geographical analysis

Geographical Analysis (2024) **56**, 62–78

Temporal Network Kernel Density Estimation

Jérémy Gelb[1] and Philippe Apparicio[1,2]

[1]Institut National de la Recherche Scientifique centre Urbanisation Culture Société, Montréal, Quebec, Canada , [2]Department of Applied Geomatics, Université de Sherbrooke, Sherbrooke, Quebec, Canada

Kernel density estimation (KDE) is a widely used method in geography to study concentration of point pattern data. Geographical networks are 1.5 dimensional spaces with specific characteristics, analyzing events occurring on networks (accidents on roads, leakages of pipes, species along rivers, etc.). In the last decade, they required the extension of spatial KDE. Several versions of Network KDE (NKDE) have been proposed, each with their particular advantages and disadvantages, and are now used on a regular basis. However, scant attention has been given to the temporal extension of NKDE (TNKDE). In practice, when the studied events happen at specific time points and are constrained on a network, the methodologies used by geographers tend to overlook either the network or the temporal dimension. Here we propose a TNKDE based on the recent development of NKDE and the product of kernels. We also adapt classical methods of KDE (Diggle's correction, Abramson's adaptive bandwidth and bandwidth selection by leave-one-out maximum likelihood). We also illustrate the method with Montreal road crashes involving a pedestrian between 2016 and 2019.

Introduction

Analysis of events in space and KDE

Analysis of events represented with points in GIS is common in numerous research and professional fields. The main goal of such work is to identify hot and cold spots which could be further analyzed. One of the most prevalent methods in Point Pattern Analysis (PPA) for this task is the kernel density estimate (KDE). This nonparametric method allows for the estimation of the intensity of function of a spatial process observed with a finite sample of events and, thus, the analysis of first order properties of the phenomenon (density variation). Spatial Kernel Density Estimate (SKDE) is an extension in the geographical space of the KDE method used to estimate the density function of random variables in a one-dimensional space.

Because of its simplicity in use and interpretation, this method has been widely employed in geography. Recent use cases include vehicular crashes (Abdulhafedh, 2017; Shad and Rahimi, 2017), crimes (Setiawan et al., 2019), bird flu (Shi et al., 2018), cases of dengue (Mala

Correspondence: Jérémy Gelb, Institut National de la Recherche Scientifique centre Urbanisation Culture Société, 385 R. Sherbrooke E, Montréal, QC H2X 1E3 Canada
e-mail: jeremy.gelb@inrs.ca

Submitted: December 12, 2022. Revised version accepted: May 23, 2023.

doi: 10.1111/gean.12368

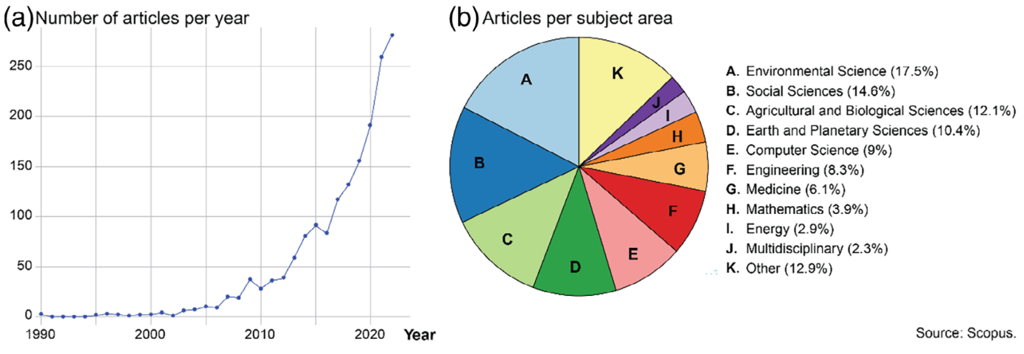

Figure 1. Number of published journal articles related to SKDE, 1990–2022.

and Jat, 2019), wildlife habitats (Zhang et al., 2018), and so on. A quick search on Scopus limited to published papers in peer-reviewed journals between 1990 and 2022 with the following request: TITLE-ABS-KEY ("kernel density" AND estimat* AND (spatial OR geographical OR gis)) AND (LIMIT-TO (DOCTYPE, "ar")) yields more than 1,682 results and reveals the increasing use of the method in a great diversity of study fields (Fig. 1).

Recently, the SKDE method has been adapted for use on a special type of space often encountered in geography: spatial networks. A spatial network (or graph) is a collection of nodes with coordinates in a space using a metric (like the Euclidean distance) and connected by directed or undirected edges (Barthélemy, 2011). Nodes and edges in a spatial network often have associated features which can be used to calculate distances or represent relationship strength between nodes. Graphs are mathematical objects defined in graph-theory and are very useful to model a great number of physical objects such as road networks, bicycle networks, public transit networks, river systems, gas and water networks, electrical circuits, and so on.

The classical SKDE assumes that the considered space is two dimensional, is infinite in all directions, and homogeneous. These assumptions are violated when the SKDE is applied to a spatial network. Indeed, a network can be seen as a 1.5 dimensional space because it is only possible to move along edges and to change direction at nodes (Steenberghen, Aerts, and Thomas, 2010). Using an SKDE to study a phenomenon constraint on a network presents two issues. First, the SKDE will estimate the intensity of the spatial process for the entire two-dimensional space when the events can only occur on the network space. Second, the Euclidean (or geodesic, depending on the reference system) distance will underestimate the actual distance between events and, consequently, lead to systematic overestimation of calculated intensity (Yamada and Thill, 2007). Thus, there is a strong need for a specific method for spatial analysis. Baddeley et al. (2021) propose an insightful literature review of these methods.

One of the first proposals to adapt the SKDE method to the network case is in the article of Flahaut et al. (2003). They studied hot spots of car crashes along a road network and established the basis of the density estimation along linear entities. However, the studied network was simplistic because it had only one long edge without intersections. Xie and Yang (2008) proposed another approach (referred to here as Geo-NKDE) with a genuinely geographical understanding of the problem. They proposed to estimate the intensity of the spatial process only for the centroid of lixels along the network. Lixels are small parts with similar lengths of a network edge and can be seen as the one-dimensional equivalent of a pixel. They also proposed

