Communications for Statistical Applications and Methods
2023, Vol. 30, No. 1, 95–107

https://doi.org/10.29220/CSAM.2023.30.1.095
Print ISSN 2287-7843 / Online ISSN 2383-4757

Intensity estimation with log-linear Poisson model on linear networks

Idris Demirsoy[1,a], Fred W. Huffer[b]

[a]Computer Engineering, Usak University, Turkiye;
[b]Department of Statistics, Florida State University, USA

Abstract

Purpose: The statistical analysis of point processes on linear networks is a recent area of research that studies processes of events happening randomly in space (or space-time) but with locations limited to reside on a linear network. For example, traffic accidents happen at random places that are limited to lying on a network of streets. This paper applies techniques developed for point processes on linear networks and the tools available in the R-package spatstat to estimate the intensity of traffic accidents in Leon County, Florida.

Methods: The intensity of accidents on the linear network of streets is estimated using log-linear Poisson models which incorporate cubic basis spline (*B*-spline) terms which are functions of the x and y coordinates. The splines used equally-spaced knots. Ten different models are fit to the data using a variety of covariates. The models are compared with each other using an analysis of deviance for nested models.

Results: We found all covariates contributed significantly to the model. AIC and BIC were used to select 9 as the number of knots. Additionally, covariates have different effects such as increasing the speed limit would decrease traffic accident intensity by 0.9794 but increasing the number of lanes would result in an increase in the intensity of traffic accidents by 1.086.

Conclusion: Our analysis shows that if other conditions are held fixed, the number of accidents actually decreases on roads with higher speed limits. The software we currently use allows our models to contain only spatial covariates and does not permit the use of temporal or space-time covariates. We would like to extend our models to include such covariates which would allow us to include weather conditions or the presence of special events (football games or concerts) as covariates.

Keywords: spatial statistics, log-linear Poisson model, linear network, model selection

1. Introduction

Not only do accidents cause injuries and fatalities, they can also destroy a city's economy. In 2015, the US spent $871 billion on the economic and social harm caused by vehicular accidents, including one billion dollars spent toward vehicle emergency services (Lowy, 2014). Understanding traffic accidents, specifically where they occur, is an ongoing field of research. Various discussions and papers have offered solutions to this problem.

Traffic accidents occur randomly in space and time. However, the occurrence of traffic crashes on certain roads relate to the road's traffic volume, the road's physical characteristics (e.g., sharp turns and curve depth), environmental effects (rain, fog), and light conditions (daylight and dark). Factors like these are connected to spatial patterns (Xie and Yan, 2008).

[1] Corresponding author: Department of Computer Engineering, Usak University, Ankara Izmir Yolu 8.km Bir Eylul Kampusu, Usak 64000, Turkiye. E-mail: idris.demirsoy@usak.edu.tr

Published 31 January 2023 / journal homepage: http://csam.or.kr

Spatial point pattern analysis is one way to analyze this type of problem and has been widely used in geographical information systems (GIS) (Okabe *et al.*, 2009). Spatial point patterns: Methodology and applications with R by (Baddeley *et al.*, 2015) and spatial analysis along networks by (Okabe and Sugihara, 2012) are two books that describe this type of analysis in detail.

Spatial point pattern analysis can be classified into two broad topics; First-order properties focus on properties of point process events and investigate the variation of the points' density (intensity) across a study window (area). Intensity estimation and testing for complete spatial randomness are concerned with first-order properties. Second-order properties examine spatial dependency and investigate the interaction between two events (points) of the process as a function of the distance between them or, more generally, of their positions. The K and L functions are commonly used displays of second-order properties.

The estimation of (spatial) intensity or population density is very important in applications. Intensity is the expected number of points per unit in a study area. When it comes to estimating populations such as bird populations in ecology, forest fires in agriculture or disease cases in public health, intensity helps in comparing different events or the same event's different levels and investigate how intensities vary over the spatial location. For example, the intensity of subjects with a certain disease can be compared with non-disease subjects to examine how the disease varies over the study area (Zimmerman, 2008). Another example is that the probability of finding a goldfield depends on the existence of faults. When one identifies that there is a spatial modification, then statistical methods to predict this function can be used. One can investigate intensity on a linear network as being either homogeneous or inhomogeneous.

Homogeneous Intensity: Often the first step for analyzing point patterns begins with the assumption of homogeneity. Let $\mathbf{X}$ be a point process on a linear network $\mathbf{L}$. If the intensity is constant over the entire network, then we say the point process is homogeneous. If a point process on $\mathbf{L}$ is homogeneous with constant intensity λ, then the expected number of points of $\mathbf{X}$ falling in any subset of $B \subseteq L$ is equal to the total length of B, $\ell(B)$, multiplied by the intensity (Baddeley *et al.*, 2015),

$$En\left(\mathbf{X} \cap B\right) = \lambda \ell\left(B\right). \tag{1.1}$$

If $\mathbf{X}$ is homogeneous with intensity λ, then an unbiased estimator of λ on a linear network can be calculated as the ratio of the total number of points in the study area over the total length of the network: $\hat{\lambda} = n(\mathbf{x})/\ell(l)$.

