Where corrections and law enforcement agencies are concerned, if offenders do gravitate to socially disorganized neighborhoods, the ability to differentiate varying levels of social disorganization at a relatively fine geographic scale can assist agencies in better targeting local intervention efforts within a community - especially for more vulnerable or exposed neighborhoods (e.g. increased patrols, address verification, etc.). It is also possible that this added level of neighborhood differentiation can enlighten law enforcement efforts for tracking down non-compliant offenders (i.e. offenders that fail to register), although this is often a time-consuming and difficult process.[2] From a broader perspective, more accurate descriptions of neighborhood conditions and their potential linkages to sex offender residency can better inform analysts, planners and law enforcement agencies on issues of community equity when attempting to manage offender populations (Grubesic and Murray, 2008).

[1] It is important to note that empirical evidence from the Tewksbury and Mustaine (2006: 71) study suggests that there *is* a notable minority of convicted offenders that locate their residences in places where they will have "quick, easy and efficient access to a pool of potential victims."

[2] Recent data suggests that over 100,000 offenders are considered non-compliant or "missing" from federal, state and local registries.

The next section outlines a quantitative approach for evaluating the varying levels of social disorganization within a community. Leveraging a combination of socioeconomic and demographic variables for defining socially disorganized areas, a multivariate index is developed and then combined with both linear statistical and non-linear data mining approaches for developing an analytical framework that predicts sex offenders' residential locations.

Methodology

Index of social disorganization (ISDOR)

In an effort to empirically capture varying levels of social disorganization at the neighborhood level, we utilize the variables highlighted in Table 1 to construct a basic, multivariate index of social disorganization (ISDOR). By design, index construction was kept relatively simple to ensure that the methodology is easily repeatable for law enforcement agencies and associated practitioners. It is also important to note that our interpretation of social disorganization, particularly where variable selection is concerned, is somewhat flexible. While the literature typically cites poverty, residential mobility, ethnic heterogeneity, and family disruption as the core components of the social disorganization (Osgood and Chambers, 2000), we have no intention of strictly limiting ISDOR to these measures. As will be illustrated below, there are benefits in maintaining some flexibility in the composition of ISDOR for describing neighborhoods where sex offenders reside.

The first step in the construction of ISDOR is to determine how the values of each variable should be interpreted. This step is necessary because of a potential mismatch between metric *value* and metric *interpretation:* higher variable values do not necessarily correspond to higher levels of social disorganization. For example, *higher* values of percent rental vacancy, suggest higher levels of social disorganization. Conversely, *lower* values for median age suggests higher levels of social disorganization. This mismatch problem is important to resolve, particularly for constructing an additive index such as ISDOR, where high values will correspond to high levels of social disorganization. Table 2 illustrates the interpretive framework used for each metric. Next, a natural breaks method was utilized to create standardized, component variable values. This was necessary because values of median age do not neatly match values for median income, population density, etc. More importantly, this step is necessary to maintain internal consistency for each variable, ensuring that their quantitative interpretation is not diluted in the composite index. The resulting range of values for each metric is 1-10, with 10 indicating the highest level of social disorganization. This was operationalized with the Jenks natural breaks method which is specified as follows:

$$\sum_{n=i}^{j} A[n]^2 - \frac{\left(\sum_{n=1}^{j}[n]\right)^2}{j-i+1} \tag{1}$$

The Jenks natural breaks method minimizes within-class sum of squared differences, where A is the set of variable values that have been ordered from 1 to n. Specifically, $1 \leq i \leq j \leq n$. Most desktop geographic information systems (GIS), including ArcGIS, have the ability to perform this technique in the "classification" menu for creating thematic maps making it easily accessible to practitioners and law enforcement agencies.

After this process is completed, it is possible to construct the ISDOR index, which is specified as follows:

$$ISDOR_i = \sum_{j=1}^{n} x_{ij} \tag{2}$$

Where:
i = the number of spatial units (1,…, n)
j = the number of variables (1,…, n)

x_{ij} = the value of variable j in spatial unit i

$x_j \in [1,10]$

The values range between 1 and 10 for each j and represent varying levels of income, education, population density, etc. When summed, the index represents the derived level of social disorganization for each spatial unit, i.[3] The overall interpretation of ISDOR is relatively simple. Higher values indicate elevated levels of social disorganization for a spatial unit, while lower values indicate the opposite. As noted previously, there are no practical limits to the number of variables used in this index; therefore, the numeric interpretation is governed by n.[4] Thus, if ten variables are used, the unstandardized ISDOR ranges from 10 to 100. If twelve variables are used, it would range from 10 to 120. The unbounded version of ISDOR can be modified to force the index into a more regularized spectrum of potential values, regardless of the number of variables used. For example, if equation (2) is divided by the maximum value of ISDOR across all spatial units under observation, the index has a lower bound of 0.10 and an upper bound of 1.00. The bounded version of ISDOR is specified as follows:

$$ISDOR_{i*} = \frac{\sum_{j=1}^{n} x_{ij}}{\max_i \left(\sum_{j=1}^{n} x_{ij} \right)} \qquad (3)$$

Customizing ISDOR via Exploratory Data Analysis

While the initial selection of variables for inclusion in ISDOR was theoretically driven, it is also possible to utilize basic exploratory data analysis (EDA) (Tukey, 1977) for determining which combination of variables generates the most effective index for capturing an existing spatial distribution of sex offenders in a community. In essence, while an unmodified ISDOR

---

[3] Obviously, the choice of spatial units is flexible (e.g. block group, tract, etc.)

[4] It is also possible to implement a weighting scheme for this index to emphasize certain variables over others.

does an excellent job in summarizing the varying levels of social disorganization, it can be further refined to capture both social disorganization and maximize the number of offenders accounted for within a community. Specifically, because ISDOR is a composite index, it is likely that certain component variables do a better job of summarizing where offenders live than others. In some cases, this may not include the more traditional metrics associated with social disorganization theory. Further, it is also likely that these key variables vary between communities. Therefore, the following EDA approach provides some additional flexibility for analysts dealing with unique demographic or socioeconomic structures within a region.

Figure 1 shows the cumulative distribution for two variables considered for inclusion in ISDOR, public assistance and median income. Both could be considered surrogates for *poverty* in a social disorganization framework. The x-axis displays the standardized values of public assistance and median income. The y-axis displays the cumulative percentage of registered offenders accounted for as the overall level of social disorganization increases (this is outlined in Table 1 for each variable). In this instance, public assistance did a relatively poor job of describing block groups that contained large numbers of sex offenders while median income did a relatively good job. This performance is reflected in the shape of their respective cumulative distribution curves. For example, the curve for public assistance climbs rather quickly and then levels off at a value of 7. This suggests that more offenders are located in block groups with fewer households receiving public assistance – suggesting that this variable is not a particularly good descriptor of block groups where offenders reside in Hamilton County. Conversely, the curve for median income climbs rather steeply in the section of the graph associated with lower household median incomes (7-10), which means the variable is a good descriptor of the block groups where offenders reside.

This type of insight generated from EDA is an important one for constructing ISDOR. As mentioned previously, while a generalized ISDOR will perform fine for most applications, the ability to customize variable selection via EDA provides additional relevancy for accounting for sex offender populations in regions with unique socioeconomic or demographic compositions. For the purposes of this study, various combinations of the variables outlined in Table 1 were considered in constructing the optimal ISDOR, with the overall goal of using a parsimonious set of variables to account for as many sex offender residences as possible.

Artificial neural networks (ANN)

The development of an index like ISDOR is valuable in two respects. First, it can function as a descriptive tool for identifying areas with higher or lower levels of social disorganization. Second, its descriptive properties can also be used to evaluate the propensity of registered offenders to reside in socially disorganized locales. This is accomplished by a simple overlay procedure in a geographic information system, where the number of offenders in each area classified as socially disorganized area is tabulated. However, ISDOR does not, at least by itself, provide a *predictive* framework for identifying locations where sex offenders may be residing. When ISDOR is combined with tools that excel in learning or recognizing patterns, such as artificial neural networks (ANN), the variables from the derived index can serve as a crucial input for predicting where sex offenders choose to establish a residence.

