Article

Particle swarm optimization algorithm for parameter estimation in Gamma-Poisson distribution model of *k*-tree distance

Feixia Lu[1], **Dingyuan Mo**[2,3], **Meng Gao**[2]

[1]School of Mathematics and Information Science, Yantai University, Yantai, 264005, China

[2]Key Laboratory of Coastal Environmental Processes and Ecological Remediation, Yantai Institute of Coastal Zone Research, Chinese Academy of Sciences, Yantai, 264003, China

[3]Univerisity of Chinese Academy of Sciences, Beijing, 100049, China

E-mail: mgao@yic.ac.cn

Received 25 July 2015; Accepted 3 August 2015; *Published online 1 December 2015*

Abstract

Distance sampling is a flexible and efficient inventory technique in forestry and ecology, especially in highly dense plant communities, and in difficult terrain. Point-to-tree distance or tree-to-tree distance was used to estimate characteristics of the spatial point pattern mapped from particular spatial locations of plant or tree individuals. For random spatial point patterns, there is an ideal probability distribution model of point-to-tree distance resulting in unbiased density estimators. For aggregated spatial point patterns, Gamma-Poisson probability model of point-to-tree distance corresponding to Gamma-Poisson point process is one candidate model. Although the density estimator based on Gamma-Poisson model is biased, it performs satisfactorily in practical applications. However, numerical method to compute the maximum likelihood estimates of Gamma-Poisson model is very complicated. In this paper, a parameter optimization method, particle swarm optimization algorithm, is applied for parameter estimation in Gamma-Poisson model. The results showed that the new parameter estimation method was efficient and not constrained by the sample size; therefore, the computational complexity was significantly reduced. We suggest this parameter optimization for density estimation in forestry and ecology.

Keywords spatial point pattern; distance sampling; point to tree distance; density estimator.

Computational Ecology and Software
ISSN 2220-721X
URL: http://www.iaees.org/publications/journals/ces/online-version.asp
RSS: http://www.iaees.org/publications/journals/ces/rss.xml
E-mail: ces@iaees.org
Editor-in-Chief: WenJun Zhang
Publisher: International Academy of Ecology and Environmental Sciences

1 Introduction

Abundance or density of a focal species is an important indicator of the species' viability and conservation status; as such the ability to estimate abundance/density from point patterns or quadrat/grid samples of occurrence has been an area of endeavor in conservation ecology (Cassey, 1999; Schwarz and Seber, 1999;

Hui and McGeoch, 2007, 2014; Hui et al., 2009, 2011). Distance sampling, also known as plotless sampling, is more commonly applied in practice when plot (quadrat) sampling would be difficult or too costly (Sheil et al., 2003; Picard et al., 2005). Point-to-tree distance sampling (sometimes also referred to as *k*-tree sampling) that use distance from a sample point to the *k* nearest objects (trees or plants) is one simple distance sampling approach (Pielou, 1977; Sutherland, 1998). For moderate *k* values, point-to-tree distance sampling often facilitates field work considerably (Kleinn and Vilćko, 2006). With samples of point-to-tree distance, the density of trees (individuals per unit area) can be derived according to a density estimator. Generally, there are three kinds of density estimator using point-to-tree distances: empirical, model-based, and composite (Kleinn and Vilćko, 2006; Magnussen et al., 2008b; Magnussen, 2012; Magnussen et al., 2012). It is found that the performance of density estimators often depend on the spatial pattern of trees, the choice of *k*, and the sample size of point-to-tree distance *n*. In this paper, we focus our attention on model-based density estimator.

For a majority of density estimators, it is explicitly or implicitly assumed that the spatial distribution of trees follow a Poisson model of complete spatial randomness. Under Poisson model, an asymptotically unbiased maximum likelihood estimator of density was derived (Moore, 1954; Eberhardt, 1967). In natural and semi-natural populations the distribution of trees is often irregular with a variation in local density that generates an appearance of spatial clustering (Hui et al., 2010). Magnussen et al. (2008a, 2008b) proposed a density estimator based on a Gamma-Poisson distribution model of point-to-tree distance corresponding to clustered spatial point patterns. An extensive evaluation of 19 *k*-tree distance estimators revealed that the Gamma-Poisson density estimator ranked second when the sample size was greater than 10 (Magnussen et al. 2008b). Because of its good performance, Gamma-Poisson density estimator was always included in composite density estimator (Magnussen, 2012; Magnussen et al., 2012). In Gamma-Poisson model of point-to-tree distance, estimating the two model parameters is not trivial task because of the complicated form of the probability density function. In statistics, there are many numerical methods for estimating the parameters of complicated probability distribution model such as moment method, empirical method, graphical method, and maximum likelihood method. Maximum likelihood method was applied in (Magnussen et al. 2008b) for parameter estimation and the optimized estimates on the log-likelihood surface were searched by Nelder-Mead algorithm. However, maximum likelihood estimation was fraught with nontrivial numerical issues when the samples of point-to-tree distance were rare (Magnussen et al., 2008b). The aim of this paper is to apply a numerical optimization method, the particle swarm algorithm, to calculate the model parameters of Gamma-Poisson distribution model of point-to-tree distance.