Inhomogeneous intensity: If the intensity of a point process is not constant, then we say it is inhomogeneous. Let $\mathbf{X}$ be a point process on a linear network $\mathbf{L}$ with intensity function $\lambda(u)$. Then

$$En\left(\boldsymbol{X} \cap B\right) = \int_B \lambda\left(s\right) d_1 s. \tag{1.2}$$

In Equation (1.2), s denotes a point in $\mathbf{L}$, $\lambda(s)$ is the expected number of points of $\mathbf{X}$ per unit length on the linear network $\mathbf{L}$ at the point s, $n(\boldsymbol{X} \cap B)$ is the number of points of $\boldsymbol{X}$ in the subset B, and $d_1 s$ represents one-dimensional integration with respect to arc length (Baddeley *et al.*, 2015; Ang *et al.*, 2012).

2. Materials and methods

2.1. Spatial point processes on linear networks

Although spatial point processes on the line and plane have been studied since at least the 1950's, work on spatial point processes on linear networks is a recent development. There are important

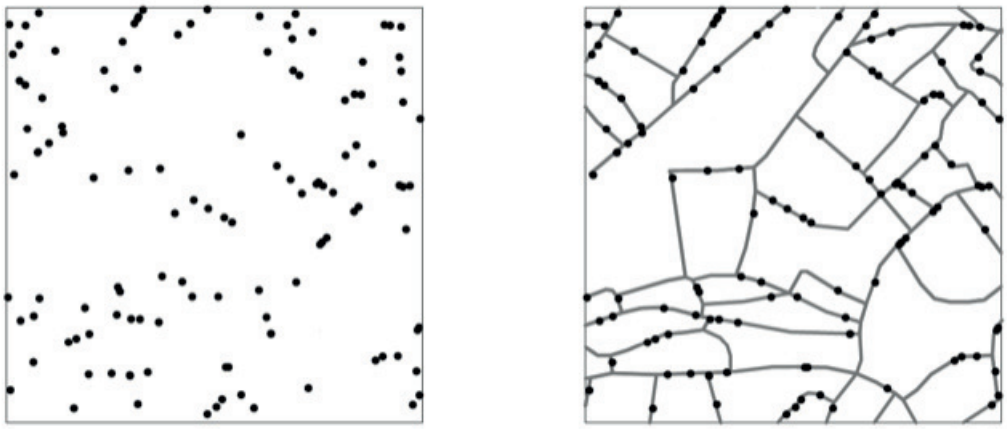

Figure 1: *Points on a planar and linear network (Note: (a) and (b) are the same data points).*

differences between the analysis of point processes on Euclidean spaces (e.g., the line or plane) and point processes on a linear network. For example, the Euclidean distance metric figures prominently in the development of point processes on the plane, but may be an insufficient or misleading distance metric for spatial point patterns on linear networks, such as those arising in the study of the locations of traffic accidents, crime on sidewalks, road-kill, the population distribution of dendrites, or the heterogeneity of tree species along a river. For point processes on a linear network, it is usually more appropriate to define the distance between two points on the linear network to be the length of the shortest path between these two points traveling along the linear network.

Another important difference between point processes on the plane and on a linear network is illustrated in Figure 1. When one looks at Figure 1(a), one may not think the points are randomly distributed, but Figure 1(b) clearly shows that the points are randomly distributed (Okabe and Sugihara, 2012) on a linear network. The notions of "randomness" or "uniformity" are very different on the plane and on a linear network.

2.2. Definitions relating to linear networks

The line segment l on the plane with endpoints $u, v \in \mathbb{R}^2$, $u \neq v$ can be written in any of the following ways:

$$l = l_{u,v} = [u, v] = \{tu + (1-t)\, v : 0 \leq t \leq 1\}\,.$$

The length of this segment can be written as

$$|\, l \,| = |\, l_{u,v} \,| = \|u - v\|,$$

where $\|.\|$ is the usual Euclidean norm in $\mathbb{R}^2$ which for $z = (z_1, z_2)$ is defined by $\|z\| = \sqrt{z_1^2 + z_2^2}$ (Ang *et al.*, 2012; Moradi *et al.*, 2019).

A linear network $\boldsymbol{L}$ is a combination of line segments (edges) l_i:

$$\boldsymbol{L} = \bigcup_{i=1}^{n} l_i\,.$$

The total length of the linear network $\boldsymbol{L}$ is defined by

$$| \boldsymbol{L} | = \sum_{i=1}^{n} \|l_i\| .$$

According to Okabe and Sugihara (2012), the endpoints of segments are called nodes or vertices, and the degree of a node u, written as $d(u)$, is the number of segments that are connected to the node. When $d(u) = 1$, then u is called a *terminal* node (Mc Swiggan *et al.*, 2017).

A path between u and v in a linear network $\boldsymbol{L}$ is a sequence $x_0, x_1, \ldots, x_m$ of points in $\boldsymbol{L}$ such that $x_0 = u$, $x_m = v$, and $[x_i, x_{i+1}] \subset \boldsymbol{L}$ for each $i = 0, \ldots, m-1$. This path is denoted by $P(u, x_1, \ldots, x_{m-1}, v)$. The *length of a path* $P(u, x_1, \ldots, x_{m-1}, v)$ on $\boldsymbol{L}$ is defined to be

$$\|u - x_1\| + \|x_1 - x_2\| + \cdots + \|x_{m-1} - v\| .$$

The *shortest-path distance* between two points u and v in a linear network $\boldsymbol{L}$ is the length of the shortest-path in $\boldsymbol{L}$ between u and v; this distance is denoted by $d_L(u, v)$.