Artificial neural networks operate on a premise similar to regression models; the goal is to predict an outcome based on selected inputs (Olligschlaeger, 1998). However, ANNs possess an advantage over parametric statistical techniques, like regression, because they do not depend upon distributional requirements, such as the normal or multivariate normal distribution to reliably generate their predictions. Further, the performance of artificial neural networks actually improves as they are trained with more data. Openshaw (1998: 1863) summarizes several of the advantages of using artificial neural networks when compared to more standard statistical approaches:

a. they are universal approximators,
b. they are equation free
c. they are highly nonlinear
d. they are robust and noise resistant

Despite their advantages over other statistical applications, there are some caveats to be considered when utilizing artificial neural networks. For example, some analysts are troubled by the fact that most ANNs appear to be equation free and function as "black-box" models (Olligschlaeger, 1998). Further, because the standard theoretical frameworks for developing predictive models no longer apply, Openshaw (1998) suggests that one needs a certain amount of 'faith' that the study constructs are well grounded. Care must also be taken to assure that the networks are not overfitted; otherwise the ANNs might be predicting noise in the data rather than actual patterns (Fischer and Gopal, 1994; Olligschlaeger, 1998; Corcoran et al., 2003). Perhaps the most significant issue is that ANNs provide limited (or no) information about the underlying processes a model seeks to represent.

Even with these limitations, artificial neural networks still provide a powerful framework for generating predictive models. One of the most popular artificial neural network models is known as *backpropogation* (McClelland and Rumelhart, 1988). The topology of this type of ANN is represented in Figure 2 (Eberhart and Dobbins, 1990). The inputs to this type of model typically represent a set of raw data or parameters that represent a single pattern (Eberhart and Dobbins, 1990), where the selection of n is a function of the type of pattern or problem one is analyzing or the way the actual data are represented. These initial inputs are passed to a layer of processing neurons which are found in the input layer. The next step is to distribute signals along multiple paths to the hidden layer neurons (Figure 2). With each individual distribution, the ANN associates a weight between the hidden neuron and the input layer.[5] Similarly, each neuron of the hidden layer is also associated with a weight and connected to every neuron in the output layer.[6] The overall goal of this training process is to minimize the average sum of squared error so that the outputs from the ANN match the observed data as closely as possible. Computationally, one critical aspect of many ANNs, including the network used in this research, is the use of the gradient descent algorithm (Snyman, 2005) for minimizing error. This type of approach helps the network avoid getting stuck at a suboptimal set of weights because of nonconvexity, flatness or local minima in the sum of squares (Openshaw, 1998). While space limitations prevent us from detailing every aspect of artificial neural networks and their associated equations, these technical details can be found elsewhere (see Eberhart and Dobbins, 1990; Fischer and Gopal, 1994; Hewitson and Crane, 1994; Lawrence, 1993; Masters, 1993; McClelland and Rumelhart, 1988; Olligshlaeger, 1998).

Study area and data

Hamilton County, Ohio and Jefferson County, Kentucky are the study areas utilized in this paper. Both locations serve as the central county for their metropolitan areas, and are home to the cities of Cincinnati (pop. 332,252) and Louisville (pop. 256,231), respectively. The choice of these counties for analysis is three-fold. First, sex offender data for both locations is readily available from local law enforcement agencies. Second, the morphological structure of both counties is relatively diverse, intermixing heavily urbanized cores with more suburban and exurban areas on the fringe. This provides a needed level of geographic diversity for obtaining

[5] Weights are randomly generated.
[6] Most backpropogation models are designed as "feedforward" networks. As a result, there are no feedback loops in the system (Eberhart and Dobbins, 1990).

robust results that may be generalizable to other areas. Finally, the authors are familiar with both locations, having conducted previous field work in these locales.

Sex offender registry data for Hamilton County were acquired from the county Sherriff, Simon L. Leis, in June of 2007. Registry data for Jefferson County were acquired from the Kentucky State Police Sex offender Registry in June of 2008. After some basic preprocessing of the data, including address cleanup and standardization, all offenders were geocoded and assigned latitude and longitude coordinates, based on their published registry address. Only those offender addresses that received a street-level address match were kept for analysis. This included information on 1,302 registered offenders in Hamilton County and 722 offenders for Jefferson County[7]. Census block group demographic and socioeconomic estimates from Caliper Corporation (2007) were utilized to construct ISDOR. All spatial data were processed and managed in TransCad 4.5 (Caliper, 2007) and ArcGIS 9.2 (ESRI, 2008). Tiberius 6.05 (2008) served as the artificial neural network for this study.

Empirical results

ISDOR results

A variety of ISDOR constructs were tested for this paper, with their compositions ranging from five to ten demographic and socioeconomic determinants. Figure 3 displays the results of ISDOR in its simplest and best-performing composition, utilizing education, population density, median age, percent white and median income. Areas shaded in red represent block groups with ISDOR values above the mean, while those shaded in blue reflect ISDOR values below the mean. Thus, red areas correspond to relatively higher levels of social disorganization while blue areas correspond to relatively lower levels of social disorganization. One of the more interesting

[7] This is to ensure that no locational biases or errors propagated through the neural network. For more details on geocoding error and spatial analysis, see Whitsel et al., 2006, Ratcliffe, 2001 and 2004.

outcomes of classifying block groups in this way is the obvious spatial overlap of sex offenders in areas that ISDOR designates as socially disorganized. In fact, the associated metrics pertaining to this overlap are remarkable. ISDOR defined areas of social disorganization contain 77% of all registered sex offenders in Hamilton County and 80% of all registered sexual predators.[8] In Jefferson County, approximately 66% of all registered offenders are currently residing in areas displaying relatively high levels of social disorganization. It is important to note that direct comparisons of unstandardized ISDOR values between areas can be problematic due to variations in the value ranges associated with the classified component variable in each of these areas. However, in the case of Jefferson County and Hamilton County, where the mean value is nearly identical (25 and 24 respectively), comparisons in overall levels of social disorganization between these counties is not completely unwarranted.

The ISDOR results lend credibility to the hypothesis that sex offenders gravitate towards socially disorganized locations (Mustaine et al., 2006; Burchfield and Mingus, 2008). However, the simple categorization of a block group as an area with a high level of social disorganization does not necessarily mean sex offender residences will be located in this area. In this context, residency restrictions or some other combination of factors (e.g. transportation access) may influence offender residence choice. To investigate the coincidence of social disorganization with residency restrictions, a cross tabulation of block groups possessing above average levels of social disorganization with their relative restricted/unrestricted status was computed.[9] Figure 4 illustrates the distribution of restricted (371) and unrestricted (365) block groups in Hamilton

[8] In Ohio, if an offender is convicted of, or pleads guilty to a sexually oriented offense, that person is automatically classified as a sexually oriented offender. If the judge feels that a higher classification (i.e. Sexual Predator, Habitual Sex Offender) may be necessary, a Sexual Predator Hearing is held. Classifications are not mandated to any specific offense (HCSO, 2005).

[9] The designation of block groups as restricted or unrestricted followed the convention established in Grubesic et al., 2008. A block group was designated as restricted if the proportion of restricted parcels was greater than the proportion of unrestricted parcels in the block group. A block group was designated as unrestricted if the proportion of unrestricted parcels was greater than the proportion of restricted parcels in the block group.

County.[10] Interestingly, 40% of unrestricted block groups and 59% of restricted block possess above average levels of social disorganization. Further, 37% of all registered sex offenders live in unrestricted block groups considered socially disorganized, while 40% live in restricted block groups considered socially disorganized. This suggests that regardless of restriction status, offenders appear to gravitate towards more socially disorganized areas within Hamilton County.

Despite the cartographic and numerical overlap of sex offender residential locations with ISDOR defined areas of social disorganization, it is important to note this index is purely *descriptive* and not *predictive* in nature. Therefore, an artificial neural network will be utilized in an effort to provide a more robust predictive-analytical framework. Specifically, the ANN will be used to evaluate whether variables indicative of social disorganization help forecast the residential choices of offenders. The neural network will be trained to predict the presence or absence of registered sex offenders for each block group in Hamilton County. The resulting series of weights identified by the ANN in Hamilton County will then be combined with an identical set of demographic and socio-economic determinants for Jefferson County to predict sex offender presence and to assess the generalization properties of the network.

