Spatial analysis of roadside Acacia populations on a road network using the network K-function

Peter G. Spooner1,*, Ian D. Lunt1, Atsuyuki Okabe2 and Shino Shiode2
1The Johnstone Centre, Charles Sturt University, PO Box 789, Albury NSW 2640, Australia; 2Centre for Spatial Information Science, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan;
*Author for correspondence (e-mail: pspooner@csu.edu.au)

Received 9 April 2003; accepted in revised form 9 October 2003

Key words: Acacia, Anthropogenic disturbance, Field margins, Kernel estimation, Road verge, Stream ecology

Abstract

Spatial patterning of plant distributions has long been recognised as being important in understanding underlying ecological processes. Ripley’s K-function is a frequently used method for studying the spatial pattern of mapped point data in ecology. However, application of this method to point patterns on road networks is inappropriate, as the K-function assumes an infinite homogenous environment in calculating Euclidean distances. A new technique for analysing the distribution of points on a network has been developed, called the network K-function (for univariate analysis) and network cross K-function (for bivariate analysis). To investigate its applicability for ecological data-sets, this method was applied to point location data for roadside populations of three Acacia species in a fragmented agricultural landscape of south-eastern Australia. Kernel estimations of the observed density of spatial point patterns for each species showed strong spatial heterogeneity. Combined univariate and bivariate network K-function analyses confirmed significant clustering of populations at various scales, and spatial patterns of Acacia decora suggests that roadworks activities may have a stronger controlling influence than environmental determinants on population dynamics. The network K-function method will become a useful statistical tool for the analyses of ecological data along roads, field margins, streams and other networks.

Introduction

Roads are transport corridors imposed on landscapes by humans for the movement of people and materials. Owing to their relatively large area, roads have a significant effect on the natural environment (Bennett 1991). Roads affect landscape structure and species survival through destruction, fragmentation and alteration of habitat, edge effects, disruption to animal movement patterns, pollution and traffic mortality (Way 1977; Andrews 1990; Forman and Alexander 1998; Carr et al. 2002). Conversely, in many fragmented landscapes, roadside environments (commonly known as road verges or road shoulders) provide important refuges for isolated populations of threatened native animal and plant species (e.g., Cale 1990; Vermeulen and Opdam 1995). The science of road ecology has recently emerged to improve our understanding of road network environments, but empirical studies have been frustrated by the lack of suitable spatial tools to analyse point patterns on road networks (Forman 1999).

The K-function (Ripley 1976, 1981) is a widely used method for studying the spatial pattern of mapped point data in plant ecology (Haase 1995; Pordan and Czárán 1997; Dale 1999). Applications include analyses of distribution patterns of Eucalyptus species (West 1984), Pinus ponderosa (Getis and Franklin 1987), Pinus banksiana (Kenkel 1988), mixed deciduous-coniferous forests (Szwagrzyk and Czerwczak 1993) and sub-tropical trees (Debski et al. 2000). The advantage of K-function analysis is that it
uses all point-point distances, not just the nearest neighbour distances, to show spatial clustering at various scales of pattern, and the distance where clustering or over-dispersal becomes significant (Bailey and Gatrell 1995). However like most spatial statistical methods, the K-function assumes a homogeneous environment to calculate the Euclidean distance between points (or straight-line distance “as the crow flies”), and therefore is an inappropriate tool for analysing point patterns confined along irregular road networks (Little et al. 1997; Rathbun 1998; Okabe and Yamada 2001). When used to analyse spatial point patterns constrained by road networks, the K-function can result in over-detection of clustering patterns, leading to possible Type I errors (Okabe et al. 2002a; Yamada and Thill 2003). Therefore an application of the K-function modified for a network system of nodes and linkages would be useful to analyse spatial point patterns along roads.

A method to conduct K-function analysis of point patterns on a network was recently developed by Okabe and Yamada (2001), for the analysis of social-economic and micro-marketing applications (e.g., traffic accidents and market area delimitation). Using this method, we analysed the spatial patterns of three Acacia species with different life-history attributes, to investigate the effects of disturbance from roadworks on roadside Acacia populations. These three Acacia species are dominant shrubs in roadside remnants in the study area: Acacia pycnantha (Golden Wattle) is an obligate seeder, A. montana (Mallee Wattle) a facultative seeder, and A. decora (Western Silver Wattle) is a facultative seeder. All three Acacia species are leguminous shrubs 1–8 m tall (Tame 1992). Spatial analysis of roadside Acacia shrub populations may reveal important insights into anthropogenic interrelations (Motzkin et al. 1996), an issue critical to conservation of these unique environments. The aims of this paper are to (1) draw the network K-function method to the attention of terrestrial and aquatic ecologists, and (2) illustrate the usefulness of this method by examining the spatial point patterns of Acacia species along roadsides in a fragmented agricultural landscape in south-eastern Australia.