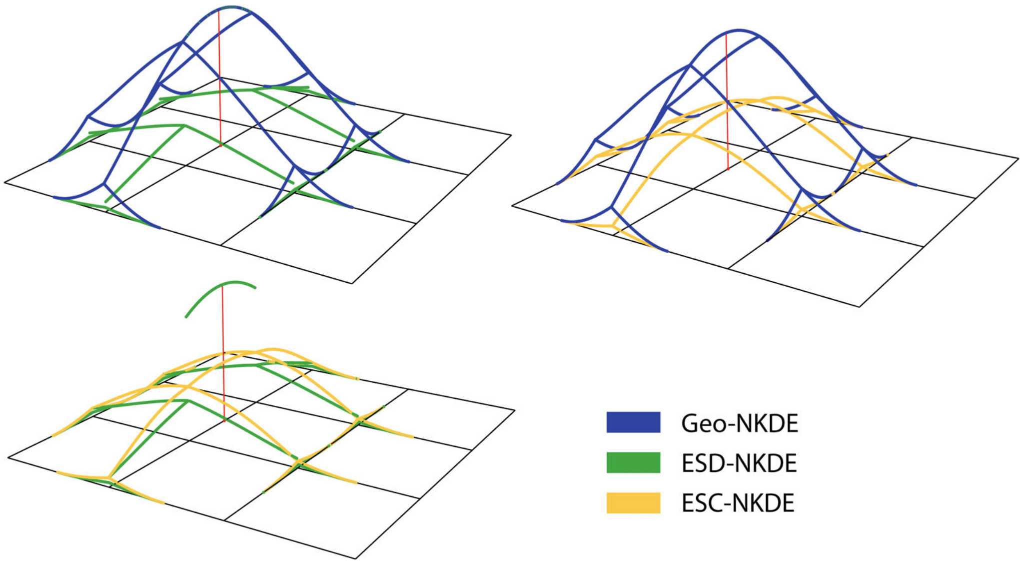

Figure 2. Comparison of GEO, ESD, and ESC NKDE.

replacing the Euclidean distance by the shortest path distance in the network in the kernel formula. This article had an important impact as this method has been used in numerous subsequent works (Li et al., 2011; Loo, Yao, and Wu, 2011; Mohaymany, Shahri, and Mirbagheri, 2013; Yu, Ai, and Shao, 2015). The method is intuitive and compatible with the previous work of Flahaut et al. (2003) on a simplistic network. However, it is statistically inexact and biased (Okabe, Satoh, and Sugihara, 2009; McSwiggan, Baddeley, and Nair, 2017). Indeed, this method multiplies the mass of events at intersections in the network. Thus, the total mass of an event is greater than one when at least one intersection lies at a distance less than the kernel's bandwidth. This is illustrated in Fig. 2 with one event (red dot) and a gridded network. The blue line represents the estimated density with bandwidth of 1.5. One can note that the density is multiplied at each intersection in every possible direction. For this example, the integral of the estimated density is 1.84, corresponding to an additional 84% of the mass of the event. As stated previously, the multiplication of the mass leads to a systematic overestimation of the density, especially in the parts of the network with many events and intersections. In this geographical approach, the kernel function is not used as a smoothing technique for the events, but rather as a function of their influence diminishing with distance.

Considering these limits, Okabe, Satoh, and Sugihara (2009) proposed two heuristics to solve the problem of the mass multiplication at intersections.

In the first approach, the mass of the events is simply divided at intersections by the numbers of lines (minus one) connected to the intersections. They call this method the Equal Split Discontinuous NKDE (ESD-NKDE). It produces sharp falls in density estimation at intersections and less smooth results. Edges of the network with many events tend to cumulate high-density values and strongly differ on produced maps. This feature can be interesting when sharp differences are desirable. However, this discontinuous nature can be counter-intuitive because, geographically speaking, it is difficult to justify the intensity of an event dropping suddenly at an intersection.

The second approach called Equal Split Continuous NKDE (ESC-NKDE) solves this issue. This heuristic divides the estimated density at encountered intersections and applies propagation correction to produce a continuous density estimation. However, it requires much more intensive calculation because it must adjust densities on edges already visited in a recursive fashion. Networks with short edges and numerous intersections involve deep recursions and extensive calculation time.

Fig. 2 illustrates the density differences obtained with the three methods. It is possible to observe that the density produced on the edges on which the event occurred is identical for the Geo-NKDE and ESD-NKDE. The ESC-NKDE produces the smoothest results.

The two NKDEs correcting the density at intersections have been widely adopted in practice (Mohaymany, Shahri, and Mirbagheri, 2013; Dai and Jaworski, 2016; Harirforoush and Bellalite, 2019; Shen et al., 2020), in part, due to its implementation in the SANET software (Okabe, Okunuki, and Shiode, 2006). The approach proposed by Xie and Yang (2008) is still used (Li et al., 2011; Lesage-Mann and Apparicio, 2016; Rui et al., 2016) because it can easily be implemented in traditional GIS software and because of its intuitive character.

Let us indicate here that McSwiggan, Baddeley, and Nair (2017) proposed another NKDE based on the heat kernel which describes heat propagation along networks. Its main appeal is its very simple calculation in comparison with the algorithm of the ESD and ESC-NKDE. It produces results equivalent to the continuous (ESC) NKDE if the latter uses a Gaussian function.

Spatiotemporal analysis

Another recent development of the SKDE method is the introduction of the temporal dimension. Most of the events happening in space also have a timestamp. Considering time is necessary to identify variations of a spatial process across years, seasons or day to day. Only considering the spatial dimension limits the analysis to a crude average of the studied spatial process for its time range.

Brunsdon, Corcoran, and Higgs (2007) proposed a first natural extension of the SKDE to include the temporal dimension. It combines the bivariate SKDE (two dimensions of geographical space) and the univariate TKDE (Temporal KDE) to obtain a TSKDE as a product of the two kernels. This approach requires the selection of two bandwidths, one for time and one for space.

Zhang et al. (2011) proposed a second approach based on the generalized product of kernels (Li and Racine, 2007). The method separates the two spatial dimensions (X and Y) to calculate the product of three kernels (X, Y and time). This formula is more flexible and can produce better adjusted non-spheric kernels when the studied phenomenon follows linear trends in space. However, this method requires the choice of three bandwidths.

Both approaches produce iso-volumes: a continuous estimation of density in a space–time cube which can be seen as a three-dimensional generalization of iso-surfaces.

Temporal dimension for NKDE

Surprisingly, there are only a few works considering the temporal dimension when calculating density of events on a network (Moradi and Mateu, 2020). Here, we give some examples of articles where a proper temporal NKDE (TNKDE) would have been more appropriate.

The article of Tang et al. (2016) used the ESD-NKDE to analyze events with a temporal dimension, that is, taxi trips. However, the proposed method does not properly consider time in the density estimation. They presented a set of maps where the ESD-NKDE was calculated for different time intervals. This is roughly equivalent to selecting a uniform kernel for the time dimension and a bandwidth equal to time intervals, with no correction for the boundary of the time intervals. This approach was also used by Yu et al. (2021) to analyze spatiotemporal use of bicycles in Beijing. Splitting time in bins is not satisfactory here considering the continuous nature of the dataset in time and contributes to overlook the temporal variation within bins.

Li et al. (2020) worked on traffic violation behavior at urban intersections and used an STKDE, despite the fact that these events can only occur on the road network. Here, the temporal dimension is well included in the analysis, but the use of the SKDE instead of a NKDE leads to the several problems described in the previous sections (estimation of the density outside of the network, underestimation of the distances between events, violation of SKDE assumptions, etc.).

Romano and Jiang (2017) recently proposed a potential TNKDE as the product of an NKDE and a TKDE. However, they used the Geo-NKDE and the TNKDE obtained has the same bias as the Geo-NKDE (multiplication of the mass at intersections).