ANN results

The backpropogation artificial neural network for Hamilton County was trained using the component variables of the ISDOR (n = 5) index (education, population density, median age, percent white and median income) in their classified form (1-10).[11] Initially, 70% of the 736 block groups (n = 515) were randomly selected for training, with the remaining 30% (n = 221) withheld from the training process – subsequently used for validating the results.

[10] There are a total of 736 block groups in Hamilton County.
[11] This represents a total of 22,256,640 potential unique combinations of values to consider in the ANN.

The resulting ANN for Hamilton County correctly predicted 72.5% (n = 534) of the block groups in Hamilton County (Figure 5a). While there is no discernable spatial pattern to the results, perhaps the outcome of these types of discrete classification problems is best summarized by Figure 6, a confusion matrix, and the potential class predictions produced by the model. The objective of this type of classification problem is to maximize the frequencies of true *positives* and *true negatives* in prediction. By classifying the predictions and actual observations in this way, a variety of diagnostic metrics can be used to capture the precision and accuracy of the resulting classification (Fawcett, 2006). One such metric is known as the AUC (area under the receiver operating curve). The AUC is a two-dimensional depiction of classifier performance that compares true positives to false positives for a discrete classification problem. As noted by Fawcett (2006, 868), an important property of the AUC metric is that the "AUC of a classifier is equivalent to the probability that the classifier will rank a randomly chosen positive instance higher than a randomly chosen negative instance." Further, the AUC is easily related to the more commonly used Gini coefficient by a relatively simple formula (Hand and Till, 2001):

$$Gini + 1 = 2 * AUC \qquad (3)$$

This metric and its accompanying curve displayed in Figure 7 highlights the quality of performance of the fitted ANN for Hamilton County. The diagonal line represents a classification strategy of random guesses (AUC = 0.50). When AUC = 1.0, the model correctly classified all observations. For this model, the AUC score of ~0.75 suggests a relatively good performance. The composite Gini curves, one based on the observations in the training set and the other based on the validation set also suggest that the model is able to effectively differentiate between block groups with and without sex offenders.

Given the strong performance of the ANN, model weights for Hamilton County were used to classify block groups in Jefferson County in an effort to gauge the transferability and generalizability of the ANN. Figure 5b highlights the resulting classifications from a geographic perspective. Without additional training, the ANN for Hamilton County produced correct classifications of sex offender presence or non-presence for 67.6% of Jefferson County's block groups. Again, this is a fairly strong performance, particularly considering that the ANN was not trained on any of the Jefferson County ISDOR data. This also suggests the fitted ANN is highly generalizable, particularly when trained with inputs from ISDOR. Further it advocates that variables indicative of social disorganization may prove useful in understanding sex offender residential location decisions in a variety of locations outside of Hamilton County.

Discussion and conclusion

Several aspects of the analysis presented in the previous sections merit additional discussion. First, the results of the artificial neural network (ANN) analysis demonstrate the utility of social disorganization, both as a theoretical framework and a tool (operationalized in ISDOR) to gain insight into the residential locations of registered sex offenders. Interestingly, empirical results suggest that offenders gravitate towards these neighborhoods regardless of their restricted or unrestricted status. In some respects, this debunks the theory that offenders are being forced into socially disorganized areas because of publically mandated residence restrictions – at least in Hamilton and Jefferson counties. Further, it is important to note that there is a big difference between 'forced' residence in socially disorganized areas because of the implementation of public policies and 'forced' residence in such areas because of one's economic circumstances. As noted by Grubesic et al. (2008), affordable housing is found throughout Hamilton County, including those areas which are not restricted to offenders.

Specifically, the empirical results of this study suggested that unrestricted block groups had a more favorable demographic and socioeconomic profile than did restricted block groups. Where the results of this paper are concerned, it is interesting to note that median income levels appear to describe the locations of offenders better than median rent. While both metrics are highly correlated, it does suggest that housing affordability does not appear to be the sole driver of sex offender residential choice. Specifically, these results suggest that a large percentage of offenders are *choosing* to live in socially disorganized areas despite the ability to live elsewhere. As noted previously, offenders may find these areas appealing because neighborhood conditions allow them to escape the stigma associated with registration and community notification. Further, socially disorganized areas can provide enough "cover" for many offenders to simply disappear. Therefore, the ability to efficiently identify these neighborhoods via ISDOR is extremely useful.

A second aspect of the analytical approach presented in this study is the excellent performance of the artificial neural network for predicting the locations of registered sex offenders in Hamilton and Jefferson Counties. This is particularly true when one considers that only five determinants were used to construct the ISDOR index and train the ANN. A comparative analysis of the quality of the ANN performance relative to other predictive measures was performed to demonstrate the quality of the obtained results. The neural network results were compared to a naïve prediction and a logistic regression that utilized the ISDOR component variables as predictors.[12] Table 3 summarizes the results of this comparative analysis. The performance of the ISDOR trained ANN exceeded that of the naïve and logistic forecasts for Hamilton County and was comparable to the logistic forecast for Jefferson County. In practice, ISDOR could be easily extended to accommodate additional variables that would provide supplementary inputs to ANN, which may produce more accurate results in both

[12] The naïve prediction assumed all block groups contained at least one sex offender.

Hamilton and Jefferson Counties. Additional variables might include information on job availability or local employment opportunities and access to transportation networks. Nevertheless, the use of variables indicative of social disorganization certainly provides useful information that illuminates some of the characteristics of locales where offenders reside.

A third aspect of the analytical framework presented above is its prospective use for identifying block groups where unregistered or non-compliant offenders may be residing. One of the most ironic outcomes in the implementation of sex offender residence restrictions is the evidence that suggests this legislation has made the monitoring and tracking of offenders more difficult. For example, Davey (2006) notes that the state of Iowa has seen a *three-fold* increase in sex offender non-compliance with registration mandates since the implementation of residence restrictions. Recent research in the state of California also reveals that authorities lost track of approximately 33,000 sex offenders, representing 44% of the 76,350 offenders who had registered with the state at least once in 2002 (Curtis, 2003). When confronted with this information, the Attorney General of California acknowledged that a lack of human resources and funding for maintaining the state sex offender registry plays a major role in the accumulation of so many "lost" offenders (Curtis, 2003). While there are a number of statewide initiatives to find non-compliant offenders such as the Sex Offender Apprehension Felony Enforcement Initiative (SAFE) (WDOC, 2006), the vast majority of local efforts are less organized and largely understaffed. Therefore, the predictive-analytical framework presented in this paper has the potential to help maximize the allocation of sparse law enforcement resources and optimize efforts to locate neighborhoods that non-compliant offenders may be attracted to.

Finally, the ability to identify communities and neighborhoods that have a propensity to house a greater number of sex offenders is important for public policy development. As noted by Grubesic and Murray (2008), the effective management of sex offenders is contingent upon understanding the potential impacts of restriction zones, dispersion ordinances or other spatial strategies prior to their implementation. Poorly conceived public policies are ones that fail to consider these ramifications and therefore place sensitive populations at a greater risk than intended. For example, a failure to coordinate restriction zones across regions composed of many communities can lead to the unintended formation of sex offender clusters (Grubesic, 2009) on their margins. Given these challenges associated with sex offender management strategies and public policy, the methodological framework presented in this paper can provide significant descriptive detail two respects. One, it may be used to inform spatial modeling efforts designed to evaluate proposed management strategies. Two, it may be used to provide community officials and law enforcement agencies a good first-pass analysis of offender residential choice. Regardless of application, more information, particularly when generated in a local spatial context, is certainly better than incomplete information, hypotheses and guesswork. As noted previously, if offenders are found to be gravitating toward specific locales in a community, these areas represent may be targeted for additional intervention efforts (education, rehabilitation programs) and law enforcement activity (e.g. offender address verification). Further, the predictive aspects of the methodology outlined in this paper can help communities with growing populations of offenders determine which areas may require attention in the future.

In conclusion, sex offender management in the United States and abroad will continue to challenge law enforcement agencies, elected public officials and members of the community. Considering the growing concerns regarding offender reintegration and the potential threats of recidivism, a more complete understanding of the spatial dynamics of sex offenders, residential choice and the impacts of public policies is needed. The methodological framework utilized in this paper not only provides more descriptive information about the residential environments of sex offenders, but holds significant promise for predicting sex offender residential locations. Further, the inherent flexibility of ISDOR, its ease of use and the ability to seamlessly integrate the index into a GIS framework, holds great promise for local agencies concerned with sex offender management, particularly considering the time and budget constraints most of these agencies face.