Ang *et al.* (2012) notes that a point process $\boldsymbol{X}$ on a linear network $\boldsymbol{L}$ is a special case of a point process on a planar space. We assume that $\boldsymbol{X}$ is simple, meaning that it does not have any coincident points. Each realization of $\boldsymbol{X}$ is a finite set $x = \{x_1, \ldots, x_n\}$ of distinct points $x_i \in \boldsymbol{L}$, where $n \geq 0$ is (typically) random and not fixed in advance.

2.3. Log-linear Poisson models on linear networks

Poisson processes require independence and randomness of events, but in spatial data settings, this is not always the case. Therefore, Poisson processes are mainly used as a benchmark model (Last and Penrose, 2017; Baddeley *et al.*, 2015; Bivand *et al.*, 2008; Illian *et al.*, 2008; Johnson, 2010).

Definitions relating to the Poisson process come from Last and Penrose (2017), Johnson (2010) and Baddeley *et al.* (2015).

Our point processes will be defined on a set $S \subseteq \mathbb{R}^d$. Let $\lambda : S \to [0, \infty)$ be locally integrable, that is, $\int_B \lambda(u)\, du < \infty$ for all bounded sets $B \subseteq S$, and define the measure $\mu(B) = \int_B \lambda(u)\, du$. For convenience, for any sets X and B define $X_B = X \cap B$.

Definition 1. *Let f be a density function on a set $B \subseteq S$ and $n \in \mathbb{N}$. A point process X consisting of n independent and identically distributed points with common density f is called a binomial point process of n points in B with density f:*

$$X \sim binomial\,(B, n, f) = Bin\,(B, n, f) .$$

Definition 2. *Suppose X is a point process on S. If X satisfies the following properties it will be called a Poisson process with intensity λ and intensity measure μ which we denote by*

$$X \sim Poisson\,(S, \lambda) .$$

- *For any bounded set $B \subseteq S$ such that $\mu(B) < \infty$, we have $n(X_B) \sim Poisson(\mu(B))$.*
- *For any $n \in \mathbb{N}$ and $B \subseteq S$ such that $0 < \mu(B) < \infty$ we have*

$$[\, X_B \mid N(B) = n \,] \sim Bin\left(B, n, \frac{\lambda(u)}{\mu(B)}\right) .$$

Algorithm 1 Simulation of Poisson process

1-Simulate $N(S) \sim \text{Poisson}(\mu(S))$,
2-Generate $N(S)$ i.i.d. points on S from the density $\lambda(\cdot)/\mu(S)$.

The density function for the binomial point process above is $f(\cdot) = \lambda(\cdot)/\mu(B)$

Define $N(B) = n(X_B)$ *to be the number of points of* X *in* B. *If* $X \sim Poisson(S, \lambda)$ *and* $B \subseteq S$ *is bounded, then*

$$\mathbb{E}\left(N\left(B\right)\right) = \mu\left(B\right).$$

When S *is a bounded space, one can simulate* $X \sim Poisson(S, \lambda)$ *as follows.*

Definition 3. *If* $X \sim Poisson(S, \lambda)$, *then* X *is called a homogeneous Poisson process with intensity* λ_0 *if* $\mu(B) = \lambda_0 \mathsf{B}$ *for all* $B \subseteq S$ *where* B *denotes the Lebesgue measure (area) of* B. *If* X *is not homogeneous for any value of* λ_0, *then* X *is said to be an inhomogeneous Poisson process with intensity* λ. *For this process*

$$EN\left(B\right) = \mu\left(B\right) = \int_B \lambda\left(s\right)\, ds\,.$$

Property: If $X \sim \text{Poisson}(S, \lambda)$ and $B_1, B_2, \ldots, B_m$ are disjoint bounded subsets of S, then $N(B_1)$, $N(B_2), \ldots, N(B_m)$ are independent random variables.
Property: Suppose X is a homogeneous Poisson process and $B \subseteq S$ is bounded. Conditional on $N(X_B) = n$, these n points are independent and uniformly distributed in B.

Definition 4. *A point process* X *on* $\mathbb{R}^d$ *is stationary if its distribution is invariant under spatial shifts.*

Definition 5. *A point process* X *on* $\mathbb{R}^d$ *is isotropic if its distribution is invariant under rotations about the origin.*

In applications it is often necessary to simulate inhomogeneous Poisson processes. Baddeley *et al.* (2015) offered a computationally fast strategy for inhomogeneous Poisson process on planar spaces which is to split space into pixels, compute the probability that each pixel contains an event, and select pixels at random using these probabilities. Afterward, they pointed out that to create an inhomogeneous Poisson process with intensity $\lambda(u)$ on a linear network $\boldsymbol{L}$, the linear network can be split into line segments and a similar approach as in planar spaces then carried out.