In other words, previous research analyzing events with a temporal dimension and occurring along a network was not using suitable methods for density estimation, leading in biased estimation of the density or neglecting the temporal or spatial dimension. The analysis of spatiotemporal events occurring on a network is common in numerous research and professional fields and there is a need for more suitable methods. Thus, we focus here on the formulation of an unbiased temporal network kernel density estimate. We also illustrate the method by applying it to the case of road accidents with a pedestrian on the island of Montreal.

Methodology

KDE and SKDE

Kernel density estimation (KDE) is a nonparametric method used to estimate the probability density of a random variable on its domain from a sample of this variable. It is often used as an alternative to the histogram. The definition of the estimator ($\hat{f}_h$) for every point x on the variable's domain is given by the equation (1).

$$\hat{f}_h(x) = \frac{1}{Nh}\sum_{i=1}^{N} k\left(\frac{x - x_i}{h}\right) \tag{1}$$

where k is a kernel function which must respect the three conditions below, and x and h are scalars:

$$k(x) > 0 \text{ if } x < h$$

$$k(x) = 0 \text{ if } x \geq h$$

$$\int_{-\infty}^{+\infty} k(x) = 1$$

There are a great number of kernel functions; the best known are the Epanechnikov, Quartic, Triangle, Tricube, Triweight, Cosine, and Uniform functions. The Gaussian function is certainly the most widely used but it is defined on the domain ($-\infty$; $+\infty$) and does not respect the conditions presented above (special case of noncompact kernel). The second parameter to select for KDE is the bandwidth: the smoothing degree of the function. A larger bandwidth means that the mass of each data sample is spread over a wider interval. The goal is to select an appropriate bandwidth to represent well the variation in density of the random variable without over or underfitting it. The bandwidth is the most influential parameter of the estimated densities in this method (Turlach, 1993; Silverman, 2017).

The SKDE (Spatial KDE) is a modification of the KDE to estimate spatial process intensity observed through a sample of events located in space. More specifically, this method assumes that the observed events (e) are produced by a spatial process (p) and its intensity function (λ) can be estimated ($\hat{\lambda}$) at every location (u) of the space as:

$$\hat{\lambda}_h(u) = \frac{1}{h^2}\sum_{i=1}^{N} k\left(\frac{dist\left(u, e_i\right)}{h}\right) \tag{2}$$

with $dist\left(u, e_i\right)$ the Euclidean distance between the location u and the event e_i.

In practice, the studied space is divided using a uniform grid and the estimator $\hat{\lambda}_h(u)$ is calculated for the centroid (u) of each pixel. The SKDE is often used in geography to estimate relative risk as the log ratio of two estimated densities, the numerator being the density of the case, and the denominator the density of the control (Kelsall and Diggle, 1995; Davies, Hazelton, and Marshall, 2011).

NKDE

As stated in the introduction, the SKDE method is not well suited to analyze events constrained on a spatial network.

GEO-NKDE

The first NKDE for a network with intersections was proposed by Xie and Yang (2008). To adapt the SKDE to a network space, they proposed the following formulation for the NKDE:

$$\hat{\lambda}_h(u) = \frac{1}{h}\sum_{i=1}^{N} k\left(\frac{dist_{net}\left(u, e_i\right)}{h}\right) \tag{3}$$

with $dist_{net}\left(u, e_i\right)$ the network shortest path between the location u and the event e_i. Note that the divisor of the kernel is simply the bandwidth and not the squared bandwidth. This is because the intensity is not estimated over the surface, but rather for linear units. Moreover, Xie and Yang (2008) proposed splitting the network space into lixels (linear equivalents of a pixel) and estimating the intensity at the center of each lixel. We call this NKDE the Geo-NKDE because it is very close to other geographical methods of point pattern analysis extended to the case of networks such as the K-function or nearest neighbor analysis (Okabe and Sugihara, 2012).

ESD-NKDE and ESC-NKDE

The Geo-NKDE is intuitive and easy to implement in traditional GIS but produces biased estimates by multiplying the mass of events at intersections.

To address this issue Okabe, Satoh, and Sugihara (2009) proposed two other formulations: the Equal Split Discontinuous (ESD) and the Equal Split Continuous (ESC) NKDE.

The ESD-NKDE is easy to calculate because it only involves dividing the estimation of the density at encountered intersections (equation 4).

$$\hat{\lambda}_h\left(u, e_i\right) = k\left(dist_{net}\left(u, e_i\right)\right)\prod_{j=1}^{J}\left(\frac{1}{\left(n_{ij}-1\right)}\right) \tag{4}$$

with k the selected kernel function and n_{ij} the number of lines connected at intersection j encountered on the path between u and e_i, J the number of intersections encountered and $\widehat{\lambda}_h\left(u, e_i\right)$ the estimated density at u considering only the event e_i.

The ESC-NKDE is much more complex to calculate. Indeed, it needs to correct retroactively the estimated densities on edges previously crossed when an intersection is encountered. Okabe, Satoh, and Sugihara (2009) proposed a recursive algorithm, which can be converted into an iterative one to reduce calculation time. As with the SKDE, a ratio of two NKDE can be used to estimate relative risk on a network (McSwiggan, Baddeley, and Nair, 2020).

TNKDE

In this article, we propose the Temporal Network Kernel Density Estimate (TNKDE), an extension based on the three previous NKDEs and the generalized product of kernels.

The NKDEs can be seen as unidimensional kernels as well as the temporal KDEs. Therefore, the TNKDE is a bi-dimensional kernel:

$$\widehat{\lambda}_{h_n h_t}\left(u_{nt}\right) = \frac{1}{h_n h_t} \sum_{i=1}^{N} k_{net}\left(\frac{dist_{net}\left(u_{nt}, e_i\right)}{h_n}\right) \sum_{i=1}^{N} k_{time}\left(\frac{dist_{time}\left(u_{nt}, e_i\right)}{h_t}\right) \tag{5}$$

with h_n the bandwidth selected for the network part and h_t the bandwidth for the temporal part, u_{nt} a point at the location n on the network and t in time, and e_i an event. Note that if the density of one of the two dimensions of this kernel is 0, the spatiotemporal density will also be 0. In other words, if an event is too far in time or on the network, its contribution to the density at u_{nt} will be 0. Geo-NKDE, ESC-NKDE and ESD-NKDE can be used in this framework. Fig. 3 illustrates the ESC-TNKDE over time.

It is also possible to use a different kernel function for the network and time dimensions if needed (special case where $k_{net} \neq k_{time}$). In this article, we will only work with the general case, using the same kernel function for the two dimensions.