References

Adam Walsh Child Protection Act, 2006, Public Law 109-248. 42 U.S.C § 16901-16962.

Barnes J C, Dukes T, Tewksbury R, De Troye T M, 2009, "Analyzing the Impact of a Statewide Residence Restriction Law on South Carolina Sex Offenders" *Criminal Justice Policy Review* **20** 21-42

Burchfield K B, Mingus W, 2008, "Not in my neighborhood: Assessing registered sex offenders' experiences with local social capital and social control" *Criminal Justice and Behavior* **35** 356-374

Bursik R J Jr, 1988, "Social Disorganization and Theories of Crime and Delinquency: Problems and Prospects" *Criminology* **26** 519-551

Chajewski M, Calkins-Mercado C, 2009, "An Evaluation of Sex Offender Residency Restriction Functioning in Town, County, and City-Wide Jurisdictions" *Criminal Justice Policy Review* **20** 44-61

Corcoran J J, Wilson I D, Ware A, 2003, "Predicting the geo-temporal variations of crime and disorder" *International Journal of Forecasting* **19** 623-634

Curtis K, 2003, "California 'Loses' 33,000 Sex Offenders" Associated Press
URL: http://www.cbsnews.com/stories/2003/01/08/national/main535654.shtml

Davey M, 2006, "Iowa's Residency Rules Drive Sex Offenders Underground" *New York Times*, **March 15**

Eberhart R, Dobbins R, Eds., 1990, *Neural Network PC Tools, A Practical Guide*. (Academic Press, San Diego)

Edwards W, Hensley C, 2001, “Contextualizing sex offender management legislation and policy: Evaluating the problem of latent consequences in community notification laws” *International Journal of Offender Therapy and Comparative Criminology* **45** 83-101

Fawcett T, 2006, “An introduction to ROC analysis” *Pattern Recognition Letters* **27** 861-874

Fischer M M, Gopal S, 1994, “Artificial Neural Networks: A New Approach To Modeling Interregional Telecommunications Flows” *Journal of Regional Science* **34** 503-527

Grubesic T H, 2009, “Sex offender clusters” *Applied Geography* **in press**

Grubesic T H, Murray A T, 2008, “Sex Offender Residency and Spatial Equity” *Applied Spatial Analysis and Policy* **1** 175-192

Grubesic T H, Murray A T, Mack E A, 2008, “Sex offenders, housing and spatial restriction zones” *GeoJournal* **73** 255-269

Grubesic T H, Mack E, Murray A T, 2007, “Geographic exclusion: Spatial analysis for evaluating the implications of Megan’s law” *Social Science Computer Review* **25** 143-162

HCSO, 2005, “Hamilton County Sherriff’s Office” *Sex Offender FAQ*
URL: http://www.hcso.org/PublicServices/SexOffenders/FAQ.htm

Hand D J, Till R J, 2001, “A Simple Generalisation of the Area Under the ROC Curve for Multiple Class Classification Problems” *Machine Learning* **45** 171-186

Hewitson R, Crane R G, Eds., 1994, *Neural Nets: Applications in Geography* (Kluwer Academic Publishers, Boston)

Krivo L J, Peterson R D, 1996, “Extremely Disadvantaged Neighborhoods and Urban Crime” *Social Forces* **75** 619-650

Lawrence J, 1993, *Introduction to Neural Networks: Design, Theory, and Applications* (California Scientific Software, Nevada City)

Levenson J S, 2008, “Collateral consequences of sex offender residence restrictions” *Criminal Justice Studies* **21** 153-166

Levenson J S, Hern A L, 2007, “Sex Offender Residence Restrictions: Unintended Consequences and Community Reentry” *Justice Research and Policy* **9** 59-74

Levenson J S, Zgoba K, Tewksbury R, 2007, “Sex Offender Residence Restrictions: Sensible Crime Policy or Flawed Logic? *Federal Probation* **71 2-8**

Levenson J S, Cotter L P, 2005, “The Impact of Sex Offender Residence Restrictions: 1,000 Feet From Danger or One Step From Absurd” *International Journal of Offender Therapy and Comparative Criminology* **49** 168-178

Masters T, 1993, *Practical Neural Network Recipes in C+ +* (Academic Press, San Diego)

McClelland J, Rumelhart D, 1988, *Explorations in Parallel Distributed Processing: A Handbook of Models, Programs, and Exercises* (MIT Press, Cambridge)

Mulford C F, Wilson R E, Parmley A M, 2009, “Geographic Aspects of Sex Offender Residency Restrictions: Policy and Research” *Criminal Justice Policy Review* **20** 3-12

Mustaine E E, Tewksbury R, Stengel K M, 2006, “Social Disorganization and Residential Locations of Registered Sex Offenders: Is This a Collateral Consequence?” *Deviant Behavior* **27** 329-350

Olligschlaeger A M, 1998, “Artificial Neural Networks and Crime Mapping”. in *Crime Mapping and Crime Prevention*, *Crime Prevention Studies*, Vol. 8. Eds D Weisburd, T McEwen (Criminal Justice Press, Monsey)

Openshaw S, 1998, “Neural network, genetic, and fuzzy logic models of spatial interaction” *Environment and Planning A* **30** 1857-1872

Osgood D W, Chambers J M, 2000, “Social Disorganization Outside the Metropolis: An Analysis of Rural Youth Violence” *Criminology* **38** 81-116

Ratcliffe J H, 2004, “Geocoding crime and a first estimate of a minimum acceptable hit rate” *International Journal of Geographical Information Science* **18** 61-72

Ratcliffe J H, 2001, “On the accuracy of TIGER type geocoded address data in relation to cadastral and census areal units” *International Journal of Geographical Information Science* **15** 473-485

Sampson R J, 1985, “Neighborhood and crime: The structural determinants of personal victimization. *Journal of Research in Crime and Delinquency* **22** 7-40

Sampson R, Groves B, 1989, “Community Structure and Crime: Testing Social-Disorganization Theory” *American Journal of Sociology* **94** 774-802

Shaw C, McKay H, 1942, *Juvenile Delinquency and Urban Areas: A Study of Rates of Delinquents in Relation to Differential Characteristics of Local Communities in American Cities* (University of Chicago Press, Chicago)

Snyman J A, 2005, *Practical Mathematical Optimization: An Introduction to Basic Optimization Theory and Classical and New Gradient-based Algorithms* (Springer, New York)

Tewksbury R, Mustaine E E, 2006, "Where to Find Sex Offenders: An Examination of Residential Locations and Neighborhood Conditions" *Criminal Justice Studies* **19** 61-75

Tewksbury R, 2005, "Collateral Consequences of Sex Offender Registration" *Journal of Contemporary Criminal Justice* **21** 67-81

Tiberius, 2008, *Back Propagation of Error Algorithm Proof*. Retrieved August 28, 2008, from http://www.tiberius.biz/bpproof.html

Tukey J W, 1977, *Exploratory Data Analysis* (Addison-Wesley, Reading)

(USDOJ) United States Department of Justice, 2008, *The National Guidelines for Sex Offender Registration and Notification* URL:
http://www.ojp.usdoj.gov/smart/guidelines_final.htm.