Fitting a Poisson model: When applying Poisson process models on the plane or on a linear network, it is common to assume the intensity has a log-linear relationship to the covariates in the model. A log-linear model has the form:

$$\lambda_\theta\left(s\right) = \exp\left(B\left(s\right) + \boldsymbol{\theta}^T \boldsymbol{Z}\left(s\right)\right) = \exp\left(B\left(s\right) + \theta_1 Z_1\left(s\right) + \cdots + \theta_m Z_m\left(s\right)\right), \tag{2.1}$$

where

- $B(s)$ is a scalar baseline function.
- $\boldsymbol{Z}(s)$ is an m-dimensional spatial covariate vector.
- $\boldsymbol{\theta}$ is a parameter vector.

Taking the natural logarithm of both sides, the intensity becomes a linear function of the parameters (Baddeley *et al.*, 2015).

$$\log\left(\lambda_\theta\left(s\right)\right) = B\left(s\right) + \boldsymbol{\theta}^T \boldsymbol{Z}\left(s\right). \tag{2.2}$$

Working with a log-linear model has an important advantage in that the intensity is always positive regardless of the values of $B(s)$ and the vectors $\boldsymbol{\theta}$ and $\boldsymbol{Z}(s)$.

The log-likelihood function for a realization $X = \{x_1, x_2, \ldots, x_n\}$ of a Poisson process with intensity $\lambda_\theta(s)$ observed on a region $W \subseteq \boldsymbol{L}$ is given by

$$\log L\left(\theta\right) = \sum_i \log \lambda_\theta\left(x_i\right) - \int_W \lambda_\theta\left(s\right)\, d_1 s. \tag{2.3}$$

For the log-linear model in (2.1), this assumes a standard exponential family form which satisfies all the regularity conditions required for the consistency and asymptotic normality of the maximum likelihood estimates and the asymptotic distribution of likelihood ratio test statistics. (See, Baddeley *et al.* (2015) for a discussion of the likelihood theory for Poisson processes and further references.) As a result of this, when fitting our log-linear Poisson models using the `lppm` function in the `spatstat` *version 2.3.4* package, we have all the standard machinery of maximum likelihood estimation available to us: Parameter estimates with accompanying standard errors and confidence intervals, likelihood ratio tests for comparing nested models, and AIC values for use in model selection.

The integral in (2.3) has no analytic form and must be approximated by numerical methods, but other than this the maximization of the log-likelihood can be accomplished by quite standard methods. The `spatstat` package employs an approach known as the Berman-Turner device (Berman and Turner, 1992) to convert the maximization problem into a form that can be handled by the `glm` function in R.

The log-linear Poisson process models described above are closely related to the Poisson regression models commonly employed to model count data. Poisson regression is frequently applied to data consisting of regional counts of some event, for example, the number of murders in various towns, the number of homes sold in various neighborhoods, the number of cases of pancreatic cancer in various counties, etc. One can in fact approximate the log-linear model for a Poisson process X given in (2.1) by a Poisson regression model constructed by dividing the observation window W into disjoint regions $C_1, C_2, \ldots, C_m$ which are small enough so that the covariate vector $\boldsymbol{Z}(s)$ is essentially constant within each region, and then fitting a Poisson regression model to the regional counts $y_i = n(X \cap C_i)$.

3. Data analysis

Leon County, Florida traffic accident data is used for real data analysis purpose. Prior to the data analysis, two steps were done. The first is using shapefiles to create a linear network representing the roads in Leon County. The second is gathering the data set. The data set, which totals 59,773 accident records, consists of accidents occurring from 2013 through 2019, and was provided by the GeoPlan Center affiliated with the Department of Urban & Regional Planning at the University of Florida. Creating the linear network of roads in Leon County required shapefiles which were available as online resource (Florida Department of Transportation, 1996; Administration, 2009).

A shapefile is a spatial data storage unit which has information on events' geographical location and related attributes. The shapefile can contain different elements such as points, lines, and polygons (Engel, 2017). Various R packages (`sp`, `sf`, `rgdal`, `spatstat`, etc.) supply object classes and software for the storage and manipulation of spatial data.

(a) *The traffic signals* (b) *The intersections*

Figure 2: *Traffic signal and intersection locations.*

- **Points** (event locations) may be stored in `SpatialPointsDataFrame` with or without attributes. For example: The location of trees and their heights, or the location of a hospital with breast cancer patients and patients' ages.
- **Lines** are connections of vertices. Some examples of lines include rivers, roads, and neurons. They can be stored in a `SpatialLinesDataFrame`
- **Polygons** are three or more vertices that are connected and closed. Examples are lakes, islands, counties, oceans, countries and continents. They can be stored in a `SpatialPolygonsDataFrame`.

Locations of bike lanes, speed limits, annual average daily traffic (AADT), annual average daily truck traffic (TruckAADT), road width, and road direction information are downloaded as individual `SpatialLinesDataFrame` objects from the Florida department of transportation (FDOT)'s website. This network information is combined in a shapefile (shown in Figure 3) which created the Leon County linear network. The locations of the traffic signals and the locations of intersections are downloaded from the FDOT website as a `SpatialPointsDataFrame` which is shown in Figure 2.

The linear network of roads in Leon County is created from 8,417 vertices and 8,563 lines. There are 346 different segments which are individual streets or roads, each of which may contain many lines. The same covariates are defined on each segment. These covariates are constant on each segment, but may assume different values on different segments. The values of the covariates for six different segments (roads) are displayed in Table 1. The covariates are described below.