**Materials and methods**

The development of agriculture in south-eastern Australia has resulted in the clearing of over 85% of original native vegetation, and most patches of intact native vegetation occur along roadsides (Yates and Hobbs 1997). This process has created a system of cropping and grazing paddocks hostile to most native species, separated by narrow road corridors with explicit boundaries, containing road verge populations of varying quality (Hobbs and Saunders, 1994). Previous studies on the effects of soil disturbance on roadside Acacia populations showed evidence of soil disturbance from previous roadworks activities in over 90% of Acacia populations, and a significant relationship between major recruitment pulses in Acacia populations and road construction events. For roadside Acacia populations, soil disturbance along the road network from roadworks is analogous to periodic disturbance from a natural fire regime, vital for the maintenance of remaining populations (Spooner et al. 2004).

**Data collection**

One hundred and thirty-five Acacia populations (47 A. pycnantha, 48 A. montana and 40 A. decora) were selected from existing roadside survey data (Bull 1997) using a random sampling procedure (Spooner et al. 2004), in the Lockhart Shire study area, New South Wales (35.12°S, 146.43°E), in south-eastern Australia (Figure 1). The region consists of low undulating hills and flat riverine plains, and the vegetation is dominated by open woodlands of Eucalyptus species and Callitris glaucophylla. Roadside Acacia populations in the region were located and GPS coordinates from the centre of each population were recorded. Data was then entered into ESRI ArcGIS 8.2.

To visually inspect spatial point distributions, a Kernel estimation (Bailey and Gatrell 1995) of the density of each Acacia population was constructed, with a bandwidth of 4 km, using the Animal Movement Analysis GIS extension (Hooge 2002). The Kernel process constructs a set of probability polygons that represents a histogram of the intensity of the spatial point pattern, to illustrate clumping of point patterns of each Acacia species (Bailey and Gatrell 1995).

**Network K-function background and analysis**

The reduced second moment measure or K-function uses all point-to-point Euclidean distances to analyse a two-dimensional distribution pattern on an infinite homogeneous plane, where a circle of radius r is cen-
tred on each point \( p_1, \ldots, p_n \) and the number of neighbouring points within the circle are counted (Figure 2a). By varying the circle radius \( t \), spatial pattern can be examined at various scales, and the deviation of the observed number from expected number of points is then plotted against \( t \). If the points are randomly distributed, the K-function \( K(t) = \pi t^2 \) for all values of \( t \), and the null hypothesis for complete spatial randomness (CSR) is accepted (Ripley 1981; Diggle 1983). If there is significant deviation of the observed function \( \hat{K}(t) \) from a randomly generated (Poisson) point process \( K(t) \), the null hypothesis for CSR is rejected (Haase 1995). \( \hat{K}(t) \) is often plotted as:

\[
\hat{L}(t) = \sqrt{\frac{\hat{K}(t)}{\pi} - t},
\]

(for further details see Bailey and Gatrell 1995; Haase 1995).

The univariate network K-function (Miller 1994, 1999; Okabe and Kitamura 1996; Okabe and Yamada 2001) follows a similar process, but calculates the shortest path distance (e.g., \( t_2 < t_1 \)) from each point to all other points \( p_1, \ldots, p_n \) on a finite connected planar network (Figure 2b). Network links are defined as \( L_T \), and the total network distance as \( IL_T \). In this case, \( L_T \) is the set of road network links in the Lockhart Shire study area. Sets of points on the road network \( P = \{p_1, \ldots, p_n\} \) are assumed to be located on road network links \( L_T \). Okabe and Yamada's model (2001) defines the network K-function \( K(t) \) by:

\[
K(t) = \frac{1}{\rho} \frac{1}{E} \left( \text{the number of points } P \text{ within network distance } t \text{ to a point } p_i \text{ of } P \right),
\]

where \( E(\cdot) \) is the expected value with respect to \( p_1, \ldots, p_n \) (\( p_i \in P \)), which follow the binomial point process, and \( \rho \) is the density of points \( P \), such that \( \rho = n / IL_T \). The assumption of the binomial point process is based on the hypothesis that points \( P \) are uniformly and independently distributed over a finite road network. Thus if this hypothesis is rejected, points \( P \) are spatially interacting and may form non-uniform patterns. From Equation (2), the observed network K-function, \( \tilde{K}(t) \) for the observed data set \( P \), can be written as:

\[
\tilde{K}(t) = \frac{|L_T|}{n(n-1)} \sum_{i=1}^{n} \delta_{t_i, t},
\]