Whitsel E A, Quibrera P M, Smith R L, Catellier D J, Liao D, Henley, A C, Heiss G, 2006, "Accuracy of commercial geocoding: assessment and implications" *Epidemiologic Perspectives & Innovations* **3** 8

(WDOC) Wisconsin Department of Corrections, 2006, *Governor Doyle Announces Over 1,000 Noncompliant Sex Offenders Tracked Down*. Retrieved July 29, 2008, from http://www.wi-doc.com/Non-compliants_trackeddown.htm

Zandbergen P A, Hart T C, 2006, "Reducing Housing Options For Convicted Sex Offenders: Investigating The Impact of Residency Restriction Laws Using GIS" *Justice Research and Policy* **8** 1-24

Zevitz R G, Farkas M A, 2000, "Sex Offender Community Notification: Managing High Risk Criminals or Exacting Further Vengeance" *Behavioral Science and the Law* **18** 375-391

Zgoba K M, Levenson J, McKee T, 2009, "Examining the Impact of Sex Offender Residence Restrictions on Housing Availability" *Criminal Justice Policy Review* **20** 91-110

Table 1: Demographic and Socio-Economic Variables Used for Evaluating Social Disorganization

Variable	Definition	Study	Characteristic Described
Population Density	Population per square mile	Blau and Blau, 1982; Danzinger, 1976; Patterson, 1991	Higher crime rate
Pop <14	Population less than or equal to 14 years of age	Grubin, 1998	Potential Victims*
Median Age	Median Age	Burchfield and Mingus, 2008	Social Control
% White	Percent White	Carroll and Jackson, 1983; Messner, 1982; Patterson, 1991; Roncek, 1981	Racial Composition
Rental vacancy	Rental vacancy rate	Krivo and Peterson, 1996; Roncek and Maier, 1991	Neighborhood stability and/or transience
Education	Number of people with a high school education or less	Lochner and Moretti, 2001; Lochner, 2004	Socio-economic Status
Public Assist.	Number of households with public assistance	Hannon and Defronzo, 1999; Sampson et al., 1997	Socio-economic Status
Plumbing	Occupied households lacking complete plumbing	Shuerman and Kobrin, 1986	Urbanization/Neighborhood Quality
% Unemployed	% people 16+ unemployed in the labor force	Krivo and Peterson, 1996	Disenfranchised populace
Median Income	Median household income	Kawachi et al., 1999; Smith and Jarjoura, 1989	Socio-economic Status
Poverty Status	Number of families with a female householder and w/related children under 18 years for whom the poverty status is determined	Patterson, 1991; Cohen and Felson, 1979; Sampson, 1989	Social Control and Socio-Economic Status
Median Rent	Median rent	Grubesic et al., 2008	Affordable Housing

* Note: In the context of sex offender studies

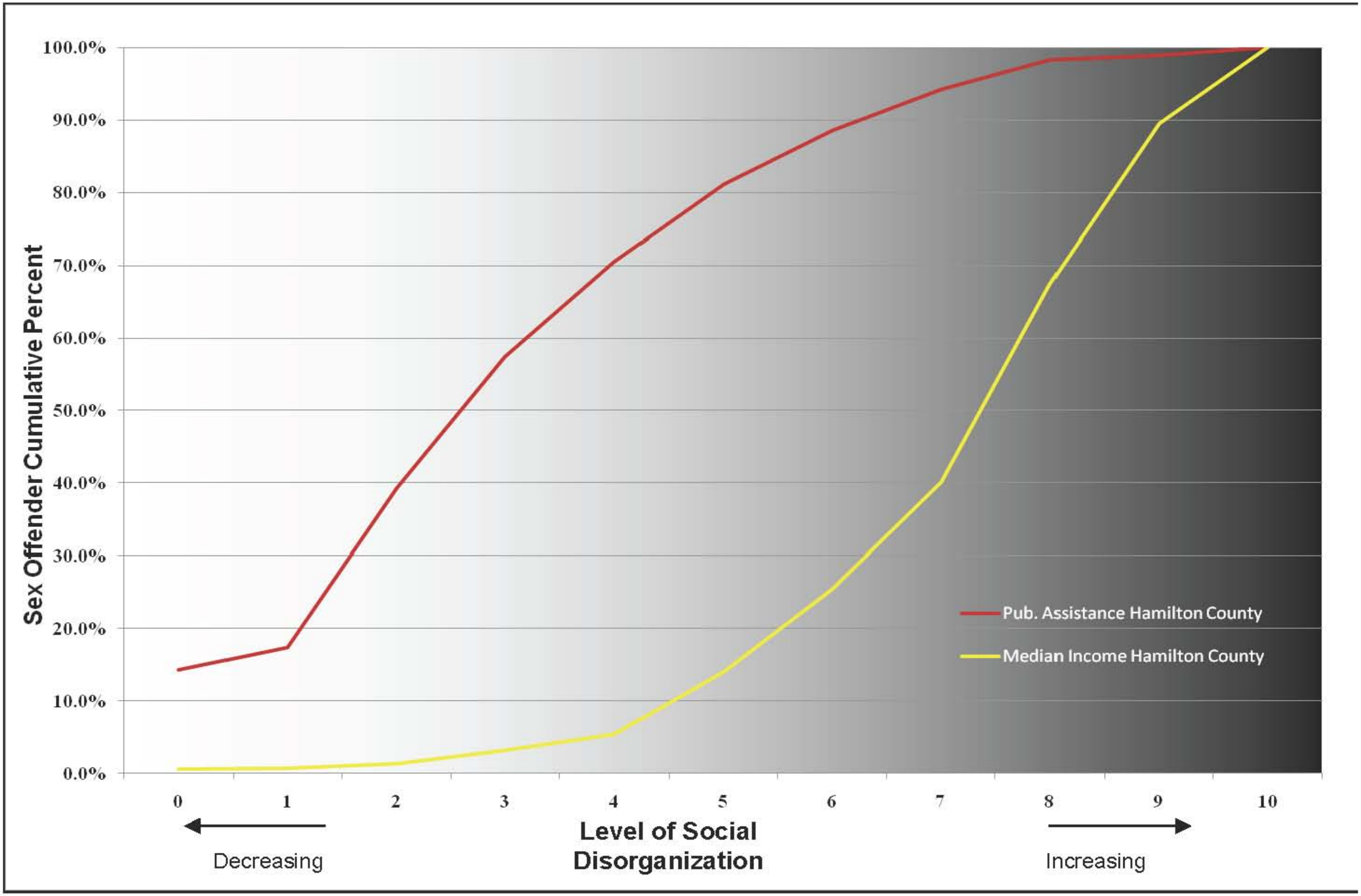
Table 2: ISDOR Variable Contribution and Interpretation

Variable	Social Disorganization	Classification Value
Median Age	lower	lowest values = 10
Percent White	lower	lowest values = 10
Rental Vacancy	higher	highest values=10
Education	higher	highest values=10
Pub. Asst.	higher	highest values=10
No. Plumbing	higher	highest values=10
% Unemployed	higher	highest values=10
Median Income	lower	lowest values = 10
Poverty Status	higher	highest values=10
Median Rent	lower	lowest values = 10
Pop. Density	higher	highest values=10

Table 3: ISDOR Variable Predictive Performance

	% Predicted Correctly (Hamilton County)	**% Predicted Correctly (Jefferson County)**
Naïve prediction	57%	54%
Logistic regression	69%	71.6%
Artificial neural network	72.5%	67.6%

Figure 1: Cumulative Distribution of Sex Offenders Residences by Social Disorganization Indicator Variable



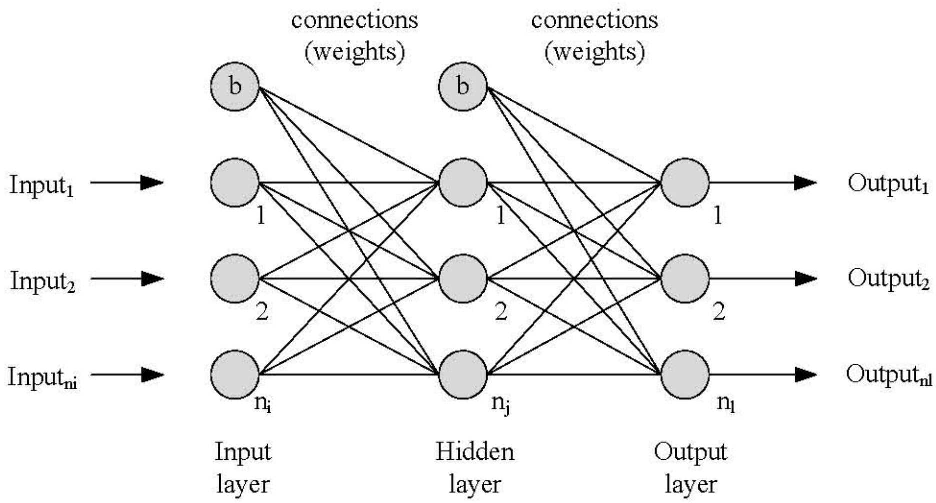


Figure 2: A Feedforward Artificial Neural Network Topology
Source: (Eberhart and Dobbins, 1990)