Diggle's correction

For the classical SKDE, the estimations of the density assume that the studied space is infinite in every direction. However, in most cases, the events are limited to a finite sampling area. At the borders of the study area, the events are scarcer because the part beyond the border is

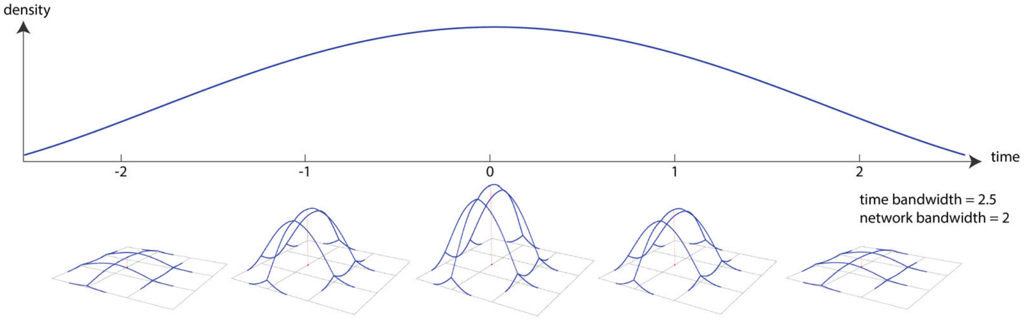

Figure 3. Product of the temporal and network kernel densities.

not sampled. Also, a part of the density of the sampled events is lost beyond the border. This concern is also valid for networks and for sampling periods of study, resulting in biased density estimation.

The classical approach to reduce this bias for the SKDE is to apply the Diggles' correction (Diggle, 1985). It increases the weights of the events located close to the boundary of the study area because a part of their mass is lost beyond the border. This additional weight is calculated as the inverse of the mass of the event inside the study area.

For the TNKDE, the Diggle's correction applied to an event e_{nt} can be calculated as follow:

$$D\left(e_{nt}\right) = \left[1 - \left(\int_{Net'}^{w} k_{net}\left(\frac{dist_{net}\left(e_{nt}, w\right)}{h_n}\right) \times \int_{Time'}^{v} k_{time}\left(\frac{dist_{time}\left(e_{nt}, v\right)}{h_t}\right)\right)\right]^{-1} \tag{6}$$

with w a point on the network *Net* and v a moment in the temporal window *Time*. *Net'* and *Time'* are the parts of the network and time outside the study domain. It can be read as the inverse of the fraction of the density of the event within the space–time boundaries.

Thus, if half of the mass of an event falls beyond the studied network and a third of its mass falls beyond the studied period, then the weight to apply is: $\left(1 - \frac{1}{2} \times \frac{1}{3}\right)^{-1} = 1.2$.

Adaptive bandwidth

Classical kernel density estimation uses a fixed bandwidth. This constraint can be problematic if the intensity of the spatial process varies in space or time. The results obtained could be oversmoothed in areas with many events and undersmoothed in areas with few events, producing biased density estimates. For the SKDE, one can use adaptive bandwidth which varies in space. The Abramson (1982) method adjusts the bandwidths locally according to an a priori estimated density using a pilot global bandwidth h_0:

$$h\left(u_{ei}\right) = h_0 \frac{1}{\sqrt{\tilde{f}\left(u_{ei}\right)/\gamma}} \text{ with } \gamma = \left(\prod_{i=1}^{n} \frac{1}{\sqrt{\tilde{f}\left(e_i\right)}}\right)^{\frac{1}{n}} \tag{7}$$

with $h\left(u_{ei}\right)$ the local bandwidth defined for the event e_i located at u, h_0, the pilot bandwidth and $\tilde{f}\left(u_{ei}\right)$ the estimated density at point u.

It is recommended to define a trimming value to avoid excessive bandwidth in regions with very rare events. This method can be directly applied to the different NKDEs without modifications. For the TNKDE, two options can be considered: separated or simultaneous adjustment of the network and time bandwidths.

In the case of a separated adjustment, the network kernel density is calculated for each event without considering the time dimension and using a network pilot bandwidth. This density is then used to calculate the local network bandwidths. The same logic is used for the temporal bandwidth (equation 8). Thus, isolated events on the network will get a larger bandwidth, as will isolated events in time. This first approach is more appropriate when the analyzed process is not characterized by a strong spatiotemporal autocorrelation and a weak interaction between the network and temporal dimensions.

$$h_{net}\left(u_{ei}\right) = h_{net0} \frac{1}{\sqrt{\widetilde{f_{net}}\left(u_{ei}\right)/\gamma_{net}}}; h_{time}\left(u_{ei}\right) = h_{time0} \frac{1}{\sqrt{\widetilde{f_{time}}\left(u_{ei}\right)/\gamma_{time}}} \tag{8}$$

Simultaneous adjustment of the two bandwidths is obtained by calculating the spatiotemporal density at each event with the two pilot bandwidths, and then adjusting both network and temporal local bandwidths (the same factor is used to multiply the temporal and network bandwidths) with this density (equation 9). In this approach, the temporal and network dimensions are interacting.

$$h_{net}\left(u_{ei}\right) = h_{net0}\frac{1}{\sqrt{\tilde{f}\left(u_{ei}\right)/\gamma}}; h_{time}\left(u_{ei}\right) = h_{time0}\frac{1}{\sqrt{\tilde{f}\left(u_{ei}\right)/\gamma}} \tag{9}$$

The k-nearest neighbors method has also been proposed to get local bandwidths varying with local event density and is known to produce interesting results for multivariate kernels (Loftsgaarden and Quesenberry, 1965; Terrell and Scott, 1992; Orava, 2011; Langrené and Warin, 2019). However, adapting this approach to the TNKDE case is complicated because it requires defining a spatiotemporal distance and it raises many methodological issues. It is probably easier to employ a two-step approach:

1. For an event, selecting its k-nearest neighbors on the network and keeping the distance to the latest as the local bandwidth for the network dimension.
2. Calculating the median of the temporal distance between the event and its neighbors. This value could be used as the local bandwidth for the temporal dimension and could be trimmed to not consider points too distant in time.

This proposal is focused on the network dimension (selected first) but considers its interaction with the temporal dimension. This is justified by the fact that kernel density estimate methods are used mainly to produce maps in geography.

Bandwidth selection

The bandwidth is the most important parameter when applying a KDE. Several methods have been proposed to select an optimal bandwidth (Turlach, 1993) and the leave-on-out maximum likelihood is probably the most popular. It is a data-driven method which can be easily adapted to multivariate kernels and thus to spatiotemporal KDE (Davies and Lawson, 2019). The idea here is to maximize the sum of the log densities observed at each event location, if that event was missing (equation 10).

$$likelihood(h) = n^{-1}\sum_{i=1}^{n}\log\left\{\tilde{f}_{X^{[-i]}}\left(x_i|h\right)\right\} \tag{10}$$

with h the bandwidth evaluated.

For a spatiotemporal kernel, two bandwidths must be selected and evaluated simultaneously as follows:

$$likelihood\left(h_{net}, h_{time}\right) = n^{-1}\sum_{i=1}^{n}\log\left\{\tilde{f}_{X^{[-i]}}\left(x_i|h_{net}, h_{time}\right)\right\} \tag{11}$$

Because densities estimated by a multidimensional kernel are smaller (the mass is distributed over several dimensions), the selected bandwidths with this method tend to be large. It is important to visually inspect the estimated densities with the selected bandwidths and to compare them with prior knowledge.