(the number of points of \( P \) on \( L_{t_i} \)).
Assessment of the significance of departure of the network \( R(t) \) from CSR corresponds to that of the 'normal' K-function application. 100 Monte Carlo simulations were used to construct a confidence envelope based on the maximum and minimum values from an equivalent number of random co-ordinates for \( K(t) \), and \( R(t) \) is compared to this envelope. If values of \( R(t) \) lie within the confidence envelope, points \( P \) are randomly distributed. Conversely, any values of \( R(t) \) which lie outside the confidence envelope were considered a significant deviation from CSR. If values of \( R(t) > K(t) \) and lie outside the upper confidence interval, then points \( P \) tend to be clustering. Conversely, if values of \( R(t) < K(t) \) and lie outside the lower confidence interval, then points are tending towards regularity (Kenkel 1988; Bailey and Gatrell 1995). Since edge effects (Haase 1995) are taken into account with distance computations, no edge adjustment factor is necessary (Okabe and Yamada 2001).

For bivariate network K-function analyses, two kinds of points, \( A = \{a_1, ..., a_n\} \) and \( B = \{b_1, ..., b_m\} \) are analysed on \( L_P \). In this case, we are interested in the hypothesis of spatial interaction between occurrences of different Acacia species. Okabe and Yamada (2001) define the network cross K-function \( K^{ab}(t) \) as:

\[
K^{ab}(t) = \frac{1}{\rho_a} \mathbb{E} \left( \frac{\text{the number of points } A \text{ within network distance } r \text{ to a point } b \text{ in } B}{\text{network distance } r} \right).
\]

(4)

where \( \mathbb{E}(\cdot) \) is the expected value with respect to \( b_1, ..., b_m, (b \in B) \), which follow the binomial point process, and \( \rho_a \) is the density of points \( A \), such that \( \rho_a = n_a / L_P \). From Equation (4), the observed network cross K-function, \( \hat{R}^{ab}(t) \) for the observed data set \( P \), can be written as:

\[
\hat{R}^{ab}(t) = \frac{[L_P]}{n_a \rho_a} \sum_{r=1}^{n_b} \left( \text{the number of points of } A \text{ on } L_P(t) \right).
\]

(5)

Statistical tests for bivariate analysis are similar to the univariate network K-function \( K(t) \), where the spatial scale at which the two groups 'attract' or 'repel' is determined (Bailey and Gatrell 1995), however the present version of SANET used for network cross K-function analyses does not construct a confidence envelope (Okabe et al. 2002b). The confidence interval of network K-function can be theoretically obtained from the binomial distribution, which is approximated by the normal distribution for a large number of points (Okabe et al. 2002b). To check for statistical significance of departures of \( \hat{R}^{ab}(t) \) from CSR, an approximation of the 95% confidence envelope was constructed by calculating the standard deviation (SD) of the normal distribution, and maximum and minimum values of \( \pm 1.65 * \text{SD using one-sided tests were accepted}. \) If \( \hat{R}^{ab}(t) > K^{ab}(t) \) and outside the confidence envelope, then points \( A \) and \( B \) are significantly 'attracted'. If \( \hat{R}^{ab}(t) < K^{ab}(t) \) and outside the confidence envelope, then points \( A \) and \( B \) are significantly 'repelled'.
Spatial point patterns of the three *Acacia* species were analysed on a road network shape-file using SANET Version 1.0 — 021125 (Okabe et al. 2002b), an ESRI ArcMap extension. The road network shape-file of the Lockhart Shire study area (area ~ 80 × 50 km; 1600 km of roads) was first pre-processed to ensure all polylines (links) were properly connected to each other. The SANET program is first used to calculate distances between all nodes on the road network, and is then used to assign points (in this case, populations of the three *Acacia* species) to the nearest point on the road network. Network K-function and cross K-function analyses were then performed using SANET, and output data were exported into MS Excel to aggregate data, calculate confidence intervals (for cross-K-function analyses), and produce graphs for visual analyses.