- "Swid" represents the ID number of the segment.
- "AADT" represents the annual average daily traffic volume in the segment.
- "TruckAADT" represents the annual average daily truck traffic in the segment.
- "Bikelane" represents whether or not that particular road has a bike lane.
- "Width" represents the width of the road.
- "Speed" represents the speed limit in the segment.
- "Road_Direct" represents the direction (numbered 1 to 8) of a particular road.

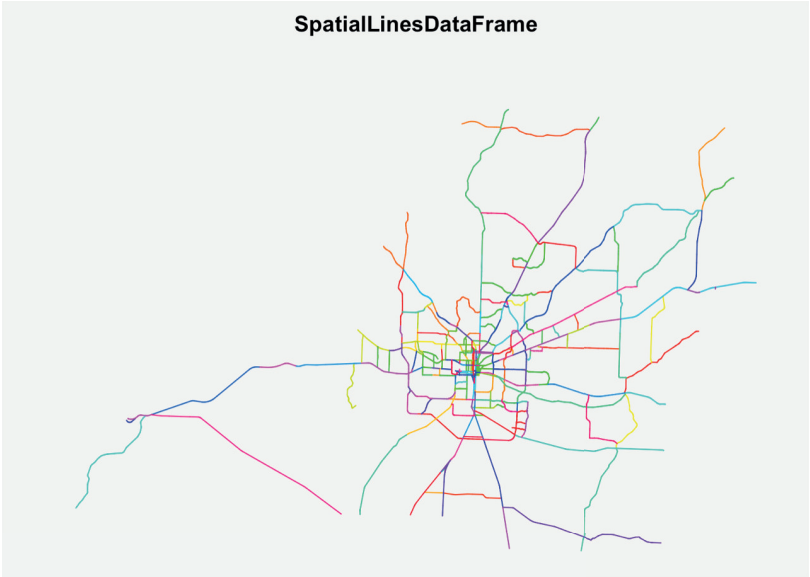

Figure 3: *Spatial lines data frame.*

- "Lane_Cnt" represents the number of lanes in the segment.
- "Length" represents the length of the segment.

The shapefile also contains the spatial locations x and y of points on the network, but these are not shown in Table 1.

There are two more covariates additional to the ones in Table 1: The distance to closest intersection and the distance to closest traffic signal. These covariates are computed using the locations of the intersections and traffic signals shown in Figure 2.

Table 2 shows a few fitted models (of increasing complexity) and their computed AIC values. The best model in Table 2 is Model 10. The covariates used in the best model are speed limit, annual average daily traffic (AADT), TruckAADT, the number of lanes, existence of bike lane (1-yes, 0-no), road direction, width of the road, distance to the nearest intersection, and distance to the nearest traffic signal. Additionally, we have used cubic basis spline (B-spline) terms which are functions of the x and y coordinates. The B-spline terms incorporate a spatial trend of the form $f(x) + g(y)$ into our model, where $f(x)$ and $g(y)$ represent smoothly varying functions of the x and y coordinates, respectively. The function $f(x)$ is a piecewise polynomial of degree 3. The points at which the pieces join are called 'knots'. In our work we choose the knots to be equally spaced between the minimum and maximum values of x in the study region, that is, when the number of knots is k, we take equally spaced points $\min(x) = x_0 < x_1 < x_2 < \cdots < x_k < x_{k+1} = \max(x)$ and use the interior points $x_1, x_2, \ldots, x_k$ as the k knot values. The degrees of freedom (number of parameters) in the polynomial $f(x)$ is $k + 3$. The discussion for the function $g(y)$ is similar. In Model 10 of Table 2 we use $k = 9$ knots for both $f(x)$ and $g(y)$. The parameter estimates, standard errors, and 95% confidence intervals for Model 10 are given in Table 3. In this table we see that all the covariates are significant. There are 12 B-spline parameters reported for both the x and y coordinates, consistent with there being 9 knots for both x and y.

The number of knots used when fitting a model containing B splines affects the accuracy and smoothness of the model fit (Atilgan and Bozdogan, 1990). Using too many knots can result in over-

Table 1: Some of the attributes carried by the spatial lines data frame for Leon County, shown for six segments

	Swid	AADT	TruckAADT	Bikelane	Width	Speed	Road_Direc	Lane_Cnt	Length
1	1	16400	1099	0.00	12.00	55	8	2	2.87
2	2	6000	150	0.00	11.00	35	2	2	0.64
3	3	35500	2769	1.00	36.00	45	8	3	0.43
4	4	5100	128	0.00	12.00	35	3	1	0.63
5	5	19100	1051	1.00	24.00	45	8	4	1.33
6	6	15300	1469	1.00	24.00	45	8	1	2.20