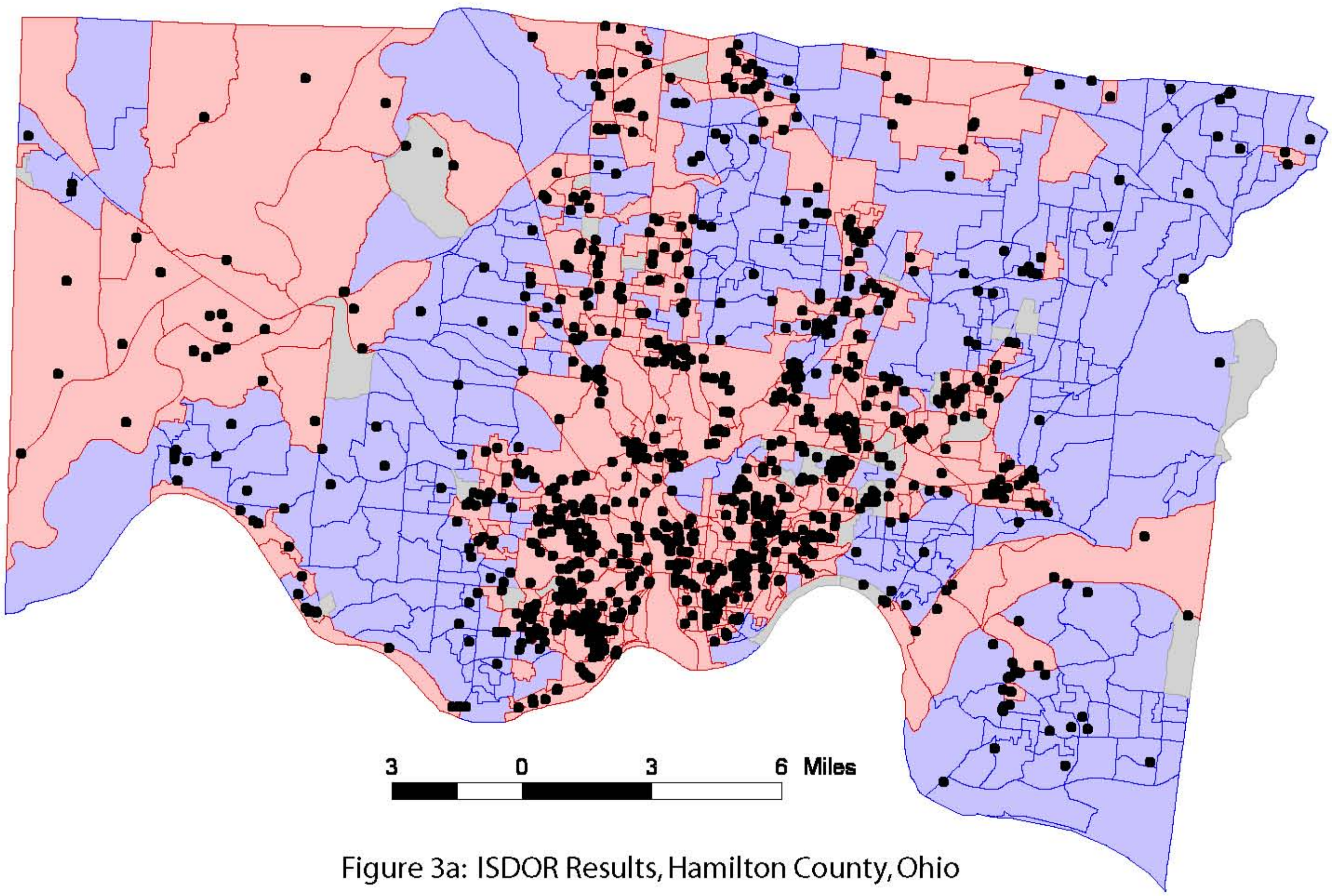


Figure 3a: ISDOR Results, Hamilton County, Ohio

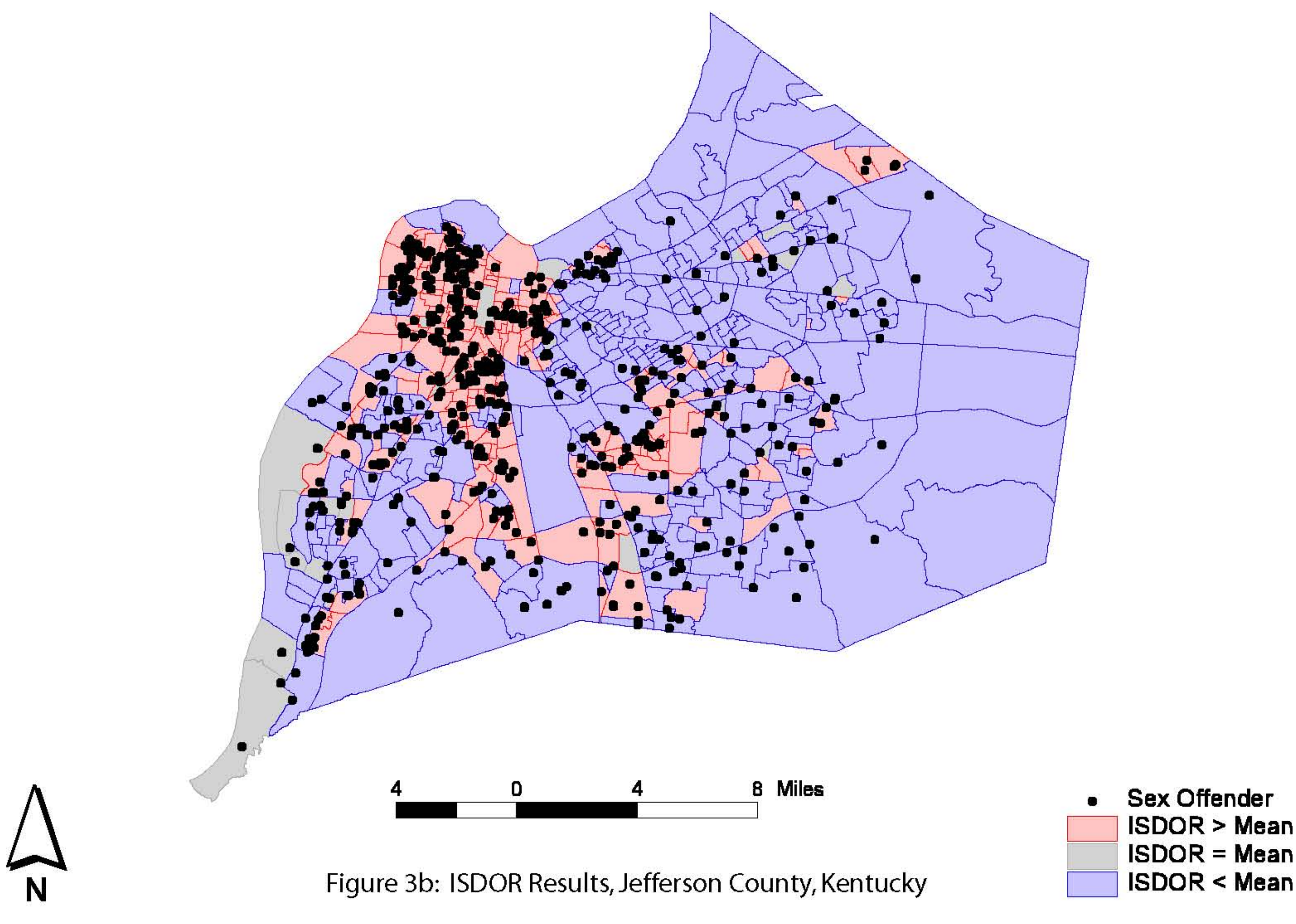
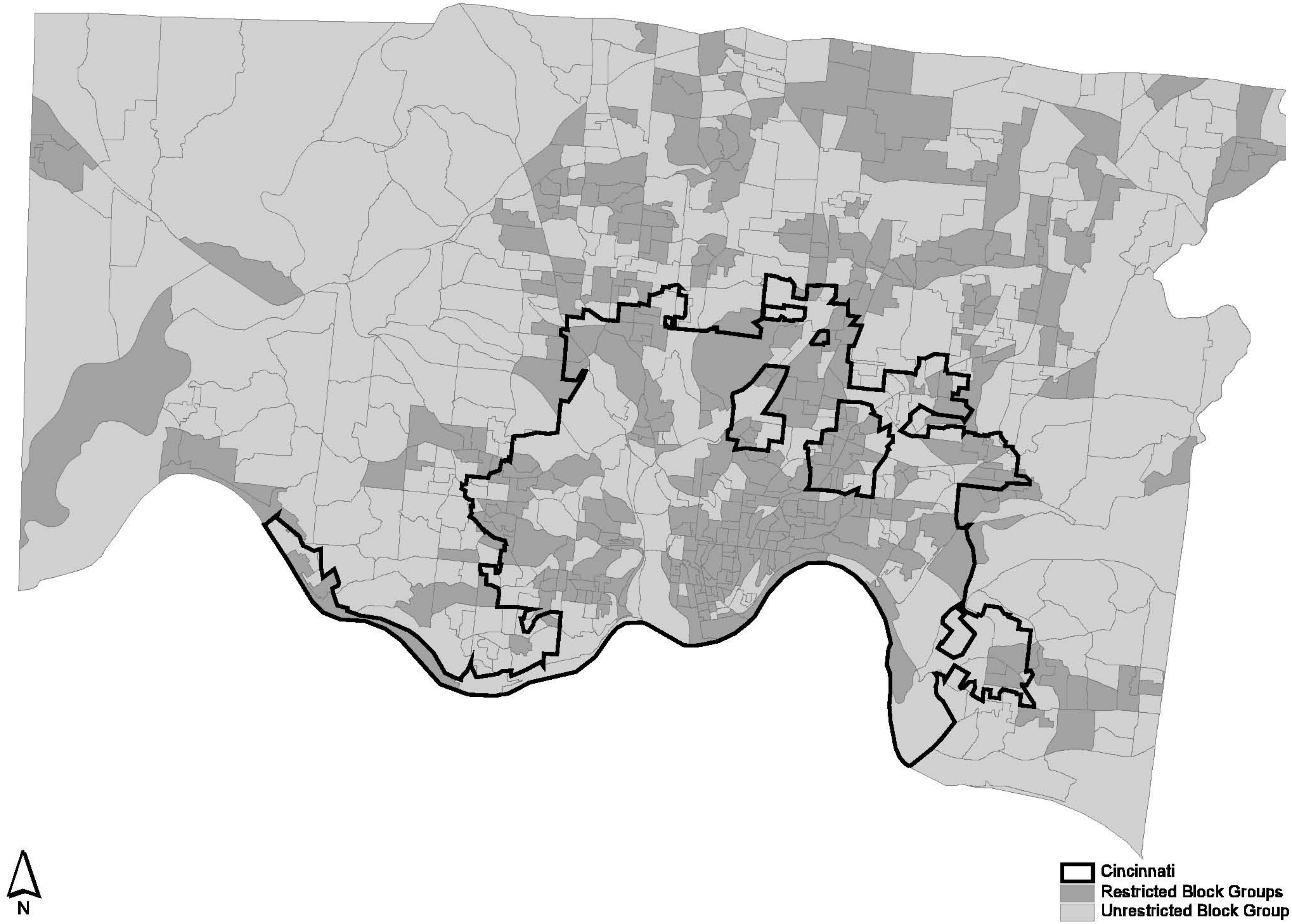