2 Materials and Methods

2.1 Gamma-Poissson model

Spatial point pattern, which is obtained by mapping the locations of each individual as points in space, is widely used for forestry statistics (Legendre and Fortin, 1989; Stoyan and Penttinen, 2000). A spatial point pattern is considered as a realisation of a spatial point process (Gatrell et al., 1996) that is a “stochastic mechanism generating countable points in the plane” (Diggle, 2003). For a spatial point process, a fundamental property is the intensity, $\lambda(s)$, which the expected number of points per unit area at the location *s*. Homogeneous Poisson process with a constant intensity λ is the simplest spatial process that is also termed as complete spatial randomness (CSR). It has two important properties (Stoyan and Penttinen, 2000). Under CSR, the number of points in any region follows the Poisson distribution with a mean of λA, where *A* is the sampling area. Then, the probability density function (pdf) of the distance r_k from a (random) sample location to the *k* nearest point is

$$g(r_k \mid \lambda) = 2(\lambda\pi)^k r_k^{2k-1} \frac{\exp(-\lambda\pi r_k^2)}{(k-1)!} \quad (1)$$

However, empirical data shows that sampling with plots of fixed size (*A*) often gives rise to a distribution of points per plot (*A*) that is closer to a negative binomial distribution (Eberhardt, 1967; White and Bennetts, 1996; Turechek and Madden, 1999; Brix and Cahdoeuf, 2002) than to a Poisson distribution. A negative binomial distribution of counts implies that the local density $\lambda(s)$ of points varies from one sample location to another as in a gamma distribution with a mean of λ (Johnson et al., 1992). Then the probability density function (pdf) of the distance r_k can be obtained from a convolution of the pdf of r_k under a Poisson model (Eq.1) with a gamma distribution

$$g(r_k \mid \alpha, \beta) = \int_0^\infty g(r_k \mid \lambda) h(\lambda \mid \alpha, \beta) d\lambda \quad (2)$$

where $h(\lambda \mid \alpha, \beta)$ is a gamma function and α and β are the two parameters. The intergal generates a new probability density funtion

$$g(r_k \mid \alpha, \beta) = 2\pi^k r_k^{2k-1} \frac{\Gamma(k+\alpha)}{\Gamma(k)\Gamma(\alpha)} \beta^{-\alpha} (\beta^{-1} + \pi r_k^2)^{-k-\alpha} \quad (3)$$

with $\lambda = \alpha \times \beta$. Eq.(3) can also be derived from negative binomial distribution directly (Thompson, 1956). Generally, parametes α and β can be estimated using either moment method (sample size ≤ 30) or maximum likelihood method (sample size > 30) (Magnussen et al., 2008a, 2008b; Gao, 2013).

2.2 Particle swarm parameter optimization

Particle swarm optimization is population based optimization tool, where the system is initialized with a population of random particles and optimizes a problem by iteratively trying to improve a candidate solution according to an objective function. Suppose that the search space is *N*-dimensional, the *i*-th particle is represented by an *N*-dimensional vector $X_i = (x_{i1}, x_{i2}, \text{L}, x_{iN})$ and its velocity is represented as $V_i = (v_{i1}, v_{i2}, \text{L}, v_{iN})$. With a given objective function, the fitness of each particle can be evaluated. Record the fitness of the personal best position $P_i = (p_{i1}, p_{i2}, \text{L}, p_{iN})$ which is the best previously visited position of the *i*-th particle as f_{pbest}. Meanwhile, the position of the best individual of the whole swarm is marked as the global best position $P_g = (p_{g1}, p_{g2}, \text{L}, p_{gN})$, and the objective function corresponding to P_g is f_{gbest}. At each step, the velocity and position of the *i*-th particle were updated according to rules defined by the following equations until the stopping limit was reached:

$$V_i = \omega * V_i + c_1 * r_1 * (P_i - X_i) + c_i * r_2 * (G - X_i) \quad (4)$$

$$X_i = X_i + V_i \quad (5)$$

where ω is the inertia weight, by which the impact of previous velocity of particle on its current one can be controlled, r_1 and r_2 are independently uniformly distributed random variables within (0,1) and c_1 and c_2 are positive constant parameters called acceleration coefficients which control the maximum step size. Fig.1 shows the computation flowchart of particle swarm algorithm.