Table 2: Log-linear models and their AIC

	Model	AIC	BIC
1	Lnet ~ 1	427831.0	427840.0
2	Lnet ~ speed	409494.0	409511.9
3	Lnet ~ bs(x)+bs(y)	355307.5	355532.5
4	Lnet ~ bs(x)+bs(y)+DistT+DistI	335137.6	335380.5
5	Lnet ~ speed+Truck+Lane+AADT+bike	353599.0	353653.0
6	Lnet ~ speed+Truck+Lane+AADT+road+bike+Width	349978.8	350104.8
7	Lnet ~ speed+Truck+Lane+AADT+road+bike+Width+DistT	339781.7	339916.7
8	Lnet ~ speed+Truck+Lane+AADT+road+bike+Width+DistT+DistI	325754.5	325898.5
9	Lnet ~ speed+Truck+Lane+AADT+road+bike+Width+DistT+DistI+bs(x)	322253.8	322505.8
10	Lnet ~ speed+Truck+Lane+AADT+road+bike+Width+DistT+DistI+bs(x)+bs(y)	320132.4	320492.4

fitting (Likhachev, 2017). One way to find the optimal number of knots is using the Akaike information criterion (AIC), which can be computed as AIC $= -2\log L_{\max} + 2p$, where $L_{\max}$ is the maximum likelihood for the model and p is the number of parameters in the model. When the number of knots is large, the parameter estimates may fail to converge or the Fisher information matrix may be singular so that standard errors cannot be produced. Difficulties can also arise when selecting the locations of the knots.

One solution to these problems is to use a limited number of knots. For the covariates in Model 10, we identified 10 as the largest number of knots for which the estimates converge. Restricting the number of knots to be at most 10 for both x and y, the minimum AIC is achieved when the number of knots is 9 for both the x and y. Table A1 shows the computed AIC numbers for various number of knots in the model, with some omitted for brevity. The estimates in Table 3 is based on using 9 knots for both x and y.

We have also tried the Bayesian information criterion (BIC) which is another method for selecting the best model, the formula for which is BIC $= -2\log(L_{max}) + p\log(N)$, where N is the sample size and p is the number of parameters used in the model. Table A2 shows that using BIC leads to the same conclusion as AIC.

Model Comparison: Model comparison is an important part of choosing the best fitting model. Agresti (2003) noted that only nested models can be compared by formal statistical tests. Nested models are two models in which one model contains all the terms of the other, and at least one additional term. Not all models are nested, and comparing non-nested models can be done using the AIC, BIC, and other model selection criteria. According to Agresti (2003) the model with the smallest AIC is considered the best. As shown in Table 2, Model 10 has the smallest AIC among those listed.

Two nested log-linear Poisson models can be compared using a likelihood ratio test (LRT). Under the null hypothesis that the additional parameters in the larger model are all zero, the decrease in the deviance has an asymptotic χ^2 distribution with degrees of freedom equal to the number of added parameters in the larger model. The sequence of 10 models in Table 2 is not nested, but the subse-

Table 3: Estimates, standard errors and confidence intervals for Model 10

	Estimate	**S.E.**	**CI95.lo**	**CI95.hi**	***p*-value**	**Zval**
(Intercept)	28.46	2.214	24.12	32.80	<2.2e-16	12.85
speed	−2.07e-02	8.77e-04	−2.24e-02	−1.90e-02	<2.2e-16	−23.67
Truck	−6.56e-05	1.20e-05	−8.93e-05	−4.19e-05	<2.2e-07	−5.429
Lane	8.31e-02	8.21e-03	6.70e-02	9.92e-02	<2.2e-16	10.11
AADT	4.68e-05	6.58e-07	4.55e-05	4.81e-05	<2.2e-16	71.17
road2	−5.31e-02	1.39e-02	−8.04e-02	−2.59e-02	<2.2e-04	−3.825
road3	1.59e-01	1.16e-02	1.36e-01	1.82e-01	<2.2e-16	13.72
road4	2.69e-01	1.70e-02	2.35e-01	3.02e-01	<2.2e-16	15.82
road5	3.51e-01	2.90e-02	2.94e-01	4.08e-01	<2.2e-16	12.09
road6	8.90e-01	7.89e-02	7.35e-01	1.04e+00	<2.2e-16	11.27
road7	−1.16e-01	4.01e-02	−1.95e-01	−3.80e-02	<2.2e-03	−2.908
road8	8.78e-02	1.97e-02	4.90e-02	1.26e-01	<2.2e-06	4.440
bike	−2.70e-02	1.10e-02	−4.86e-02	−5.39e-03	<2.2e-02	−2.449
Width	1.35e-03	6.74e-04	3.06e-05	2.67e-03	<4.5e-02	2.005
DistI	−7.55e-03	8.89e-05	−7.73e-03	−7.38e-03	<2.2e-16	−84.95
DistT	−2.37e-04	5.67e-06	−2.48e-04	−2.26e-04	<2.2e-16	−41.87
bs(x)1	−32.68	2.481	−37.55	−27.82	<2.2e-16	−13.17
bs(x)2	−26.59	1.857	−30.24	−22.95	<2.2e-16	−14.31
bs(x)3	−30.81	1.963	−34.66	−26.96	<2.2e-16	−15.69
bs(x)4	−30.23	1.901	−33.96	−26.50	<2.2e-16	−15.90
bs(x)5	−35.52	1.915	−39.27	−31.76	<2.2e-16	−18.54
bs(x)6	−32.91	1.912	−36.65	−29.16	<2.2e-16	−17.21
bs(x)7	−33.59	1.913	−37.34	−29.84	<2.2e-16	−17.55
bs(x)8	−34.42	1.911	−38.17	−30.68	<2.2e-16	−18.01
bs(x)9	−33.74	1.914	−37.49	−29.99	<2.2e-16	−17.62
bs(x)10	−31.79	1.911	−35.53	−28.04	<2.2e-16	−16.63
bs(x)11	−32.43	1.944	−36.25	−28.62	<2.2e-16	−16.67
bs(x)12	−27.19	2.080	−31.26	−23.11	<2.2e-16	−13.07
bs(y)1	4.118	1.279	1.610	6.626	<2.2e-03	3.218
bs(y)2	2.273	1.090	0.136	4.410	<3.8e-02	2.085
bs(y)3	2.780	1.154	0.517	5.043	<2.2e-02	2.408
bs(y)4	2.511	1.122	0.311	4.711	<2.6e-02	2.237
bs(y)5	3.782	1.127	1.572	5.992	<2.2e-04	3.354
bs(y)6	3.235	1.125	1.030	5.441	<2.2e-03	2.875
bs(y)7	2.930	1.127	0.719	5.140	<2.2e-03	2.598
bs(y)8	3.961	1.127	1.752	6.170	<2.2e-04	3.514
bs(y)9	3.691	1.151	1.434	5.949	<2.2e-03	3.204
bs(y)10	6.538	1.170	4.245	8.832	<2.2e-07	5.587
bs(y)11	6.539	1.189	4.208	8.870	<2.2e-07	5.498
bs(y)12	21.86	1.386	19.14	24.58	<2.2e-16	15.77