**Results**

Kernel estimations of the density of spatial point patterns for each *Acacia* species showed major clusters of *Acacia pycnantha* in central and southern areas of the Lockhart shire study area (Figure 3a), two strong clusters of *A. montana* in the south-east (Figure 3b), and two strong clusters of *A. decora* in the south-west and north-east study area (Figure 3c). The spatial patterns of *A. pycnantha* and *A. montana* cannot easily be distinguished, and appear to show similar point clusters in the southern area of the study area. However, spatial patterns of *A. decora* were markedly different to those of *A. montana*, and to a lesser extent *A. pycnantha*, with centroids of point clusters located in specific points in the north and south-west of the study area (Figure 3a-c).

Univariate spatial pattern analysis using the network K-function confirmed visual inspections using Kernel analyses, showing significant deviations from CSR for all three *Acacia* species. For *A. pycnantha*, plots of $\hat{K}(t)$ versus $t$ showed significant clustering of populations at all values of $t$, though plots were often only just outside the confidence envelope (Figure 4a). For *A. montana*, there was highly significant clustering of populations at all values of $t$, particularly at distances of 40–50 km (Figure 4b). In contrast, there was significant small-scale clustering of *A. decora* at distances up to 18 km, then significant large-scale repulsion of clusters of populations at distances of 40–65 km (Figure 4c). These scales of patterning in *A. decora* populations can easily be seen in the Kernel estimations (Figure 3c).

Observed correlations between Kernel distributions of *A. pycnantha* and *A. montana* (Figure 3) were confirmed with bivariate spatial pattern analysis using the network cross K-function. Visual examination of plots for the species pair, *A. pycnantha* — *A. montana* showed noticeable but non-significant attraction between the two distributions for all values of $t$, suggesting comparable spatial distributions (Figure 5a). For *A. pycnantha* — *A. decora*, plots of $\hat{K}_{xy}(t)$ versus $K_{xy}(t)$ were almost identical, which suggests that there was no spatial interaction between distributions of these two species (Figure 5b). For *A. montana* — *A. decora*, the distributions of these two species were somewhat but not significantly attracted at scales $t > 30$ km (Figure 5c).

**Discussion and conclusions**

The results of the Kernel and network K-function analyses showed that populations of *A. pycnantha* were tending to cluster, and *A. montana* was significantly clustered on roads in central and southern areas of the study region at most spatial scales. *A. pycnantha* and *A. montana* have similar life-history attributes, had similar distributions, and network cross K-function analysis showed that the spatial patterns of these two species were to some extent 'attracted' at all scales. These patterns are likely to reflect natural environmental factors, as both species were largely confined to higher topography in the south of the study area. This would suggest that at large scales at least, the spatial patterns of these two species are more strongly influenced by landscape variation in environmental than disturbance processes (Turner and Dale 1990).

In contrast, the spatial distribution of *A. decora* is peculiar, with significant clustering of populations in areas < 18 km wide in specific locations in the study area. The reasons for these patterns cannot be determined from spatial patterns alone, and further analyses are required to disentangle environmental and disturbance effects using patch, road and landscape scale environmental and disturbance variables. However, given that (1) centroids of clusters for *A. decora* appear to be located near human localities, (2) *A. decora* is a vigorous re-sprouting species, (3) the distribution patterns of *A. decora* do not conform to to-
Figure 3. Kernel range estimation of the spatial point density (shown as % confidence intervals) of (a) Acacia pycnantha, (b) A. montana and (c) A. decorra roadside populations.
Figure 4. Univariate spatial pattern analysis using the network K-function for distributions of three Acacia species in the Lockhart Shire study area, showing the observed statistic $K(t)$ (black lines) plotted against $t$. The grey lines give the 95% confidence envelope and values of $K(t)$ (central grey line) for CSR.

Figure 5. Bivariate spatial pattern analysis using the network K-function for distributions of pairs of three Acacia species in the Lockhart Shire study area, showing the observed statistic $K_{xy}(t)$ (black lines) plotted against $t$ (distance in km). The grey lines give the 95% confidence envelope and values of $K_{xy}(t)$ (central grey line) for CSR.

Topography or other environmental gradients, and (4) previous studies have shown how soil disturbance from roadworks promotes Acacia species, suggests that anthropogenic disturbances may have a stronger controlling influence than environmental determinants on population dynamics. Anthropogenic disturbances on road networks are usually highly selective and spatially arrayed in a logical way (de Blois et al. 2002). The effects of anthropogenic disturbances on plant populations depend on complex interactions between the spatial and temporal structure of the prevailing disturbance regime and individual species life history attributes (Clarke 1991; McIntyre et al. 1995; Moloney and Levin 1996). Further research is required to investigate the role of disturbance processes on shrub dynamics in roadside environments.