For the probability density function (Eq.3), we can compute the expectation and variance of point-to-tree distance r_k (Magnussen et al., 2008a):

$$\bar{r}_k = E[r_k \mid \alpha > 0.5, \beta] = \frac{\Gamma(\alpha - 0.5)\Gamma(k + 0.5)}{\sqrt{\pi\beta}\Gamma(\alpha)\Gamma(k)} \quad (6)$$

$$\sigma^2 = Var[r_k \mid \alpha > 1, \beta] = \frac{k}{\pi\beta(\alpha - 1)} \quad (7)$$

From Eq. (6-7), we found that parameter β can be cancelled out in $\sigma^2 / \bar{r}_k^2$. Then the residual value ε defined below is used as the objective function

$$\varepsilon = \frac{\sigma^2}{\bar{r}_k^2} - \frac{k\Gamma(k)^2\Gamma(\alpha)^2}{(\alpha-1)\Gamma(\alpha-0.5)^2\Gamma(k+0.5)^2} + 1 \quad (8)$$

With this objective function (Eq. 8), the parameter α can be estimated using particle swarm optimization algorithm, and then parameter β can be further computed using Eq. (6). In this paper, we set $\omega = 0.5$, $c_1 = c_2 = 0.5$, and the population size $N = 100$. It must to be noted that the value of parameter α was larger than one, but this restriction on α was not important in practical applications (Magnussen et al., 2008a).

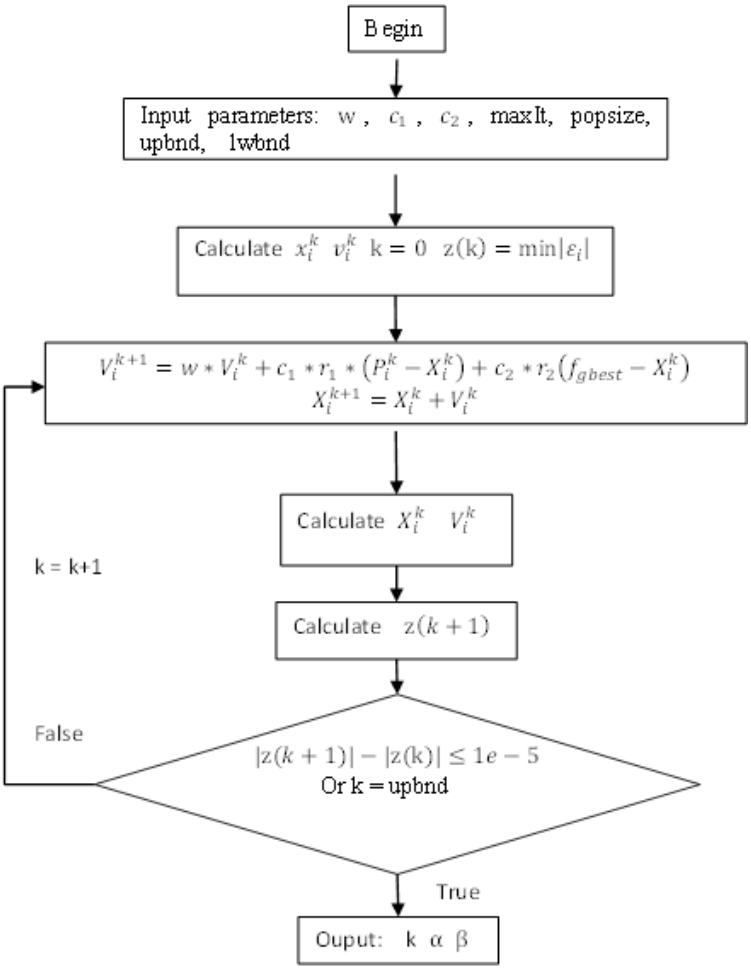

Fig. 1 Computation flowchart of Particle Swarm Algorithm in this paper.

