

Meng Gao

Detecting spatial aggregation from distance sampling: a probability distribution model of nearest neighbor distance

Received: 4 September 2012 / Accepted: 8 January 2013 / Published online: 16 February 2013
© The Ecological Society of Japan 2013

Abstract Spatial point pattern is an important tool for describing the spatial distribution of species in ecology. Negative binomial distribution (NBD) is widely used to model spatial aggregation. In this paper, we derive the probability distribution model of event-to-event nearest neighbor distance (distance from a focal individual to its n -th nearest individual). Compared with the probability distribution model of point-to-event nearest neighbor distance (distance from a randomly distributed sampling point to the n -th nearest individual), the new probability distribution model is more flexible. We propose that spatial aggregation can be detected by fitting this probability distribution model to event-to-event nearest neighbor distances. The performance is evaluated using both simulated and empirical spatial point patterns.

Keywords Spatial point pattern · Negative binomial · Distance sampling · Barro Colorado Island, Panama

Introduction

In ecology, the spatial point pattern, which is obtained by mapping the locations of each individual as points in space, is a very important tool for describing the spatial distribution of species (Legendre and Fortin 1989). Spatial analysis of point patterns is helpful in revealing the underlying ecological mechanisms behind the spatial distribution patterns (Condit et al. 2000; Stoyan and Penttinen 2000; John et al. 2007). There are three generally accepted types of spatial point patterns: regular, random, and aggregated (Pielou 1960). To detect spatial

patterns, quadrat sampling is one useful method in ecology. Quadrats are randomly thrown on the space and then the number of individuals gained in the each quadrat is counted. The quadrat count data can be well fitted by three discrete probability distribution models, generalized binomial distribution, Poisson distribution, and negative binomial distribution (NBD). They correspond to regular, random, and aggregated spatial point patterns, respectively (Bliss and Fisher 1953; Pielou 1960; Boswell and Patil 1970; He and Gaston 2000; Grevstad 2010; Zillio and He 2010). Poisson distribution usually serves as a null model of complete spatial randomness. Particularly, NBD are the most widely used models as aggregated populations have been found to be very common in nature (Pielou 1960, 1961; He and Gaston 2000).

The detection of spatial distribution was firstly implemented by computing kinds of indices based on quadrat count data, such as Green's dispersion index (Green 1966) and the Poisson index of dispersion (Cox and Lewis 1966), etc. However, spatial patterns inferred from quadrat sampling seriously depend on quadrat size and shape. Moreover, quadrat sampling also suffers from the problem of loss of information, such as spatial location of each quadrat and relative positions of individuals within quadrats. This information cannot be effectively reflected in the frequency distributions and various indices of aggregation (Hurlbert 1997).

Another class of indices is calculated based on distance sampling, such as the aggregation index of Clark and Evans (1954) and the index of segregation (Pielou 1960), etc. These indices either use distances from randomly selected points (sampling points) to the nearest individuals (also referred to as events or mapped points in references) or use distances from selected individuals to their nearest neighbors. We denote these two kinds of nearest neighbor distance (NND) as point-to-event NND and event-to-event NND in this study. The advantage of these indices is that they do not depend on the arbitrary choices of quadrat size or shape, especially the information of relative positions of individuals is

Electronic supplementary material The online version of this article (doi:10.1007/s11284-013-1029-x) contains supplementary material, which is available to authorized users.

M. Gao (✉)
Key Laboratory of Coastal Zone Environmental Processes,
Yantai Institute of Coastal Zone Research, Chinese
Academy of Sciences, Yantai, Shandong 264003, China
E-mail: gaomeng03@hotmail.com

incorporated. The disadvantage of these indices is that they only use the ensemble information of NND, sometimes the mean or the variance (Stoyan and Penttinen 2000). It is natural to consider the probability distribution function of NND of higher rank (n -th NND). Stoyan and Penttinen (2000) noted that the probability distribution function of NND can give a better parameter estimate than general second-order characteristics such as Ripley's K -function (Ripley 1976, 1977), L -function (Ripley 1988; Diggle 2003), and the pair correlation function (Stoyan and Stoyan 1996). The probability distribution models of n -th NND can be derived theoretically and are equivalent to the null models, Poisson model, and NBD model (Clark and Evans 1954; Thompson 1956; Eberhardt 1967). Compared with fitting Poisson or NBD models to quadrat count data, fitting the probability distribution models of NND to nearest neighbor distances has many advantages. First, distance sampling is easy to implement in the field. Second, the information of relative position of individuals in space can be reflected in NNDs, and the problem of information loss can be alleviated.

In this paper, we first present the probability distribution models of n -th NND, which are directly derived from the Poisson model and negative binomial model, respectively. Then we fit these probability distribution models to simulated point patterns and real point patterns of trees species in a tropical forest. The performance of fitting is then evaluated by a synthetic comparison and analysis.