Application to real data

Data

We apply the proposed TNKDE to road crashes with a pedestrian in Montreal between January 1, 2016 and December 31, 2019. This dataset is freely available on the open data website of the City of Montreal and, thus, on the road network used (named as *Geobase*). Because we are interested in accidents involving a pedestrian, highways were removed from the road network. Fig. 4 illustrates the two datasets.

Descriptive analysis

There were 1,280 accidents in 2016, 1,260 in 2017, 1,279 in 2018, and 1,315 in 2019, for a total of 5,134. They are evenly spread over the four years and there is no indication of any significant increase or decrease during the period. We know each accident's day of occurrence, so the temporal base unit will be the day.

Fig. 5 represents the densities of accidents over time with a bandwidth of 21 days (selected using a cross-validation method). One can note that the density is lowest during the summer, which can be explained by the fact that the city closes several main streets to cars at that time and by the lowest level of activity due to vacations. The density increases sharply during autumn and reaches a peak in December. People coming back from vacation and the winter weather are the two main drivers of this phenomenon. The dataset is strongly seasonal.

The Ripley's K-function applied to the dataset (Fig. 6) confirms that the events are more clustered than expected from a random distribution ($P < 0.01$ with 1,000 permutations). The clustering is the highest in a radius of 45 days and is not significant after 100 days. Note that the Fig. 6 is displaying the difference between the observed K-function and the 99th percentile calculated from 1000 random permutations.

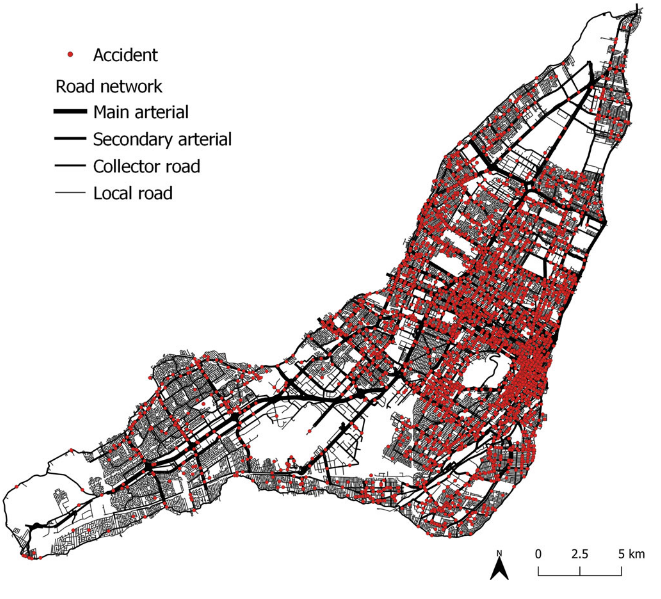

Figure 4. Road crash with a pedestrian between 2016 and 2019.

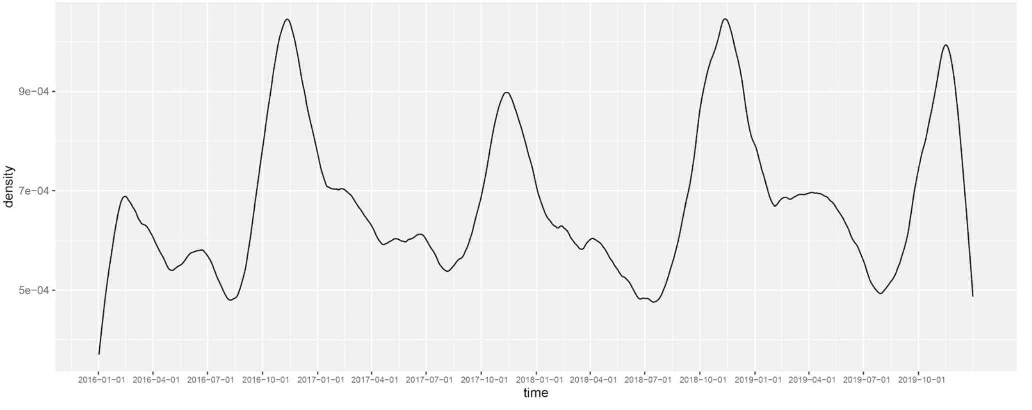

Figure 5. Estimated temporal density of accidents.

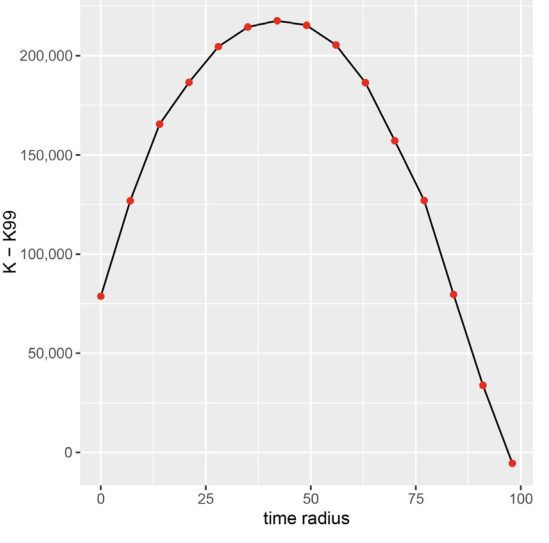

Figure 6. Ripley's K-function of accidents in time.

Fig. 7 is a map of the estimated density of events with the ESC-NKDE (multiplied by 1,000 to increase legend readability) and a Quartic kernel function. A bandwidth of 800 m was selected by leave-one-out cross validation. All bandwidths from 50 to 1,500 m were tested with a step of 50 m. To reduce computation time, the ESD-NKDE was used for this step. The Diggle correction at the borders was not applied because the network is naturally confined to an island. Considering the high variation of events concentration in space, we used adaptive bandwidths (equation 9) with a maximum value of 1,500 m. The quartiles of the obtained local bandwidths are 618 m (25%), 832 m (50%), and 979 m (75%).

Calculating the TNKDE

For the TNKDE, the network and temporal bandwidths were selected with the leave-one-out-cross validation method. Again, the ESD-TNKDE was used to reduce calculation time during the bandwidth selection.

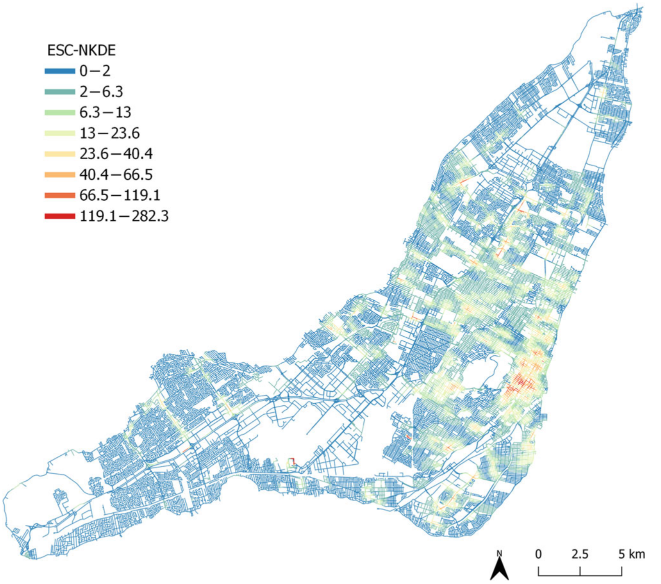

Figure 7. Map of the estimated density of accidents with the ESC-NKDE.