2.3 Simulated and empircal point patterns

We test the new parameter optimization method using both simulated and empirical data. We generated one random spatial pattern and three spatial patterns with the degree of clustering varying from low to to high levels of aggregation using a computational statistical toolbox under MATLAB Environment (Martinez and Martinez, 2002). Random point pattern was generated by a homogeneous Poisson process with intensity λ=0.001, and all points were distributed in a 100 ha (1000×1000 m) area. The aggregate spatial point patterns

were generated by Poisson Cluster process, where the parents were generated by a homogeneous Poisson and the children were distributed around the parents according to a bivariate normal distribution. There were three parameters that defined the number of parents (μ), the expected number of children in each cluster (v), and the mean distance from parent to children (φ) were used. In Fig. 2, we illustrated four spatial point patterns of random, low, moderate, and high levels of aggregation. To eliminate edge effect, a 50m wide buffer zone was chosen. That is to say, randomly selected sampling points are merely located in the inner 900×900 m square, and the number of sampling points equals to that of points in the inner square. 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 1000 ×500m rectangular plot (50 ha). The spatial patterns of most tree species are aggregate (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 also randomly distributed in the inner rectangular region.

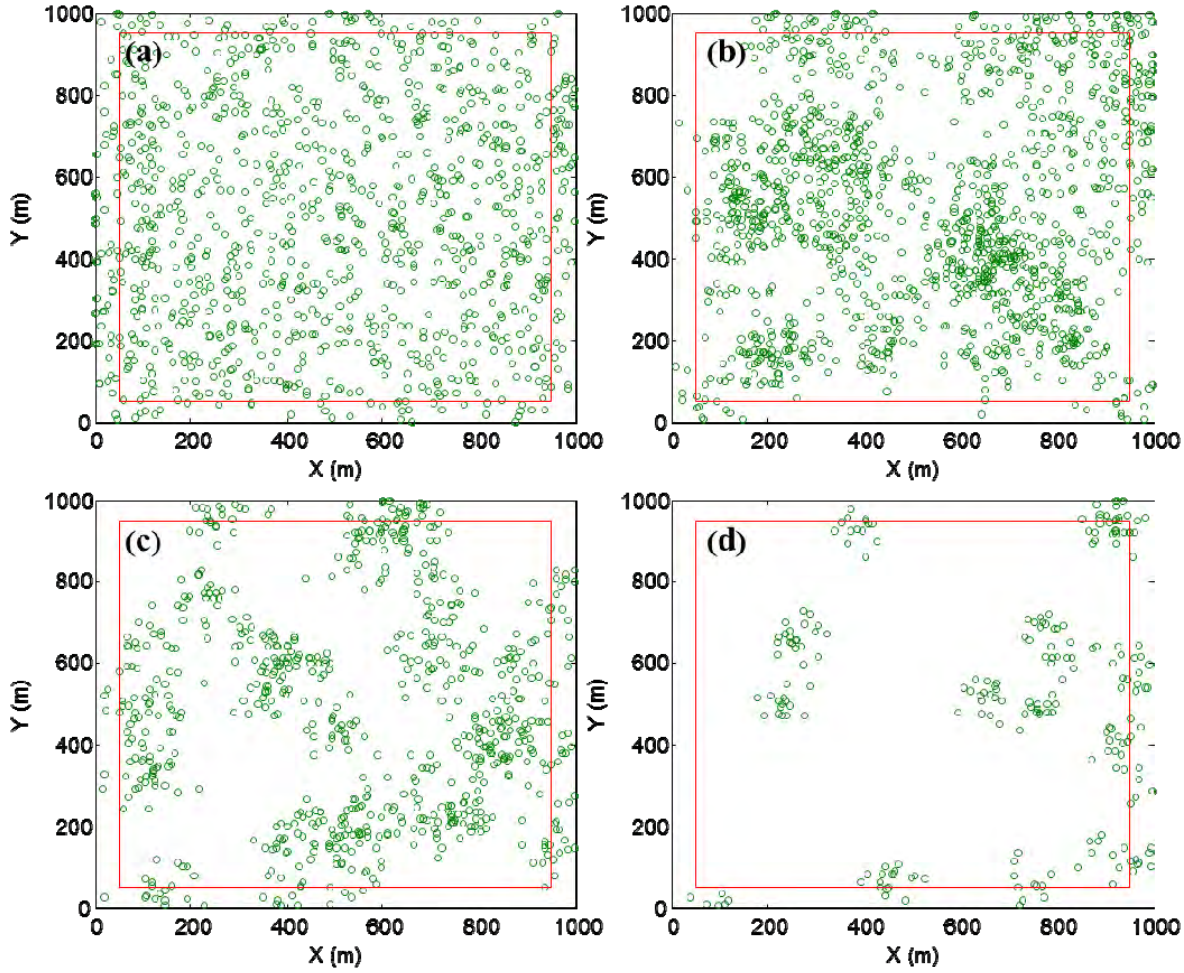

Fig. 2 Four simulated spatial point patterns (green circles). (a) random pattern generated by a homogeneous Poisson process with intensity λ=0.001; (b-d) lowly, moderately and highly aggregate spatial point patterns generated by Poisson Cluster process with parameters with parameters (μ, v, φ) equals to (50,30,50), (40,20,40), and (20,15,30), respectively. The red lines show the boundaries of the buffer zone.

The performance of particle swarm parameter optimization algorithm was evaluated according to two criteria: Chi-square error of model fitting and relative error (RE) of density estimate. Chi-square error (CSE) of model fitting, is a goodness-of-fit statistic given by

$$CSE = \sum_{i=1}^{n} \frac{(p_{obs} - p_{fit})^2}{p_{fit}} \quad (9)$$