Probability distribution model of n -th nearest neighbor distance

Poisson model vs. negative binomial model

The Poisson model describes the random spatial distribution in which the number of individuals in a region A follows the Poisson distribution (Thompson 1956; Pielou 1960):

$$p(N(A) = x) = \frac{(\lambda A)^x e^{-\lambda A}}{x!}, \quad x = 1, 2, \dots \quad (1)$$

where λ is the density of individuals and $N(A)$ is the number of individuals in region A . The negative binomial model is a very flexible probability distribution model and widely used to model spatial aggregation (Pielou 1960; Boswell and Patil 1970). The number of individuals in a region A follows the negative binomial distribution:

$$p(N(A) = x) = \frac{\Gamma(k+x)}{x! \Gamma(k)} \left(1 + \frac{k}{\lambda A}\right)^{-x} \left(1 + \frac{\lambda A}{k}\right)^{-k}, \quad (2)$$

$$x = 0, 1, 2, \dots$$

where k is a flexible aggregation parameter varying from 0 to $+\infty$ representing spatial patterns changing from extremely aggregated to extremely random. $\Gamma()$ represents the gamma function ($\Gamma(n) = (n-1)!$). Sometimes, $1/k$ is used as an index of aggregation. NBD distribution

becomes Poisson model when $k \rightarrow +\infty$. In practice, it is difficult to distinguish Poisson distribution from negative binomial distribution when k is larger than 10 (Zillio and He 2010). In modern statistics, the spatial point pattern can be generated by *stationary* and *isotropic* point process (Stoyan and Penttinen 2000; Diggle 2003). The Poisson distribution pattern can be easily simulated from the homogeneous Poisson point process. NBD can describe point patterns generated from the Neyman-Scott, Thomas, or Cox point process, although there is no stationary point process that directly generates the NBD (Diggle 2003; Zillio and He 2010).

Point-to-event nearest neighbor distance

The probability distribution model of point-to-event NND has been derived for both Poisson and negative binomial distribution cases (Clark and Evans 1954; Thompson 1956; Eberhardt 1967). For the Poisson case, the probability density function (pdf) of n -th NND (point-to-event) can be formulated as (Thompson 1956):

$$p_n(r) = 2\lambda\pi r e^{-\lambda\pi r^2} \frac{(\lambda\pi r^2)^{n-1}}{\Gamma(n)} \quad (3)$$

where r is the distance, $n (\geq 1)$ is the rank of NND, and λ is the density of individuals. For the negative binomial case, the pdf of n -th NND (point-to-event) can be formulated as (Eberhardt 1967):

$$g_n(r) = \frac{2(\lambda\pi)^n r^{2n-1}}{k^n} \frac{\Gamma(n+k)}{\Gamma(n)\Gamma(k)} \left(1 + \frac{\lambda\pi r^2}{k}\right)^{-n-k} \quad (4)$$

where k is the same as that in NBD model. As k tends to positive infinity, $g_n(r)$ tends to be $p_n(r)$. Starting from this probability distribution, Eberhardt (1967) proposed an index of non-randomness, which is referred to as Eberhardt's index, to detect the spatial aggregation. Prayag and Deshmukh (2000) derived the asymptotic distribution of Eberhardt's index theoretically.

Event-to-event nearest neighbor distance

The probability distribution model of n -th event-to-event NND, for either Poisson or negative binomial case, has not been presented before. Analogously, we also assume that the spatial distribution pattern satisfy *stationarity* and *isotropy*. The pdf of n -th NND (event-to-event) can be derived simply by applying Bayes' rule and finding the limit of a conditional probability. In addition, the derivation needs two virtual concentric circles. In Fig. 1, the two concentric circles of radii r and Δr are schematically illustrated. The details of the derivation are shown in the Appendix. For the Poisson case, the pdf of n -th NND (event-to-event) is exactly the expression in Eq. (3). For the negative binomial case, we have a new probability distribution model and the pdf of n -th NND (event-to-event) is

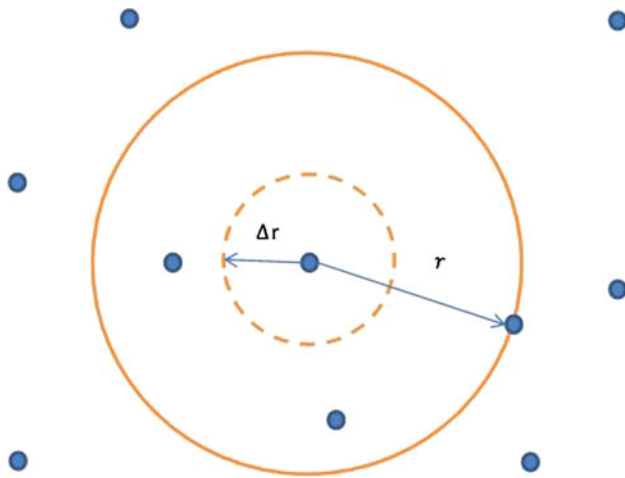


Fig. 1 Schematic illustration of the nearest neighbor distances from individual to individual. The *inner circle (dotted)* is just a virtual one used to derive $F_n(r)$

$$f_n(r) = \frac{2(\lambda\pi)^n r^{2n-1}}{k^n} \frac{\Gamma(n+k+1)}{\Gamma(n)\Gamma(k+1)} \left(1 + \frac{\lambda\pi r^2}{k}\right)^{-n-k-1} \quad (5)$$

It can be easily verified that the distribution $f_n(r)$ also becomes $p_n(r)$ when k tends to infinity. In supplementary materials, probability distribution models $p_n(r)$, $g_n(r)$ and $f_n(r)$ are schematically plotted for different values of aggregation parameter k and rank parameter n . These three probability distribution models differ significantly for a small k , but are indistinguishable when k becomes large. These relationships are also illustrated as figures there.

Parameter estimation

The aggregation parameter k can be estimated by using either the moment method or the likelihood method. From Eqs. (A7–A8) we can see that there is no restriction on the expectation of event-to-event nearest neighbor distance but the aggregation parameter k must be larger than 0.5 for the point-to-event nearest neighbor distance. As the estimate of parameter k cannot be obtained analytically, we have to resort to numerical solution. The alternative approach to estimate aggregation parameter k via likelihood method is also natural. Although the maximum-likelihood estimate of k is analytically intractable either, the approximations can be obtained numerically by maximizing the log-likelihood function.