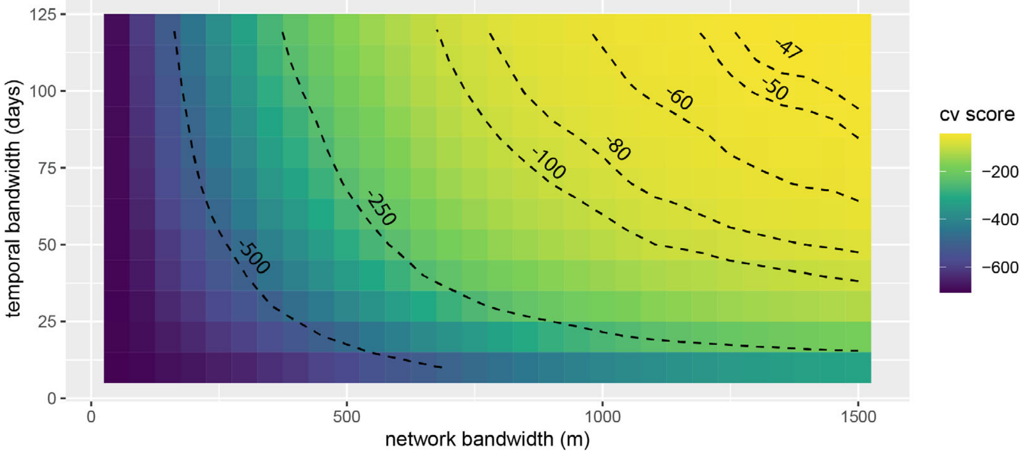

Figure 8. Leave one out cross validation scores for pairs of network and temporal bandwidths for the ESD-TNKDE.

Fig. 8 shows that beyond 1,000 m and 70 days, increasing the bandwidths leads to small gains for the selection criterion. The gain between 1,000 m and 70 days and 1,500 m and 115 days is less than 10% of the gain between 300 m and 30 days. Selecting large bandwidths tends also to produce oversmoothed results, especially for the temporal dimension. We compared several sets of bandwidths and finally selected 1,000 m and 55 days as global bandwidths. We also use adaptive bandwidths adjusted simultaneously to limit the impact of strong spatiotemporal variation of events density. The maximum of the network bandwidths was set to 1,500 m and

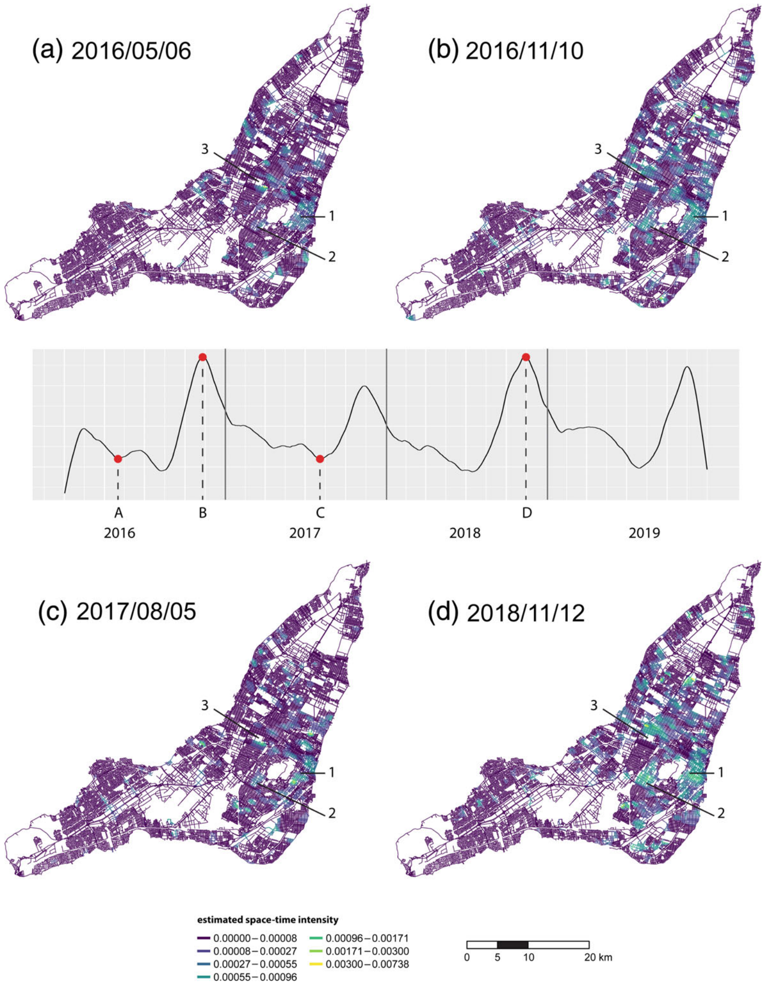

Figure 9. Mosaic maps of accidents' estimated density in Montreal.

70 days for the temporal bandwidths. The quartiles of the obtained local bandwidths are 854 m (25%), 1,020 m (50%), and 1,185 m (75%) and 47 days (25%), 56 days (50%), and 65 days (75%).

To obtain detailed results, the network was split into lixels of 150 m and into five day periods.

Interpreting the results

To analyze the results of a TNKDE, the most efficient approach is to produce animated maps (GIF or video). In traditional GIS software, it is also possible to display features according to a specific time range and to play with a slider for the temporal dimension. Here we propose four maps during two periods of high and low density of accidents for the entire region (Fig. 9). As Supporting Information, we also provide a video to visualize full TNKDE.

The differences between the low and high density periods are evident on the maps. During the entire period studied, the City Center (1) is characterized by continuously high levels of estimated density. Lixels close to the University of Montreal (2) and Jarry Park (3) experience oscillations between periods of higher and lower density, matching the regional pattern.

The maps A and C represent two periods of lower density for the whole region; however, we can distinguish local hot spots of accidents with high-density values. This highlights the need to not limit the analysis to periods with high-density values at the regional level. The animated maps are very efficient tools to identify such hot or cold spots and their behavior over time: amplification, stability, oscillation, or decrease.

Conclusion

In this article, we presented a temporal extension of the NKDE based on the general product of kernels. It can be used to estimate spatiotemporal density of phenomena occurring on a network and over time. This specific type of dataset is often encountered in several research and professional fields (road crashes, delays in public transit, leaks along pipes, species distributions along rivers, etc.). However, the methods commonly used to date were not commonly used on such datasets, neglecting the network or temporal dimension.