where p_{obs} and p_{fit} are the observed and fitted probability density function, n is the number of bins of observed probability density function. A smaller CSE means a better fitting of model (3) to samples of point-to-tree distance, and a smaller RE means a more accurate density estimate. The ranks of point-to-tree distance are chosen from 1 to 6. A larger k will be less attractive for practical application.

3 Results

The results for the simulated spatial point patterns were shown in Fig. 3 and Table 1 and 2. In Fig. 3, we showed the fitting of point-to-tree distance to model (3). For point patterns with random structure or low level of aggregation, the probability distribution models fit the point-to-tree distance well; while for point patterns with moderate to high level of aggregation, the fitting becomes worse (Fig. 3). Table 1 and 2 summarized the results of 300 independent simulations of point patterns and fitting of probability distribution functions. For random point patterns, the fitting for different k were almost identical. But for all aggregate spatial patterns, an increased CSE indicated worse fitting. When estimates of the two parameters α and β were multiplied to calculate the density, we got more interesting results. For random spatial point patterns, the density was usually over-estimated, but for aggregate spatial patterns the density was usually under-estimated. Moreover, the estimator became less biased as the increasing of k (Table 2), although the fitting of probability density function become worse (Table 1). The inconsistency between model fitting and density estimation has been explained in Magnussen et al. (2008a) because density was not directly optimized. In reality, density or abundance is unknown, that's why we can't use density as the target for optimization.

Table 1 Chi-square error (CSE) of model fitting. Entries are averages of 300 replicated independent simulations of spatial point patterns and model fitting. Standard deviations of CSE are in parentheses.

	k=1	k=2	k=3	k=4	k=5	k=6
Random	0.0129	0.0123	0.0131	0.0146	0.0149	0.0145
	(0.0058)	(0.0045)	(0.0046)	(0.0065)	(0.0076)	(0.006)
Low level of Aggregation	0.0149	0.0195	0.022	0.0234	0.0243	0.0244
	(0.0073)	(0.0099)	(0.0121)	(0.013)	(0.0152)	(0.0156)
Moderat level of aggregation	0.0293	0.0422	0.0497	0.0553	0.0578	0.0598
	(0.0112)	(0.0181)	(0.0264)	(0.0325)	(0.0365)	(0.0426)
High level of aggregation	0.0352	0.0782	0.1314	0.1971	0.247	0.2956
	(0.0142)	(0.0367)	(0.0941)	(0.2241)	(0.3053)	(0.421)

Table 2 Relative error (RE) of density estimate. Entries are averages of 300 replicated independent simulations of spatial point patterns and model fitting. Standard deviations of RE are in parentheses.