Applications

In this study, our primary objective is to test the probability distribution model of n -th NND and estimate the aggregation parameter k from NND. The parameter λ is assumed to be already known in advance. The maximum-likelihood method is used to estimate the parameter k .

Simulation test

As widely recognized, a careful study of fitting the probability distribution model to simulated data is necessary before moving on to empirical data. In this section, we first test the three probability distribution models $p_n(r)$, $g_n(r)$, and $f_n(r)$ using simulated point patterns. The first point pattern is generated by the homogeneous Poisson point process, which produces a random point pattern. The second and third point patterns are aggregated and highly aggregated point patterns, which are both generated by Neyman-Scott point process. The difference between these two aggregated patterns is the intensity of parent process and the dispersion distance of offspring. The study area is a 100×100 -m square, and a 10-m-wide buffer zone is chosen to eliminate edge effect. That is to say, randomly selected sampling points are merely located in the inner 80×80 -m square, and only individuals within such a square are selected as focal individuals to obtain the event-to-event NND. For the purpose of comparison, the number of randomly selected sampling points equals to that of individuals in the inner square. Then we fit the three probability distribution models to samples of point-to-event and event-to-event NND. Figure 2 shows the simulated point patterns and the fitted probability density functions. For random point pattern, all three probability distribution models fit the nearest neighbor distances well. For the general aggregated point pattern, probability distribution models $g_n(r)$ and $f_n(r)$ that are derived from NBD can fit the nearest neighbor distances well. However, $p_n(r)$ derived from the Poisson model does not produce a satisfactory fitting. In Fig. 2, we merely show the results for $n = 3$. For other ranks of NND, such as $n = 1, 2, 4, 5$, the results are qualitatively similar. Based on the fitting results, we note that the probability distribution models $f_n(r)$ can fit spatial point patterns ranging from highly aggregated to complete random, if the two basic assumptions *stationarity* and *isotropy* of point processes are not significantly violated. For highly aggregated point pattern, $f_n(r)$ performs better than $g_n(r)$ in fitting nearest neighbor distances.

Empirical test

To further test the probability distribution models, we use a stand-mapping data set of tree species in a tropical rain forest in Barro Colorado Island (BCI), Panama. The study area is a $1,000 \times 500$ -m rectangular plot (50 ha). The spatial patterns of most tree species are aggregated (Condit et al. 2000; Zillio and He 2010). Analogously, we also set a buffer zone along the four edges, and the width is 25 m. For each species, sampling points are randomly distributed in the inner rectangular region, and the points' number equals to the individuals in such region.

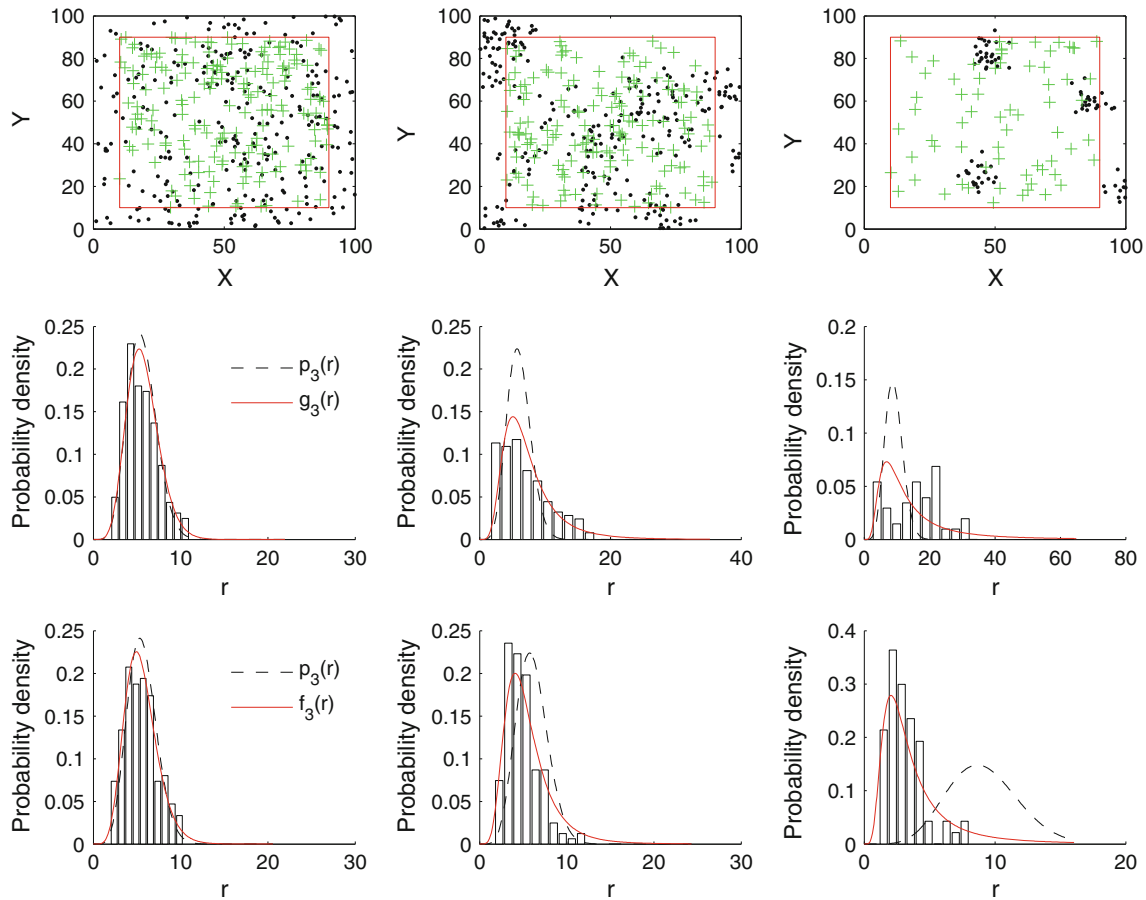


Fig. 2 (Top row) Simulated spatial point patterns (black points) and randomly distributed sampling points (plus). (Middle row) The frequency distribution of point-to-event NND (histogram) and the fitted probability density functions ($p_3(r)$ dashed lines; $g_3(r)$ solid lines). (Bottom row) The frequency distribution of event-to-event