quence of models 1, 2, 5, 6, 8, 10 is nested, and so is the subsequence 1, 3, 4, 10. Table A4 gives a sequence of LRT's (an analysis of deviance) for the first subsequence, comparing each model with its immediate predecessor, and Table A3 does the same for the second subsequence. The columns in these tables are: The number of parameters in each model, the change in the number of parameters (Df), the change in the deviance (Deviance), and the p-value of the χ^2 test. From these tables one can conclude that Model 10 is the best of the models in Table 2.

4. Conclusion

The model 10 (in Table 3) states that the intensity of traffic accidents declines by 0.9733 with existence of a bike lane. Here is another interesting output from the model is that when speed limit

increases by 1 miles on a road, intensity of the accidents decrease by 0.9794. We believe this might be because when people start getting faster, they become more careful. One can possible claim that number of lanes increase, intensity of accident also increase. Table 3 supports that claim, when number of lane increase by 1, accident intensity increases by 1.086.

Log-linear Poisson model is useful and easy to fit. Since it uses a generalized linear model (GLM) idea, it is straightforward to interpret. Same time, the model is vulnerable to over-fitting. However, in our models, the danger of over-fitting is minimal because we are using only a few covariates and not trying many transformations or including high-order interactions. So long as we use a strict criterion like BIC (in Table A2), we are not worried about over-fitting. We can consider adding more complicated terms like interactions to our models so long as they decrease the BIC. AIC is a less strict criterion than BIC. The scope of this paper is limited to the main effects in order to compare intensity estimation on a linear network with machine learning which we have been working on and going to submit our initial result soon. Therefore, two-way or more way interactions are not considered. It can be also a future work as well. Another development might be to use a thin plate spline or any other spline instead of B-spline.

We more focus on spatial covariates to model our analysis but demographic and weather information could be used in the model as well. Additionally, one can use other R packages or software to accommodate temporal settings in the model since we couldn't make it work with temporal covariates with the current R-package we used to fit the model.

Availability of data and material

The data would be provided when connected by the correspondence author.

Conflict of interest

The authors declare no competing interests.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Appendix

Table A1: Number of equally spaced knots used in the model and the corresponding AIC values

	1	2	3	4	5	6	7	8	9	10
1	323061.1	323006.5	322905.3	322579.5	322715.5	322528.4	322531.5	322564.3	322433.1	322451.5
2	323014.9	322960.3	322855.9	322523.1	322660.6	322471.1	322475.7	322508.7	322376.9	322395.4
3	322126.8	322072.3	321901.4	321365.7	321571.8	321275.3	321294.0	321307.2	321192.4	321229.9
4	322385.0	322330.6	322139.3	321691.0	321868.5	321608.9	321634.4	321647.8	321520.9	321557.9
5	321514.7	321460.5	321269.9	320696.1	320913.3	320583.6	320629.1	320614.9	320498.3	320556.5
6	321593.4	321539.4	321339.0	320782.8	321000.7	320670.7	320711.2	320691.6	320576.1	320633.4
7	321267.6	321213.7	321004.4	320412.1	320626.0	320310.8	320352.7	320336.0	320239.1	320291.5
8	321367.4	321313.7	321116.5	320519.6	320744.1	320407.5	320447.8	320428.8	320325.9	320381.1
9	321158.6	321104.7	320888.4	320292.7	320503.4	320201.3	320236.8	320226.2	**320132.5**	320178.2
10	321215.31	321161.5	320954.9	320355.5	320572.6	320261.8	320291.4	320284.1	320188.7	320234.2

The row and column labels are the number of knots for x and y, respectively.