	k=1	*k*=2	*k*=3	*k*=4	*k*=5	*k*=6
Random	0.0282	0.0194	0.0198	0.0195	0.0195	0.0187
	(0.0523)	(0.0314)	(0.0239)	(0.0213)	(0.0208)	(0.0191)
Low level of Aggregation	-0.2041	-0.1601	-0.1307	-0.1093	-0.0932	-0.08
	(0.1165)	(0.0918)	(0.0783)	(0.0709)	(0.0648)	(0.0606)
Moderat level of aggregation	-0.4688	-0.3429	-0.2647	-0.2083	-0.1639	-0.1286
	(0.1028)	(0.0939)	(0.0874)	(0.0834)	(0.0798)	(0.0771)
High level of aggregation	-0.742	-0.5586	-0.4227	-0.3149	-0.2258	-0.152
	(0.0675)	(0.0908)	(0.1029)	(0.1115)	(0.1178)	(0.1229)

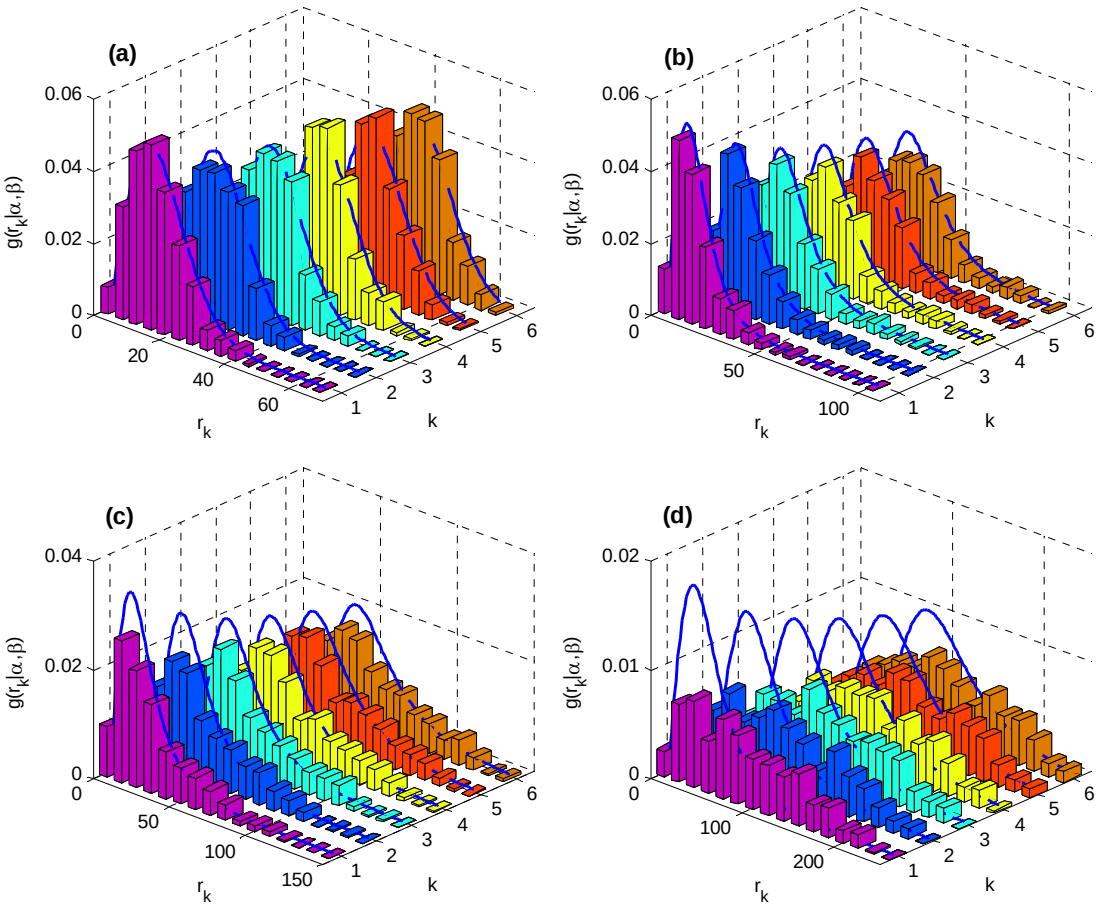

Fig. 3 Observed (histogram) and fitted (solid lines) probability density function of four spatial point patterns shown in Fig.1.