NND (histogram) and the fitted probability density functions ($p_3(r)$ dashed lines; $f_3(r)$ solid lines). From left to right, each column corresponds to random, aggregated, and highly aggregated spatial point patterns

In this study, spatial distribution data of 183 species with individuals more than 50 in the inner rectangular region is used. The values of parameter n are chosen as $n = 1, 2, \dots, 10$. The goodness-of-fit is evaluated using χ^2 -test. The results of fitting are shown in Table 1. For Poisson case, it is found that fitting $p_n(r)$ to either point-to-event or event-to-event NND cannot pass the χ^2 -test for most species. Especially, there is no species that can pass the χ^2 -test when n is larger than 7 and 9, respectively. This result again verifies the argument that most tree species in BCI are aggregated distributed. Thus, we will not test the probability distribution model $p_n(r)$ anymore in the following sections.

For the negative binomial case, about two-thirds of the 183 tree species can pass the χ^2 -test when $n = 1$, but this number decreases as n increases. Figure 3 shows the well-fitted probability distribution models for three arbitrarily selected tree species in BCI. It is clear that both $g_n(r)$ and $f_n(r)$ can fit the nearest neighbor distances well. Species that pass the χ^2 -test of $g_n(r)$ and $f_n(r)$ have some common features: population is widely distributed and there are not too many small clusters. We further classified the remaining species into three categories:

Table 1 Results of χ^2 -test of fitting probability distribution models $p_n(r)$, $g_n(r)$, and $f_n(r)$ to point-to-event and event-to-event nearest neighbor distances for 183 Barro Colorado Island, Panama (BCI) species that have more than 50 individuals in the inner rectangular area (950×450 m)

n	Poisson case		Negative binomial case	
	Point-to-event (fails)	Event-to-event (fails)	Point-to-event (fails)	Event-to-event (fails)
1	161	163	57	63
2	175	177	65	84
3	178	178	73	96
4	179	180	91	103
5	182	180	93	99
6	182	181	101	102
7	183	181	102	99
8	183	181	106	99
9	183	183	103	99
10	183	183	105	104

The number in the table is the total number of species that have not passed the Chi-square test. The density is assumed to be known, and the aggregation parameter k is estimated using the maximum likelihood method. The number of bins for Chi-square is 10, and the degree of freedom is 8. The number of species that pass or fail the Chi square test may vary slightly when the number of bins is changed

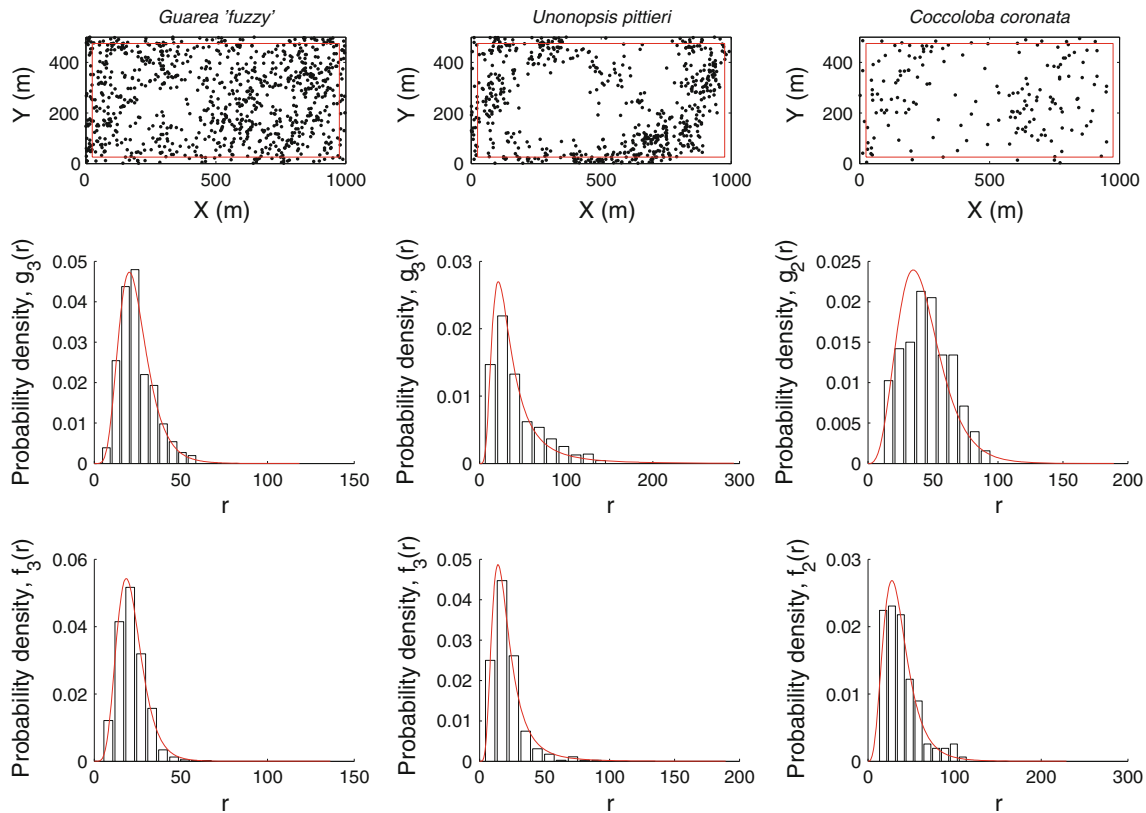


Fig. 3 (Top row) Spatial distributions of three Barro Colorado Island, Panama (BCI) tree species (*Guarea* “fuzzy”, *Unonopsis pittieri*, *Coccoloba coronata*). (Middle row) The corresponding frequency distribution of point-to-event NND (histogram) and the