The results of the TNKDE can easily be visualized and interpreted with animated maps. Moreover, the method produces a regular and complete time-series for each lixel, which can be used for further analysis. Indeed, the principal limitation of kernel density methods is to produce only descriptive results. Here, we propose two avenues to deepen their analysis:

1. The detection of clusters based on the spatiotemporal versions of local autocorrelation indexes (LISA and Getis and Ord G). This has already been proposed for the NKDE (Yamada and Thill, 2007; Xie and Yan, 2013) and could be combined with the method of emerging hot spots analysis proposed by ESRI (2021).
2. The unsupervised clustering of produced time-series to identify groups of lixels sharing similar temporal patterns. With an appropriate metric, one could even use the recently developed algorithms for unsupervised clustering taking spatial distance between observations into account (Chavent et al., 2018; Gelb and Apparicio, 2021).

To conclude this article, we would stress again that data describing events occurring on a network with a temporal dimension are very common. It is, thus, necessary that appropriate methods and tools be made available to analyze them properly. The code used to generate the results in this article is available in Supporting Information. More specifically, we used the open source software R (R Core Team, 2020) and the package *spNetwork* (Gelb, 2021), to compute the TNKDE presented in this article.

References

Abdulhafedh, A. (2017). “Identifying Vehicular Crash High Risk Locations along Highways Via Spatial Autocorrelation Indices and Kernel Density Estimation.” *World Journal of Engineering and Technology* 05, 198. https://doi.org/10.4236/wjet.2017.52016

Abramson, I. S. (1982). “On Bandwidth Variation in Kernel Estimates – A Square Root Law.” *Annals of Statistics* 10, 1217–23. https://doi.org/10.1214/aos/1176345986

Baddeley, A., G. Nair, S. Rakshit, G. McSwiggan, and T. M. Davies. (2021). “Analysing Point Patterns on Networks – A Review.” *Spatial Statistics* 42, 100435. https://doi.org/10.1016/j.spasta.2020.100435

Barthélemy, M. (2011). “Spatial Networks.” *Physics Reports* 499, 1–101. https://doi.org/10.1016/j.physrep.2010.11.002

Brunsdon, C., J. Corcoran, and G. Higgs. (2007). "Visualising Space and Time in Crime Patterns: A Comparison of Methods." *Computers Environment and Urban Systems* 31, 52–75. https://doi.org/10.1016/j.compenvurbsys.2005.07.009

Chavent, M., V. Kuentz-Simonet, A. Labenne, and J. Saracco. (2018). "ClustGeo: An R Package for Hierarchical Clustering with Spatial Constraints." *Computational Statistics* 33, 1799–822. https://doi.org/10.1007/s00180-018-0791-1

Dai, D., and D. Jaworski. (2016). "Influence of Built Environment on Pedestrian Crashes: A Network-Based GIS Analysis." *Applied Geography* 73, 53–61. https://doi.org/10.1016/j.apgeog.2016.06.005

Davies, T. M., M. L. Hazelton, and J. C. Marshall. (2011). "Sparr: Analyzing Spatial Relative Risk Using Fixed and Adaptive Kernel Density Estimation in R." *Journal of Statistical Software* 39, 1–14. https://doi.org/10.18637/jss.v039.i01

Davies, T. M., and A. B. Lawson. (2019). "An Evaluation of Likelihood-Based Bandwidth Selectors for Spatial and Spatiotemporal Kernel Estimates." *Journal of Statistical Computation and Simulation* 89, 1131–52. https://doi.org/10.1080/00949655.2019.1575066

Diggle, P. (1985). "A Kernel Method for Smoothing Point Process Data." *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 34, 138–47. https://doi.org/10.2307/2347366

ESRI. (2021). "Emerging Hot Spot Analysis (Space Time Pattern Mining) – ArcGIS Pro|Documentation [WWW Document]". https://pro.arcgis.com/en/pro-app/latest/tool-reference/space-time-pattern-mining/emerginghotspots.htm.

Flahaut, B., M. Mouchart, E. S. Martin, and I. Thomas. (2003). "The Local Spatial Autocorrelation and the Kernel Method for Identifying Black Zones: A Comparative Approach." *Accident Analysis & Prevention* 35, 991–1004. https://doi.org/10.1016/S0001-4575(02)00107-0

Gelb, J. (2021). "spNetwork: A Package for Network Kernel Density Estimation." *R Journal* 13, 561–77. https://doi.org/10.32614/RJ-2021-102

Gelb, J., and P. Apparicio. (2021). "Apport de la classification floue c-means spatiale en géographie: essai de taxinomie socio-résidentielle et environnementale à Lyon." *Cybergeo: European Journal of Geography*. https://doi.org/10.4000/cybergeo.36414

Harirforoush, H., and L. Bellalite. (2019). "A New Integrated GIS-Based Analysis to Detect Hotspots: A Case Study of the City of Sherbrooke." *Accident; Analysis and Prevention* 130, 62–74. https://doi.org/10.1016/j.aap.2016.08.015

Kelsall, J. E., and P. J. Diggle. (1995). "Kernel Estimation of Relative Risk." *Bernoulli* 1, 3–16. https://doi.org/10.2307/3318678

Langrené, N., and X. Warin. (2019). "Fast and Stable Multivariate Kernel Density Estimation by Fast Sum Updating." *Journal of Computational and Graphical Statistics* 28, 596–608. https://doi.org/10.1080/10618600.2018.1549052

Lesage-Mann, E., and P. Apparicio. (2016). "Geography of Discarded Needles in Montreal." *Cahiers de géographie de Québec* 60, 551–73. https://doi.org/10.7202/1041222ar

Li, Q., and J. S. Racine. (2007). *Nonparametric Econometrics: Theory and Practice*. Princeton, NJ: Princeton University Press.

Li, Q., T. Zhang, H. Wang, and Z. Zeng. (2011). "Dynamic Accessibility Mapping Using Floating Car Data: A Network-Constrained Density Estimation Approach." *Journal of Transport Geography* 19, 379–93. https://doi.org/10.1016/j.jtrangeo.2010.07.003

Li, Y., M. Abdel-Aty, J. Yuan, Z. Cheng, and J. Lu. (2020). "Analyzing Traffic Violation Behavior at Urban Intersections: A Spatio-Temporal Kernel Density Estimation Approach Using Automated Enforcement System Data." *Accident Analysis & Prevention* 141, 105509. https://doi.org/10.1016/j.aap.2020.105509

Loftsgaarden, D. O., and C. P. Quesenberry. (1965). "A Nonparametric Estimate of a Multivariate Density Function." *Annals of Mathematical Statistics* 36, 1049–51. https://doi.org/10.1214/aoms/1177700079

Loo, B. P. Y., S. Yao, and J. Wu. (2011). "Spatial Point Analysis of Road Crashes in Shanghai: A GIS-based Network Kernel Density Method." In *2011 19th International Conference on Geoinformatics*. Presented at the 2011 19th International Conference on Geoinformatics, 1–6. https://doi.org/10.1109/GeoInformatics.2011.5980938