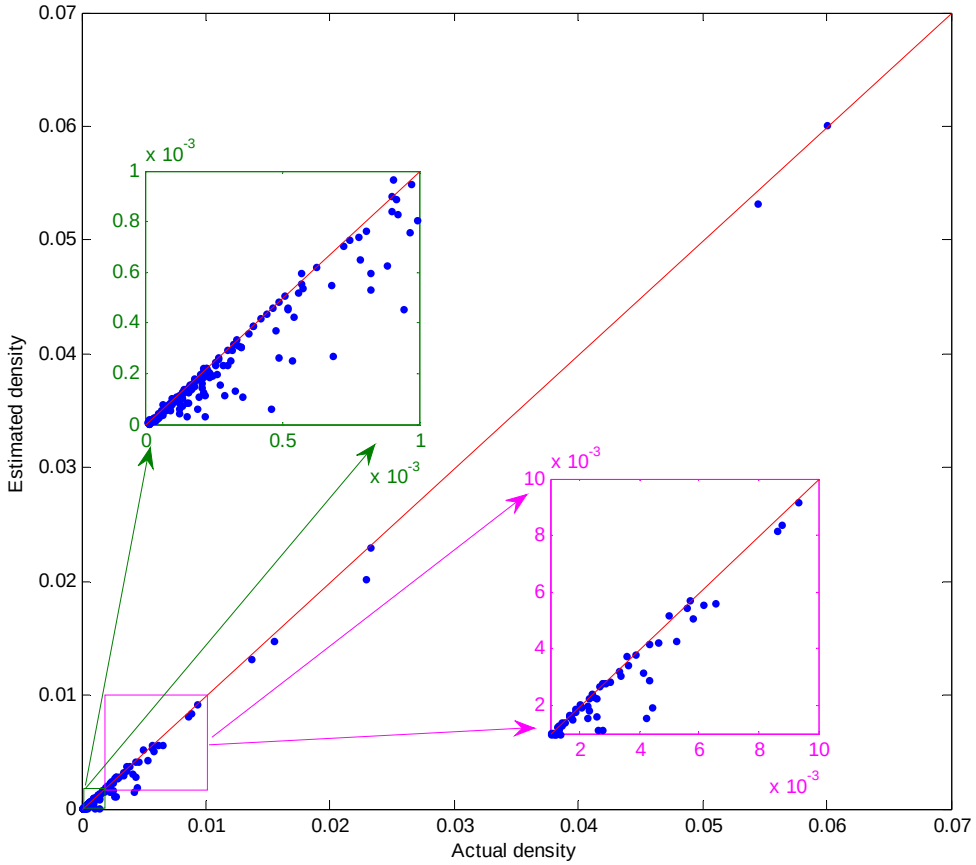

Fig. 4 Estimated vs. actual density of 259 tree species on Barro Colorado Island (BCI), Panama. The two inlets showed the details of bottom-left corner.

Fig.4 illustrates the results of density estimation for 259 tree species on Barro Colorado Island (BCI), Panama. Here, only density estimation for k = 6 was shown and the sample size was 50. We found that the estimated density is closer to the actual density for most tree species, especially for abundant tree species. However, tree density was usually under-estimated when tree species was rare. This result was consistent with that of simulated spatial point patterns indicating spatial clustering of rare species. Actually, previous study already verified that rare tree species was highly aggregate but not widely distributed within the study area (Condit et al., 2000).

4 Conclusion

Point-to-tree distance sampling was a popular plant inventory technique, especially in highly dense forest, and in difficult terrain (Picard et al., 2005; Kleinn and Vilćko, 2006). Previous study verified that Gamma-Poisson model performed well in fitting point-to-tree distance and density estimation (Magnussen et al., 2008ab). Although the density estimator based on Gamma-Poisson model suffered from the problem of bias, it was usually applied as option in constructing composite density estimator (Magnussen, 2012; Magnussen et al., 2012). In this paper, we proposed a new numerical optimization method, particle swarm optimization algorithm, to estimate the two parameters of Gamma-Poisson model. Compared with classical parameter estimation method, particle swarm optimization algorithm was more efficient and not constrained by the

sample size. It would improve the computational efficiency of composite density estimator that was recognized less biased and robust in forestry and ecology.

Acknowledgements

The study was supported by NSFC (31000197, 31570423), Knowledge Innovation Project of CAS (KZCX2-EW-QN209), and National High Technology Research and Development Program of China (2013AA06A211-4). Valuable comments from Cang Hui and other reviewer are also acknowledged.