fitted probability density functions (solid lines). (Bottom row) The corresponding frequency distribution of event-to-event NND (histogram) and the fitted probability density functions (solid lines)

species that pass the χ^2 -test of $g_n(r)$ but not $f_n(r)$, species that pass the χ^2 -test of $f_n(r)$ but not $g_n(r)$, and species that neither pass χ^2 -test of $g_n(r)$ nor that of $f_n(r)$. The spatial point pattern of these three types of species and the fitted probability distribution models are shown in detail in the supplemental materials. It is concluded that species that are highly aggregated but not widely distributed cannot pass the χ^2 -test of $g_n(r)$, and species with small clusters and widely distributed single individuals cannot pass the χ^2 -test of $f_n(r)$.

As we know, the major problem of quadrat sampling is the dependence on quadrat shape and size. Although the probability distribution models of n -th NND are thought to be equivalent to the original NBD model, it is deserved to compare the aggregation index estimated from quadrat sampling and distance sampling on different spatial scales. In this study, the comparison can be probably implemented as follows: (1) obtain the samples of n -th event-to-event NND ($n = 1, 2, \dots$); (2) compute the mean $M(r_n) = \sum r_n^{(i)} / N$; (3) take the mean as radius of circular quadrat and randomly throw equally sized quadrats on the study area; (4) fit NBD model to the quadrat count data and estimate the aggregation parameter k the using moment method; (5) fit the probability distribution model $f_n(r)$ to event-to-event NND and estimate parameter k using the maximum likelihood method. The reason why we use circular

quadrat is to link the sample scale of quadrat sampling (area) to sampling scale of distance sampling (distance). The results for four species that have passed the χ^2 -test of $g_n(r)$ and $f_n(r)$ are shown in Fig. 4. For parameter k , estimated from quadrat sampling, we find the value of k varies so much as the sample size increases. Sometimes, the detected pattern are changing from aggregated to random and then regular. However, the value of parameter k estimated from distance sampling is relatively stable. This result indicates that distance sampling is not seriously dependent on sampling scales.

Discussion

In this paper, we first showed two discrete probability distribution models, the Poisson model and the negative binomial model, of quadrat count data of spatial point patterns. The two models correspond to complete spatial random and aggregated spatial point patterns, respectively (Pielou 1960). Next, we presented two kinds of NND, point-to-event NND and event-to-event NND. For complete spatial random point pattern, these two kinds of NND have the same probability distribution model denoted as $p_n(r)$. For aggregated spatial point pattern, the probability distribution model of point-to-event NND has been given a long time ago