Figure 3b: ISDOR Results, Jefferson County, Kentucky

Figure 4: Spatial Restriction Zones by Block Group, Hamilton County Ohio



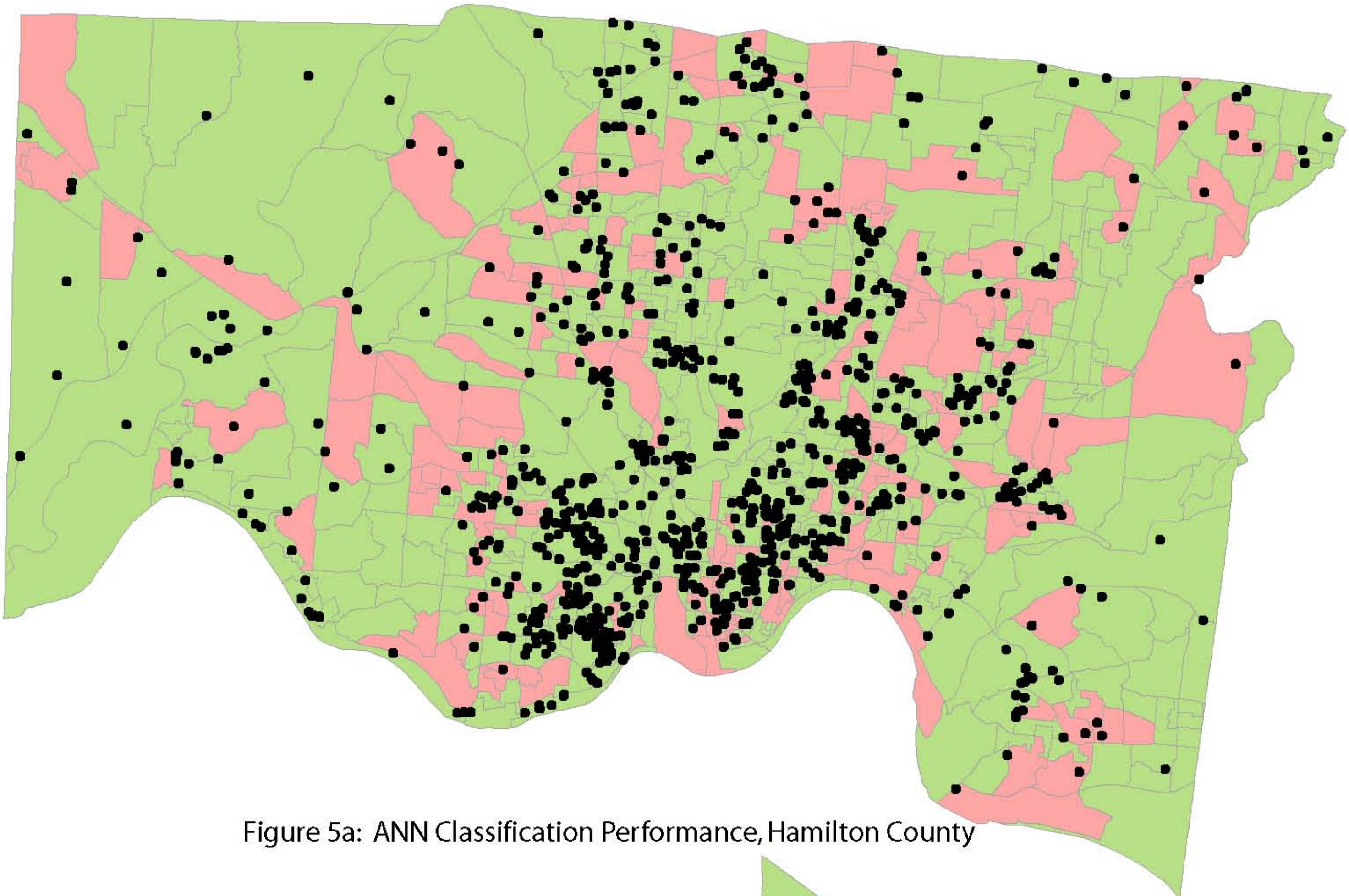


Figure 5a: ANN Classification Performance, Hamilton County

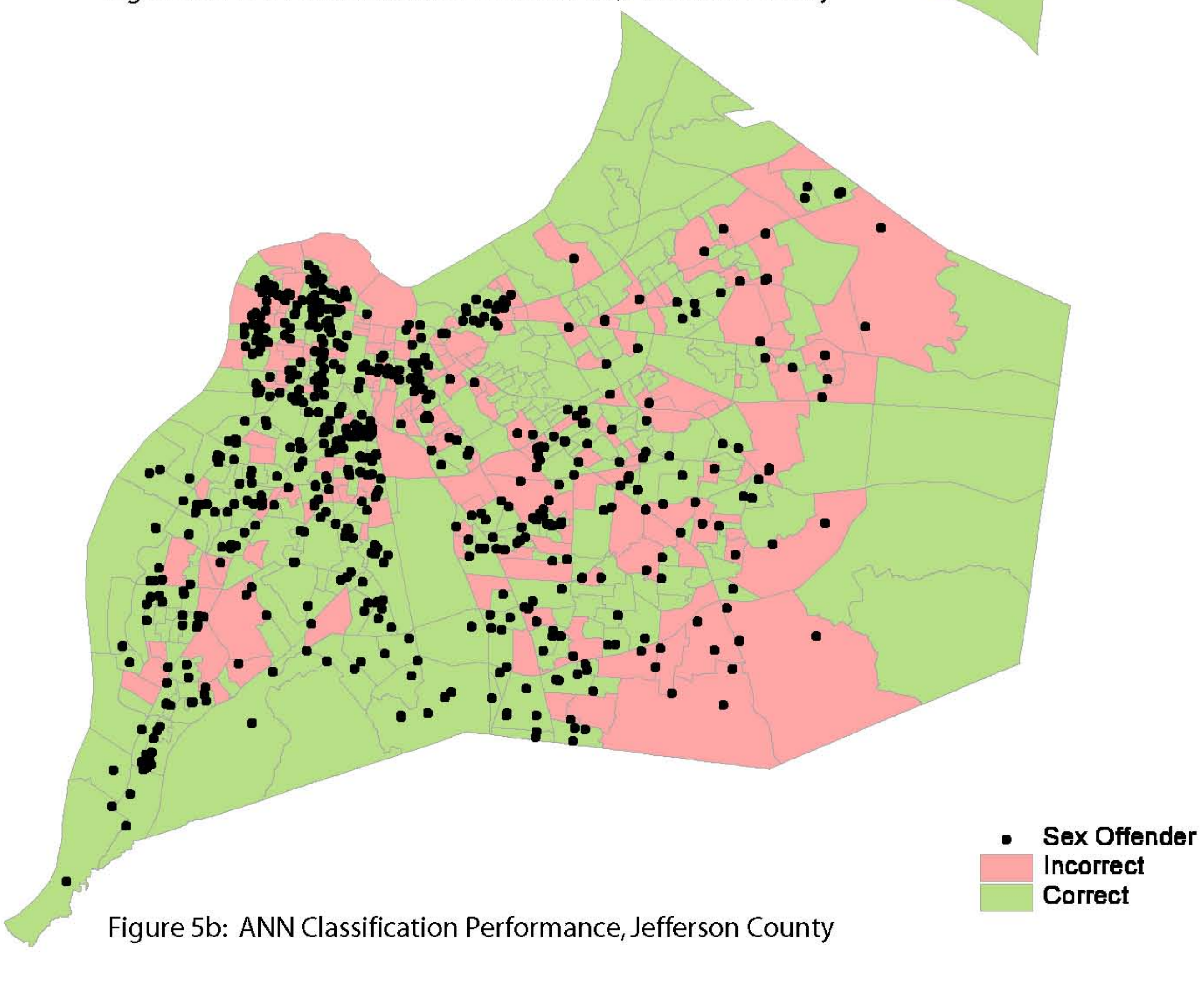


