Dispersal distance determines the exponent of the spatial Taylor’s power law

Pei-Jian Shi a,*, Hardev S. Sandhu b, Gadi V.P. Reddy c

a Bamboo Research Institute, Nanjing Forestry University, 159 Longpan Road, Xuanwu District, Nanjing 210037, China
b Everglades Research and Education Center, University of Florida, 3200 E. Palm Beach Road, Belle Glade, FL 33430, USA
c Western Triangle Agricultural Research Center, Montana State University, 9546 Old Shelby Road, P.O. Box 656, Conrad, MT 59425, USA

ARTICLE INFO

Article history:
Received 16 April 2016
Received in revised form 6 May 2016
Accepted 15 May 2016

Keywords:
Intensity
Investment
Mean–variance
Pair correlation function
Point process

ABSTRACT

The equation describing a power–law relationship between the mean and variance of population abundance in space or time is known as Taylor’s power law (TPL), initially observed in samples of insects. Factors determining the TPL exponent are of particular concern to ecologists because the observations of the exponent usually range 1–2. Recent studies have suggested that TPL is caused solely by statistical artifacts rather than biological processes, with the corresponding statistical models lacking linkages to explicit population demography. In this study, we used two special forms of the Neyman–Scott cluster point process to study the effect of offspring dispersal distance from the parents on the TPL exponent. Results showed that dispersal distance could largely affect the TPL exponent. The response curve of TPL exponent to dispersal distance is similar to the shape of the left-skewed gamma distribution function multiplied by a constant which can permit its maximum value to exceed 1. That means, short-distance dispersals could produce large TPL exponents relative to the whole response curve. However, the TPL exponent will decline in the case that the dispersal is extremely short or long. To better understand the function of the exponent of TPL on fitness, we attempted to link plant seed dispersal ability to the TPL exponent, and we discussed the trade-off between investing in propagation energy and in performance energy of plants. Dispersal overlap of offspring each other to an extent can cause a large TPL exponent, providing maximum fitness in a population. A novel theoretical frame was proposed to explain the role of spatial TPL relationships in affecting the fitness of plants.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Taylor’s power law (TPL) describes the relationship between the mean (M) and variance (V) of population abundance in both space and time (Taylor, 1961; Anderson et al., 1982). There is a great deal of empirical and theoretical evidence that demonstrates a power–law relationship between these two variables:

\[ V = aM^b. \]

where, \( a \) and \( b \) are constants. The value of \( b \) usually ranges from 1 to 2 (Taylor, 1961; Kilpatrick and Ives, 2003; Ballantyne and Kerkhoff, 2007). Several studies indicated that the TPL exponent could exceed 2 (Yamamura, 1990; Fronczak and Fronczak, 2010). This law has been accepted since the 1960s. The previous studies have proposed many biological mechanisms that could explain TPL. For example, Taylor and Taylor (1977) used an animal immigration model to study the migratory effect on TPL. Anderson et al. (1982) revealed TPL with simple population models under demographic and environmental influences, and they emphasized that the precise form of TPL might be determined by the relative magnitude of the various rate processes that govern the dynamics of population change and by the degree of spatial and temporal heterogeneity. Perry (1988, 1994) used biological relevant population dynamics models to analyze TPL. Kilpatrick and Ives (2003) used a stochastic logistic model to study the effect of interspecific competition on the TPL exponent. They found a severe degree of competition among species produced an exponent of less than two. Ballantyne and Kerkhoff (2007) showed that correlation among individual reproduction is related to different values of TPL exponent. Recent studies showed that TPL might be a general statistical pattern rather than being driven by a biological mechanism (Giometto et al., 2015; Xiao et al., 2015). Cohen and Xu (2015) demonstrated that the estimate of the TPL exponent is proportional to the skewness of a single fre-
quency distribution where observations are randomly sampled in blocks. *Xiao et al.* (2015) proposed that the TPL exponent might still contain ecological information. However, previous studies have failed in presenting an explicit biological factor that determines the exponent of spatial TPL. The work from Cohen’s group provides very detailed relationships between statistical variables and the mean–variance scaling patterns in TPL, but the statistical variables such as the skewness of the distribution (Cohen and Xu, 2015) and the transition probability (*Giometto et al.*, 2015) lack clear biological meanings. In addition, these stochastic multiplicative models are not spatially explicit and cannot account for spatial pattern. The usefulness of spatial point processes for plants in studying TPL has not been paid much attention before (Picard and Favier, 2011). We attempted to analyze the influence of dispersal distance of seeds on the TPL exponent using two spatial point process models with clear biological parameters.