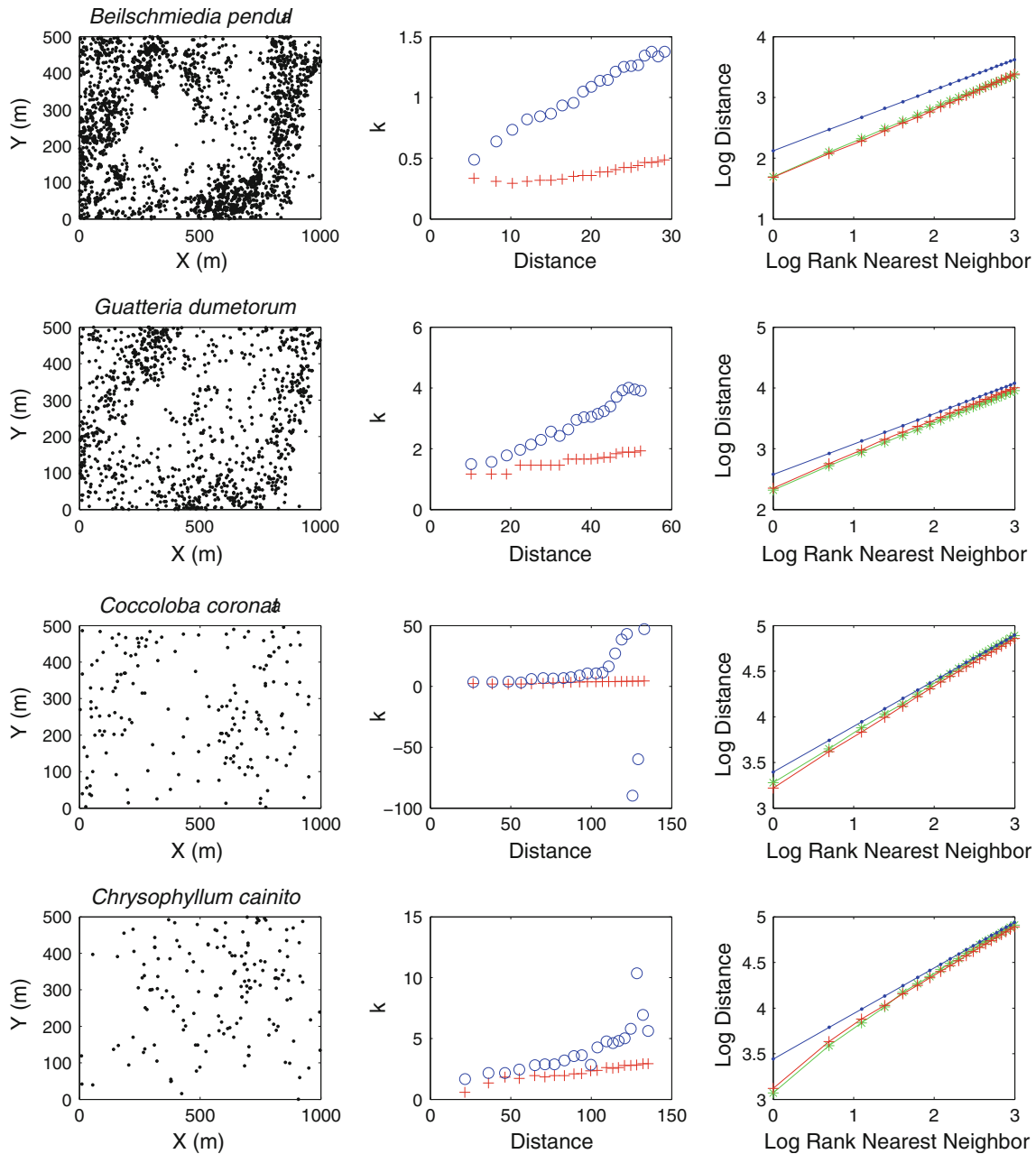


Fig. 4 (Left) Spatial distributions of three Barro Colorado Island, Panama (BCI) species (*Beilschmiedia pendula*, *Guatteria dumetorum*, *Cocoloba coronata*, *Chrysophyllum cainito*). (Middle) the change of aggregation parameter obtained by fitting quadrat count data to NBD (“o”) and fitting probability distribution model $f_n(r)$ to n -th

event-to-event NND (“+”). (Right) Log-log relationship between the mean or expectation of n -th event-to-event NND and the rank of nearest neighbor, symbols: asterisk the actual mean value $M(r_n)$; plus the expectation $E[r_n]$ for the negative binomial case (Eq. A8); dot the expectation $E[r_n]$ for the Poisson case (Eq. A6)

(Eberhardt 1967); however, the probability distribution model of event-to-event NND was firstly presented in this paper. We denoted these two models as $g_n(r)$ and $f_n(r)$. Although the mathematical formulae of $g_n(r)$ and $f_n(r)$ look similar, they are two different probability distribution models and differ remarkably when the shape parameter k is very small.

Theoretically, the two probability distribution models are both applicable to all kinds of spatial point patterns ranging from highly aggregated to complete random, if

the two basic assumptions *stationarity* and *isotropy* of point process are both satisfied. In simulation test, we found that $f_n(r)$ performed better than $g_n(r)$ for highly aggregated point patterns. This phenomenon can be explained by a simple example. We assume that there is a random point pattern in a 10×10 -m square. No doubt, both $f_n(r)$ and $g_n(r)$ can fit the NND well. Next, we extend the study area to a larger 20×20 -m square and assume there are no points in the new extended area. Then, the point pattern is considered to be aggregated in

the 20×20 -m square. The new density becomes one-quarter of the previous one, because density is inversely proportional to area. Accordingly, the point-to-event NND increase since the sampling points are always randomly distributed in the whole study area. In this process, the spatial aggregation cannot be realistically reflected in probability distribution model $g_n(r)$. Thus, detecting aggregation from point-to-event NND becomes unrealistic. However, all measures of event-to-event NND do not change in the process of area extension. The estimate of parameter k using probability distribution model $f_n(r)$ will become smaller indicating that aggregation can be detected. This is the superiority of $f_n(r)$ over $g_n(r)$. In an empirical test, we found that $g_n(r)$ and $f_n(r)$ were suitable for the majority of tree species in BCI. Meanwhile, there are still a few tree species in BCI that cannot pass the χ^2 -test. Our results are summarized as follows: $g_n(r)$ is not suitable for species that are highly aggregated but not widely distributed; $f_n(r)$ is not suitable for species with both small clusters and widely distributed single individuals. As we know, the derivation of both $g_n(r)$ and $f_n(r)$ have some common basic assumptions: individuals are distributed in infinitely large areas (inherit from the assumption of NBD model); the spatial distribution is invariant with respect to translations and rotations (stationary and isotropic). Obviously, these two basic assumptions cannot be strictly met in reality. Moreover, in the derivation of $g_n(r)$ and $f_n(r)$, the n individuals in the circular area are assumed to be randomly distributed for simplicity. This assumption is not realistic either. In addition, samples of higher rank of NND are not reliable because the edge effect cannot be fully eliminated by setting a narrow buffer zone. Therefore, fitting $g_n(r)$ and $f_n(r)$ to n -th NND both become worse when the order n is increased. Although the basic assumptions of continuous probability distribution models $g_n(r)$ and $f_n(r)$ are much stricter and obtaining a good fitting is not easy, they are both useful in modeling spatial point patterns. Especially, we have gotten a new flexible model $f_n(r)$.