Spatial pattern analysis using the network K-function has shown to be useful in identifying different
patterns of mapped point data on a road network, and
the scale at which these patterns occur. The combina-
tion of graphical Kernel (for visual interpretation) and
network K-function analysis was especially informa-
tive in interpreting results. However results differed;
for example, spatial distributions of A. pycnantha
which appeared highly clustered in Kernel analysis,
were in fact only marginally significant using network
K-function analysis. Though useful for visual analysis
of spatial data, Kernel estimations of spatial point
intensity are based assumptions on a homogenous en-
vironment. Therefore like the ‘normal’ K-function,
Kernel estimations do not compensate for spatial dif-
f erences in road networks and their effect on point
patterns observed, e.g., the low density of roads in the
north-west of the Lockhart study area. As Figure 3,
Figure 4 demonstrate, it can be difficult for the hu-
man eye to distinguish a random pattern from a clus-
tered pattern on a network (Yamada and Thill 2002).
This point serves to illustrate the importance of using
correct tools for spatial analysis of point patterns on
networks such as the network K-function.

There were some technical difficulties experienced
with this method. Firstly, further development of bi-
variate analyses in SANET is required so that a con-
fidence envelope is automatically generated to assess
significance of departure from CSR. It is expected
that a future release of SANET will incorporate this
function. Secondly, SANET has quite lengthy compu-
tational times for some analyses (e.g., bivariate
analysis of 48 points on this road network using a
Pentium3 750 Mhz computer > 1 hour). Compu-
tational times will vary depending on available
computer hardware, the amount of points processed,
and the amount of nodes and links in the network GIS
shapefile. However with future improvements in soft-
ware and advancements in computer hardware tech-
nology, and considering that the ‘normal’ k-function
result in over detection of clustered patterns, the
network K-function can be regard as the most suit-
able method to analyse point data on road and other
networks (Yamada and Thill 2002).

Okabe and Yamada’s (2001) network K-function
using the SANET Arcmap extension provides a valu-
able method for conducting spatial point pattern
analysis on networks, and will greatly benefit spatial
analysis in road ecology and field boundary studies
(Forman 1998; Le Cœur et al. 2002). Possible eco-
logical applications include analysis of preferred
habitat of threatened fauna (including human inter-
relations), animal movement patterns from survey and
traffic mortality (road-kill) data, as well as many as-
pects of plant and community ecology along confined
corridors such as hedgerows and greenbelts. Applica-
tions may also extend to aquatic stream networks, a
field that has received little attention regarding spa-
tial statistics, and similarly frustrated by lack of suit-
able spatial analysis tools (Cooper et al. 1997). It is
 envisaged that the network K-function will become a
standard GIS application for use in ecological studies
on networks.

Acknowledgments

Financial support to Peter Spooner was provided by
postgraduate scholarships from Charles Sturt Univer-
sity and CSIRO Sustainable Ecosystems, Australia.
We would like to thank Kei-ichi Okunuki, Depart-
ment of Geography, Nagoya University for assistance
in this project; Ikuho Yamada, Department of Geog-
raphy, State University of New York at Buffalo for
advice and support in this project; and Simon
McDonald, CSU Spatial Data Analysis Network, for
GIS and spatial analysis assistance. We would also
like to thank Phillip Dixon and two other anonymous
referees whose comments have significantly im-
proved the quality of this manuscript.

References

Andrews A. 1990. Fragmentation of habitat by roads and utility
Bailey T.C. and Garrell A.C. 1995. Interactive Spatial Data Analy-
sis. Longman Scientific and Technical, Essex, UK.
review. In: Saunders D.A. and Hobbs R.J. (eds), Nature Con-
servation 2: The Role of Corridors, pp. 99-118. Surrey Beatty
and Sons, Chipping North, New South Wales, Australia.
Bull L. 1997. Lockhart Shire Roadside Vegetation Survey and
Recommendations. Lockhart Shire Council, Lockhart, New
South Wales, Australia.
Cale P. 1990. The value of road reserves to the avifauna of the
central wheat-belt of Western Australia. Proceedings of the Eco-
logical Society of Australia 16: 359-367.
Carr L.W., Fahrig L. and Pope S.E. 2002. Impacts of landscape
transformation by roads.. In: Gutzwiller K.J. (ed.), Applying
Landscape Ecology in Biological Conservation. Springer-Verlag,
Clarke J.S. 1991. Disturbance and tree life history on the shifting
Quantifying spatial heterogeneity in streams. Journal of the