Figure 5b: ANN Classification Performance, Jefferson County

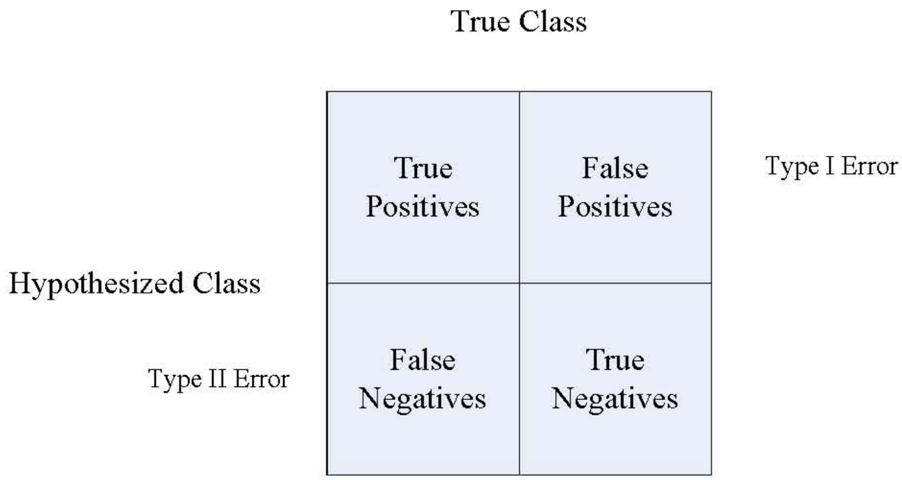


Figure 6: Confusion Matrix for Discrete Classification Problems

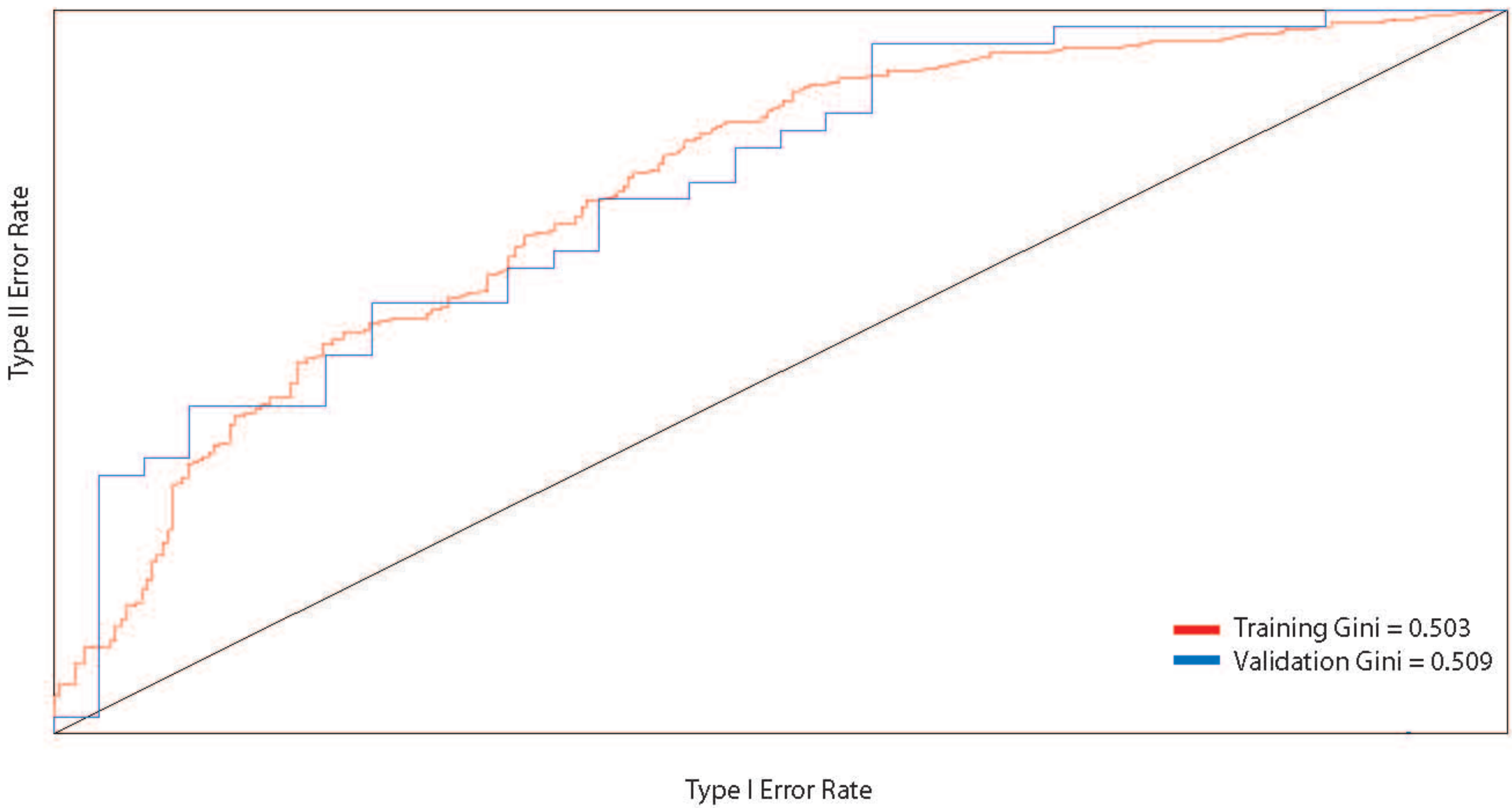


Figure 7: Artificial Neural Network Performance for Hamilton County