Spatial pattern of species is the result of different ecological processes and forces on different spatial scales (Wiegand and Moloney 2004). Over the last few decades, a number of methods have been developed to quantify the aggregation of spatial point patterns (Diggle 2003; Wiegand and Moloney 2004). These methods use either quadrat sampling or distance sampling. In earlier methods, the spatial characteristic was usually summarized as a simple index, and spatial scales are not considered (Diggle 2003). However, it is not reasonable to set spatial scales aside when we detect spatial aggregations. The relationship between spatial pattern and scales are not properly considered until the second-order statistics were proposed, such as Ripley's K -function and L -function etc. (Ripley 1976, 1977, 1988). Over the last few years, methods based on Ripley's K -function have undergone a rapid development and are now being widely used, especially in plant ecology (Wiegand and Moloney 2004 and references therein). Like these

second-order statistics, the probability distribution models $g_n(r)$ and $f_n(r)$ can also capture the multiscale spatial characteristic by showing the relationship k vs. r_n . Now we already have many methods to detect spatial aggregation; however, to our best knowledge, there is no a generally acceptable criterion to determine the best one in practice. The best choice depends on the biological question asked and the quality of field data. For second-order statistics, such as the commonly used Ripley's K -function or the pair-correlation function, the null model is complete spatial randomness (Ripley 1976, 1977; Bailey and Gatrell 1995; Diggle 2003). The second-order statistics are usually used to answer the questions of whether a pattern is random, aggregated, or regular on some specific spatial scales (Wiegand and Moloney 2004). To compute the second-order statistics, we have to know the positions of all individuals in the study area. In other words, the study area must be fully surveyed. For $g_n(r)$ and $f_n(r)$, the null model is the ideal aggregated pattern described by NBD model. They can be used to answer the question of how aggregated the spatial pattern is on multiple spatial scales. In practice, there is no need to use all samples of n -th NND to estimate the parameter k . It means that a partly surveyed study area is enough. Similarly, we cannot rashly judge which is the better one between $g_n(r)$ and $f_n(r)$ in practical applications, although $f_n(r)$ is much flexible than $g_n(r)$ theoretically. When $g_n(r)$ is used, the sampling points to obtain samples of point-to-event NND must be randomly distributed. In some previous studies, this requirement was basically met through field ecologists' expertise, so $g_n(r)$ was applicable in detecting spatial aggregation and even estimating density (Magnussen et al. 2008). Our empirical test also verified this point. In this paper, the new probability model was only tested using one accessible data set of tree species. We note that more empirical tests using other data sets are also necessary in future work.

Acknowledgments The author is grateful to the Center for Tropical Forest Science for providing the BCI data. Help from Prof. Fangliang in reading the original manuscript and Dr. G.C. Shen in preparing the data are also acknowledged. The authors also wish to thank the editors and two anonymous reviewers for their useful comments and suggestions related to this manuscript. This work was supported by NNSF of China (No.31000197), as well as Knowledge Innovation Project of CAS (No. KZCX2-EW-QN209).

Appendix

To derive the probability distribution model of event-to-event NND, we also assume that the spatial distribution pattern satisfies *stationarity* and *isotropy*. Then the pdf of n -th NND (event-to-event) can be derived simply by applying Bayes' rule and finding the limit of a conditional probability. Here, we only take the NBD case as an example to illustrate the derivation. In Fig. 1, there are two concentric circles of radii r and Δr , and the center is a randomly selected individual. First, the

number of individuals in a circle of radius r follows a negative binomial distribution with mean $\lambda\pi r^2$. Define a random variable r_n as the n -th NND (event-to-event), then probability that r_n is smaller than r is given by $P(r_n < r)$, which is also the probability distribution function of r_n (denoted by $F_n(r)$). Now, it is in position to derive the analytical expression of $P(r_n < r)$. We have

$$\begin{aligned} F_n(r) &= P(r_n < r) \\ &= P(\text{more than } n + 1 \text{ individuals in the circle, one at the center}) \\ &= \frac{P(\text{more than } n + 1 \text{ individuals in the circle, one at the center})}{P(\text{one at the center})} \\ &= \lim_{\Delta r \rightarrow 0} \frac{P(n \text{ individuals in the annulus, one in the inner circle})}{P(\text{one in the inner circle})} \end{aligned} \quad (\text{A1})$$

The numerator in Eq. (A1) can be approximated by

$$\begin{aligned} &\sum_{s=n+1}^{\infty} \binom{s}{1} q(1-q)^{s-1} \\ &\times \frac{\Gamma(k+s)}{s!\Gamma(k)} \left(1 + \frac{k}{\lambda\pi r^2}\right)^{-s} \left(1 + \frac{\lambda\pi r^2}{k}\right)^{-k} \end{aligned} \quad (\text{A2})$$

where $q = \Delta r^2/r^2$. The formula after the summation in Eq. (A2) represents the probability that only one individual is located at the inner circle while other $s - 1$ individuals are at the annulus, given there are s individuals falling in the circle. The basic assumption here is that individuals are randomly distributed in the outer circle and follow a binomial distribution. The denominator is relatively simple, which is just the probability of an individual falling in the inner circle,

$$p(x=1) = \lambda\pi\Delta r^2 \left(1 + \frac{\lambda\pi\Delta r^2}{k}\right)^{-k-1} \quad (\text{A3})$$