References

Brix A, Cahdoeuf J. 2002. Spatio-temporal modelling of weeds by shot-noise Cox processes. Biometrica Journal, 44(l) : 83-99

Cassey P. 1999. Estimating animal abundance by distance sampling techniques. Conservation Advisory Science Noies No. 237, Department of Conservation, Wellington, New Zealand

Condit R et al. 2000. Spatial patterns in the distribution of tropical tree species. Science, 288: 1414-1418

Diggle PJ. 2003. Statistical Analysis of Spatial Point Patterns (2nd ed). Arnold, London UK

Eberhardt LL. 1967. Some developments in distance sampling. Biometrics, 23: 207-216

Gao M. 2013. Detecting spatial aggregation from distance sampling: a probability distribution model of nearest neighbor distance. Ecological Research, 28: 397-405

Gatrell AC, Bailey TC, Diggle PJ, Rowlingson BS. 1996. Spatial point pattern analysis and its application in geographical epidemiology. Transactions of The Institute of British Geographers, 21:256-274

He F, Gaston KJ. 2000. Estimating species abundance from occurrence. American Naturalist, 156:553-559

Hui C, McGeoch MA. 2007. Capturing the ‘droopy tail’ in the occupancy-abundance relationship. ÉcoScience, 14: 103-108

Hui C, McGeoch MA. 2014. Zeta diversity as a concept and metric that unifies incidence-based biodiversity patterns. American Naturalist, 184: 684-694

Hui C, Foxcroft LC, Richardson DM, MacFadyen S. 2011. Defining optimal sampling effort for large-scale monitoring of invasive alien plants: a Bayesian method for estimating abundance and distribution. Journal of Applied Ecology, 48: 768-776

Hui C, McGeoch MA, Reyers B, le Roux PC, Greve M, Chown SL. 2009. Extrapolating population size from the occupancy-abundance relationship and the scaling pattern of occupancy. Ecological Applications, 19, 2038-2048

Hui C, Veldtman R, McGeoch MA. 2010. Measures, perceptions and scaling patterns of aggregated species distributions. Ecography, 33: 95-102

Jonsson B, Holm S, Kallur H. 1992. A forest inventory method based on density-adapted circular plot size. Scandinavian Journal of Forest Research, 7: 405-421

Kleinn C, Vilćko F. 2006. A new empirical approach for estimation in k-tree sampling. Forest Ecology and Management, 237: 522-533

Legendre P, Fortin MJ. 1989. Spatial pattern and ecological analysis. Vegetatio, 80: 107-138

Magnussen S, Kleinn C, Picard N. 2008a. Two new density estimator for distance sampling. European Journal of Forest Research, 127:213-224

Magnussen S, Picard N, Kleinn C. 2008b. A Gamma-Poisson distribution of point to k nearest event distance, Foreat Science, 54(4): 429-441

Magnussen S, Fehrman L, Platt WJ. 2012. An adaptive composite density estimator for k-tree sampling, European Journal of Forest Research, 131: 307-320

Magnussen S. 2012. A new composite k-tree estimator of stem density. European Journal of Forest Research, 131: 1513-1527

Martinez WL, Martinez AR. 2002. Computational Statistics Handbook with MATLAB. Chapman & Hall, USA

Moore PG. 1954. Species in plant populations. Ecology, 35(2): 222-227

Picard N, Kouyaté AM, Dessard H. 2005. Tree density estimations using a distance method in Mali Savanna. Forest Science, 31:7-18

Pielou EC. 1960. A single mechanism to account for regular, random and aggregate populations. Journal of Ecology, 48: 575-584

Pielou EC. 1977. Mathematical Ecology. Wiley, New York, USA

Schwarz CJ, Seber GAF. 1999. Estimating animal abundance: review III. Statistical Science, 14: 427-456

Sheil D, Ducey MJ, Sidiyasa K, Samsoedin I. 2003. A new type of sample unit for the efficient assessment of diverse tree communities in complex forest landscapes. Journal of Tropical Forest Science, 15: 117-135

Stoyan D, Penttinen A. 2000. Recent applications of point processmethods in forestry statistics. Statistical Science, 15: 61-78

Sutherland WJ. 1998. Ecological Census Techniques: A Handbook. Cambridge University Press, USA

Thompson HR. 1956. Distribution of distance to n-th nearest neighbor in a population of randomly distributed individuals. Ecology, 37: 391-394

Turechek W, Madden LV. 1999. Spatial pattern analysis of strawberry leaf blight in perennial production systems. Phytopathology, 89 (5): 421-433

White GC, Bennetts RE. 1996. Analysis of frequency count data using the negative binomial distribution. Ecology, 77: 2549-2557

Zillio T, He F. 2010. Modeling spatial aggregation of finite populations. Ecology, 91(12): 3698-3706