Mala, S., and M. K. Jat. (2019). “Geographic Information System Based Spatio-Temporal Dengue Fever Cluster Analysis and Mapping.” *The Egyptian Journal of Remote Sensing and Space Science* 22, 297–304. https://doi.org/10.1016/j.ejrs.2019.08.002

McSwiggan, G., A. Baddeley, and G. Nair. (2017). “Kernel Density Estimation on a Linear Network.” *Scandinavian Journal of Statistics* 44, 324–45. https://doi.org/10.1111/sjos.12255

McSwiggan, G., A. Baddeley, and G. Nair. (2020). “Estimation of Relative Risk for Events on a Linear Network.” *Statistics and Computing* 30, 469–84. https://doi.org/10.1007/s11222-019-09889-7

Mohaymany, A. S., M. Shahri, and B. Mirbagheri. (2013). “GIS-Based Method for Detecting High-Crash-Risk Road Segments Using Network Kernel Density Estimation.” *Geo-Spatial Information Science* 16, 113–9. https://doi.org/10.1080/10095020.2013.766396

Moradi, M. M., and J. Mateu. (2020). “First- and Second-Order Characteristics of Spatio-Temporal Point Processes on Linear Networks.” *Journal of Computational and Graphical Statistics* 29, 432–43. https://doi.org/10.1080/10618600.2019.1694524

Okabe, A., K.-I. Okunuki, and S. Shiode. (2006). “The SANET Toolbox: New Methods for Network Spatial Analysis.” *Transactions in GIS* 10, 535–50. https://doi.org/10.1111/j.1467-9671.2006.01011.x

Okabe, A., T. Satoh, and K. Sugihara. (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. https://doi.org/10.1080/13658810802475491

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

Orava, J. (2011). “K-Nearest Neighbour Kernel Density Estimation, the Choice of Optimal K.” *Tatra Mountains Mathematial Publiations* 50, 39–50. https://doi.org/10.2478/v10127-011-0035-z

R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Romano, B. and Z. Jiang (2017). “Visualizing Traffic Accident Hotspots Based on Spatial-Temporal Network Kernel Density Estimation.” In *Proceedings of the 25th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, SIGSPATIAL'17*. Association for Computing Machinery, New York, NY, USA, 1–4. https://doi.org/10.1145/3139958.3139981

Rui, Y., Z. Yang, T. Qian, S. Khalid, N. Xia, and J. Wang. (2016). “Network-Constrained and Category-Based Point Pattern Analysis for Suguo Retail Stores in Nanjing, China.” *International Journal of Geographical Information Science* 30, 186–99. https://doi.org/10.1080/13658816.2015.1080829

Setiawan, I., M. Dede, D. Sugandi, and M. Agung Widiawaty. (2019). “Investigating Urban Crime Pattern and Accessibility Using Geographic Information System in Bandung City.” *KnE Social Sciences*, 3, 535–48. https://doi.org/10.18502/kss.v3i21.4993

Shad, R., and S. Rahimi. (2017). “Identification of Road Crash Black-Sites Using Geographical Information System.” *International Journal of Traffic and Transportation Engineering* 7, 368–80.

Shen, B., X. Xu, J. Li, A. Plaza, and Q. Huang. (2020). “Unfolding Spatial-Temporal Patterns of Taxi Trip Based on an Improved Network Kernel Density Estimation.” *ISPRS International Journal of Geo-Information* 9, 683. https://doi.org/10.3390/ijgi9110683

Shi, B., X.-M. Zhan, J.-X. Zheng, H. Qiu, D. Liang, Y.-M. Ye, G.-J. Yang, Y. Liu, and J. Liu. (2018). “Identifying Key Bird Species and Geographical Hotspots of Avian Influenza A (H7N9) Virus in China.” *Infectious Diseases of Poverty* 7, 97. https://doi.org/10.1186/s40249-018-0480-x

Silverman, B. W. (2017). *Density Estimation for Statistics and Data Analysis*. Boca Raton: Routledge. https://doi.org/10.1201/9781315140919

Steenberghen, T., K. Aerts, and I. Thomas. (2010). “Spatial Clustering of Events on a Network.” *Journal of Transport Geography* 18, 411–8. https://doi.org/10.1016/j.jtrangeo.2009.08.005

Tang, L., Z. Kan, X. Zhang, F. Sun, X. Yang, and Q. Li. (2016). “A Network Kernel Density Estimation for Linear Features in Space–Time Analysis of Big Trace Data.” *International Journal of Geographical Information Science* 30, 1717–37. https://doi.org/10.1080/13658816.2015.1119279

Terrell, G. R., and D. W. Scott. (1992). “Variable Kernel Density Estimation.” *Annals of Statistics* 20, 1236–65. https://doi.org/10.1214/aos/1176348768

Turlach, B. A. (1993). *Bandwidth Selection in Kernel Density Estimation: A Review*. Citeseer: CORE and Institut de Statistique.

Xie, Z., and J. Yan. (2008). "Kernel Density Estimation of Traffic Accidents in a Network Space." *Computers, Environment and Urban Systems* 32, 396–406. https://doi.org/10.1016/j.compenvurbsys.2008.05.001

Xie, Z., and J. Yan. (2013). "Detecting Traffic Accident Clusters with Network Kernel Density Estimation and Local Spatial Statistics: An Integrated Approach." *Journal of Transport Geography* 31, 64–71. https://doi.org/10.1016/j.jtrangeo.2013.05.009

Yamada, I., and J.-C. Thill. (2007). "Local Indicators of Network-Constrained Clusters in Spatial Point Patterns." *Geographical Analysis* 39, 268–92. https://doi.org/10.1111/j.1538-4632.2007.00704.x

Yu, Q., Y. Gu, S. Yang, and M. Zhou. (2021). "Discovering Spatiotemporal Patterns and Urban Facilities Determinants of Cycling Activities in Beijing." *Journal of Geovisualization and Spatial Analysis* 5, 16. https://doi.org/10.1007/s41651-021-00084-9

Yu, W., T. Ai, and S. Shao. (2015). "The Analysis and Delimitation of Central Business District Using Network Kernel Density Estimation." *Journal of Transport Geography* 45, 32–47. https://doi.org/10.1016/j.jtrangeo.2015.04.008

Zhang, G., A.-X. Zhu, S. K. Windels, and C.-Z. Qin. (2018). "Modelling Species Habitat Suitability from Presence-Only Data Using Kernel Density Estimation." *Ecological Indicators* 93, 387–96. https://doi.org/10.1016/j.ecolind.2018.04.002

Zhang, Z., D. Chen, W. Liu, J. S. Racine, S. Ong, Y. Chen, G. Zhao, and Q. Jiang. (2011). "Nonparametric Evaluation of Dynamic Disease Risk: A Spatio-Temporal Kernel Approach." *PLoS One* 6, e17381. https://doi.org/10.1371/journal.pone.0017381

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web site.