Substituting Eqs. (A2–A3) into Eq. (A1) and evaluating the limit, we can obtain the probability distribution function:

$$\begin{aligned} F_n(r) &= \frac{2\Gamma(n+k+1)}{rk\Gamma(k+1)\Gamma(n+2)} \left(1 + \frac{k}{\lambda\pi r^2}\right)^{-n} \left(1 + \frac{\lambda\pi r^2}{k}\right)^{-k-1} \\ &\times \left[\lambda\pi r^2 k(k+n+1) F_1^2\left(2, 2+k+n, 2+n, \frac{\lambda\pi r^2}{k+\lambda\pi r^2}\right) \right. \\ &\quad \left. - (n+1)(k+\lambda\pi r^2)(\lambda\pi r^2(k+1) - nk) \right. \\ &\quad \left. \times F_1^2\left(1, 1+k+n, 1+n, \frac{\lambda\pi r^2}{k+\lambda\pi r^2}\right) \right] \end{aligned} \quad (\text{A4})$$

where F_1^2 is Gauss hypergeometric function. The expression of $F_n(r)$ is very complex, but its differential about distance r is relatively simple, which is exactly the

probability density function of the n -th NND (event-to-event),

$$f_n(r) = \frac{2(\lambda\pi)^n r^{2n-1}}{k^n} \frac{\Gamma(n+k+1)}{\Gamma(n)\Gamma(k+1)} \left(1 + \frac{\lambda\pi r^2}{k}\right)^{-n-k-1} \quad (\text{A5})$$

It can be easily verified that Eq. (A5) meets all the conditions of general probability density functions. Following the same procedure, the probability density function of n -th NND (event-to-event) for the Poisson case can also be derived. The result is exactly $p_n(r)$ shown in Eq. (3).

With the analytical expressions of $p_n(r)$, $g_n(r)$, and $f_n(r)$, the moments can be simply derived. For Poisson case, Thompson (1956) presented that the expectation of n -th NND, either point-to-event or event-to-event, was

$$E[r_n] = \frac{1}{\lambda} \frac{n(2n)!}{(2^n n!)} \approx \frac{1}{\sqrt{\lambda\pi}} n^{0.5} \quad (\text{A6})$$

The approximation is based on Sterling's formula, which holds very well even for small n . For negative binomial case, the expectations of point-to-event NND and event-to-event NND are

$$\begin{aligned} \text{Point-to-event: } E[r_n] &= \frac{1}{\sqrt{\lambda\pi}} \frac{\Gamma(n+0.5)}{\Gamma(n)} \frac{\Gamma(k-0.5)\sqrt{k}}{\Gamma(k)} \\ &\quad (k > 0.5), \end{aligned} \quad (\text{A7})$$

$$\begin{aligned} \text{Event-to-event: } E[r_n] &= \frac{1}{\sqrt{\lambda\pi}} \frac{\Gamma(n+0.5)}{\Gamma(n)} \frac{\Gamma(k+0.5)\sqrt{k}}{\Gamma(k+1)} \end{aligned} \quad (\text{A8})$$

Other higher-order of moments can also be obtained, and we do not show all of them here.

References

- Bailey TC, Gatrell AC (1995) Interactive spatial data analysis. Longman Scientific and Technical, Harlow, UK
Bliss CI, Fisher RA (1953) Fitting the negative binomial distribution to biological data. *Biometrics* 9:176–200

- Boswell MT, Patil GP (1970) Chance mechanisms generating the negative binomial distribution. In: Patil GP (ed) *Random counts in scientific work*. Pennsylvania State, University Press, University Park
- Clark PJ, Evans FC (1954) Distance to nearest neighbor as a measure of spatial relationships in populations. *Ecology* 35:445–453
- Condit R et al (2000) Spatial patterns in the distribution of tropical tree species. *Science* 288:1414–1418
- Cox DR, Lewis PAW (1966) *The statistical analysis of series of events*. Methuen, London, UK
- Diggle PJ (2003) *Statistical analysis of spatial point patterns*. Academic Press, London
- Eberhardt LL (1967) Some developments in distance sampling. *Biometrics* 23:207–216
- Green RH (1966) Measurement of non-randomness in spatial distributions. *Res Popul Ecol* 8:1–7
- Grevstad N (2010) Mapping spatial aggregation from counts data: a penalized likelihood approach. *Environmetrics* 21:834–848
- He F, Gaston KJ (2000) Estimating species abundance from occurrence. *Am Nat* 156:553–559
- Hurlbert SH (1997) Spatial distribution of the montane unicorn. *Oikos* 58:257–271
- John R et al (2007) Soil nutrients influence spatial distributions of tropical tree species. *Proc Natl Acad Sci* 104:864–869
- Legendre P, Fortin MJ (1989) Spatial pattern and ecological analysis. *Vegetatio* 80:107–138
- Magnussen S, Kleinn C, Picard N (2008) Two new density estimator for distance sampling. *Eur J For Res* 127:213–224
- Pielou EC (1960) A single mechanism to account for regular, random and aggregated populations. *J Ecol* 48:575–584
- Pielou EC (1961) Segregation and symmetry in two-species population as studied by nearest-neighbor relationships. *J Ecol* 49:255–269
- Prayag VR, Deshmukh SR (2000) Testing randomness of spatial pattern using Eberhardt's index. *Environmetrics* 11:571–582
- Ripley BD (1976) The second-order analysis of stationary point processes. *J Appl Probab* 13:255–266
- Ripley BD (1977) Modelling spatial patterns (with discussion). *J R Statist Soc Ser B* 39:172–212
- Ripley BD (1988) *Statistical inference for spatial processes*. Cambridge University Press, Cambridge, UK
- Stoyan D, Penttinen A (2000) Recent applications of point process methods in forestry statistics. *Stat Sci* 15:61–78
- Stoyan D, Stoyan H (1996) Estimating pair correlation functions of planar cluster processes. *Biom J* 38:259–271
- Thompson HR (1956) Distribution of distance to n-th nearest neighbor in a population of randomly distributed individuals. *Ecology* 37:391–394
- Wiegand T, Moloney KA (2004) Rings, circles, and null-models for point pattern analysis in ecology. *Oikos* 104:209–229
- Zillio T, He F (2010) Modeling spatial aggregation of finite populations. *Ecology* 91(12):3698–3706