Table A2: Number of equally spaced knots used in the model and the corresponding BIC values

	1	2	3	4	5	6	7	8	9	10
1	323277.1	323231.5	323139.2	322822.5	322967.4	322789.3	322801.4	322843.2	322721.0	322748.5
2	323239.9	323194.3	323098.9	322775.1	322921.6	322741.0	322754.6	322796.7	322673.9	322701.4
3	322360.8	322315.3	322153.4	321626.6	321841.8	321554.2	321581.9	321604.1	321498.3	321544.9
4	322627.9	322582.6	322400.3	321960.9	322147.4	321896.8	321931.3	321953.8	321835.8	321881.9
5	321766.6	321721.5	321539.8	320975.0	321201.2	320880.6	320934.9	320929.9	320822.3	320889.4
6	321854.4	321809.4	321617.9	321070.7	321297.7	320976.6	321026.1	321015.5	320909.1	320975.3
7	321537.5	321492.6	321292.4	320709.1	320931.9	320625.8	320676.7	320668.9	320581.0	320642.5
8	321646.4	321601.6	321413.4	320825.6	321059.1	320731.5	320780.7	320770.8	320676.9	320741.0
9	321446.5	321401.6	321194.4	320607.6	320827.3	320534.2	320578.7	320577.1	**320492.4**	320547.1
10	321512.3	321467.4	321269.8	320679.5	320905.6	320603.8	320642.3	320644.0	320557.7	320612.1

The row and column labels are the number of knots for x and y, respectively.

Table A3: Analysis of deviance for a subsequence of nested models from Table 2

Model	Number of Param	Df	Deviance	p-value
1	1			
2	2	1	18339	< 2.2e-16
5	6	4	55903	< 2.2e-16
6	14	8	3636	< 2.2e-16
8	16	2	24228	< 2.2e-16
10	40	24	5670	< 2.2e-16

Table A4: Analysis of deviance for another nested subsequence of models from Table 2

Model	Number of Param	Df	Deviance	p-value
1	1			
3	25	24	72572	< 2.2e-16
4	27	2	20174	< 2.2e-16
10	40	13	15031	< 2.2e-16

References

Administration U (2009). TIGER/Line Shapefile, 2018, county, Leon County, FL, All Roads County-based Shapefile, *Data.Gov*, **1**, Available from: https://catalog.data.gov/dataset/tiger-line-%20shapefile-2018-county-leon-county-%20fl-all-roads-county-based-shapefile

Agresti A (2003). *Categorical Data Analysis*, John Wiley & Sons, New Jersey.

Ang Q, Baddeley A, and Nair G (2012). Geometrically corrected second order analysis of events on a linear network, with applications to ecology and criminology, *Scandinavian Journal of Statistics*, **39**, 591–617.

Atilgan T and Bozdogan H (1990). Selecting the number of knots in fitting cardinal B splines for density estimation using AIC, *Journal of the Japan Statistical Society, Japanese Issue*, **20**, 179–

190.

Baddeley A, Rubak E, and Turner R (2015). *Spatial Point Patterns: Methodology and Applications with R*, Chapman, New York.

Berman M and Turner T (1992). Approximating point process likelihoods with GLIM, *Applied Statistics*, **41**, 31–38.

Bivand R, Pebesma E, Gomez-Rubio V, and Pebesma E (2008). *Applied Spatial Data Analysis with R*, Springer, New York.

Engel C (2017). Introduction to spatial data types in R, *Introduction to Spatial Data Types in R*, **5**, Available from: https://cengel.github.io%20/rspatial/2%5CspDataTypes.nb.html

Transportation F (1996). undefined. *FDOT Open Data Hub*, **1**, Available from: https://gis-fdot.opendata.arcgis.com/

Illian J, Penttinen A, Stoyan H, and Stoyan D (2008). *Statistical Analysis and Modelling of Spatial Point Patterns*, John Wiley & Sons, Chichester.

Johnson T (2010). *Introduction to Spatial Point Processes*, University of Warwick, Coventry.

Last G and Penrose M (2017). *Lectures on the Poisson Process*, Cambridge University Press, Cambridge.

Likhachev D (2017). Selecting the right number of knots for *B*-spline parameterization of the dielectric functions in spectroscopic ellipsometry data analysis, *Thin Solid Films*, **636**, 519–526.

Lowy, J (2014). Traffic accidents in the U.S. cost $871 billion a year, federal study finds, *PBS*, **5**, Available from: https://www.pbs.org/newshour/nation/motor-vehicle-crashes

Mc Swiggan G, Baddeley A, and Nair G (2017). Kernel density estimation on a linear network, *Scandinavian Journal of Statistics*, **44**, 324–345.

Moradi M, Cronie O, Rubak E, Lachieze-Rey R, Mateu J, and Baddeley A (2019). Resample-smoothing of Voronoi intensity estimators, *Statistics and Computing*, **29**, 995–1010.

Okabe A, Satoh T, and Sugihara K (2009). A kernel density estimation method for networks, its computational method and a GIS-based tool, *International Journal of Geographical Information Science*, **23**, 7–32.

Okabe A and Sugihara K (2012). *Spatial Analysis along Networks: Statistical and Computational Methods*, John Wiley & Sons, New York.

Xie Z and Yan J (2008). Kernel density estimation of traffic accidents in a network space, *Computers, Environment and Urban Systems*, **32**, 396–406.

Zimmerman D (2008). Estimating the intensity of a spatial point process from locations coarsened by incomplete geocoding, *Biometrics*, **64**, 262–270.

Received June 02, 2022; Revised November 02, 2022; Accepted November 04, 2022