2. Materials and methods

2.1. General formulae of mean and variance in spatial point processes

The intensity function and pair correlation function are the basic concepts of spatial point process theories (Diggle, 2014). The first-order properties of a spatial point process are defined by an intensity function as:

$$\lambda(x) = \lim_{|dx| \to 0} \left\{ \frac{E[N(dx)]}{|dx|} \right\}.$$  (2)

where $\lambda(x)$ denotes the intensity function at location $x$ on the two-dimensional Cartesian coordinate system. $E(\cdot)$ represents the expectation, and $|dx|$ denotes an infinitesimal region of $x$, while $|dy|$ denotes the area of $dy$ in an infinitesimal region. The second-order intensity function is defined as:

$$\lambda_2(x, y) = \lim_{|dx|, |dy| \to 0} \left\{ \frac{E[N(dx)N(dy)]}{|dx||dy|} \right\}.$$  (3)

Based on the first- and second-order intensity functions, the pair correlation function $g$ is defined as:

$$g(x, y) = \frac{\lambda_2(x, y)}{\lambda(x)\lambda(y)}.$$  (4)

where $\lambda(y)$ denotes the intensity function at location $y$.

Suppose $B$ denote any bounded area on the plane, i.e., $B \subset \mathbb{R}^2$, and let $N(B)$ denotes the number of points in $B$. The mean and variance of $N(B)$ can be calculated using the following formulae (Jalilian et al., 2012):

$$E[N(B)] = \int_B \lambda(x)dx$$  (5)

and

$$\text{Var}[N(B)] = \int_B \lambda(x)dx + \int_B \int_B \lambda(x)\lambda(y)[g(x, y) - 1]dx dy.$$  (6)

The above mean–variance formulae are general for all cases of two-dimensional point processes. For a complete spatial random point process (Diggle, 2014), the intensity function $\lambda(x)$ is a constant, $g(x, y) = 1$, and variance equals mean. According to the definition of the pair correlation function, it should be a positive real number. The case of $g > 1$ (or $g < 1$) corresponds to the possible attraction (or repulsion) between points (Jalilian et al., 2012). If $g < 1$, variance can be less than mean.

2.2. Two special forms of the Neyman–Scott cluster point process

The Neyman–Scott cluster point process can be used to describe many plants’ spatial distributions. In a given region, parent points are realizations of homogeneous or inhomogeneous Poisson point processes. Each parent point generates a cluster of offspring points based on a specific distributional pattern (e.g., Gaussian distribution). The locations of offspring points form a realization of the Neyman–Scott cluster point process. Here, we are concerned with two special forms of the Neyman–Scott cluster point process: the Matérn and Thomas (cluster point) processes. Both point processes have a ‘scale’ parameter that can be used to reflect the dispersal distance from a parent point. There are other forms of the Neyman–Scott cluster point process (e.g., *Tanaka et al.*, 2008), but for convenience we will only focus on the Matérn and Thomas processes.

Suppose that the intensity of parent points is $k$. For a Matérn cluster process, the locations of offspring points are independent and uniformly distributed inside a circle of radius $R$ centered on each parent point. The theoretical pair correlation function of a Matérn process is as follows (Baddeley et al., 2015):

$$g(r) = 1 + \frac{\lambda}{\pi} \left\{ \frac{1}{4} \left( \frac{r}{R} \right)^2 \right\},$$  (7)

where

$$h(z) = \begin{cases} 
\frac{16}{\pi} \left( z \arccos(z) - z^2 \sqrt{1 - z^2} \right) & \text{if } z \leq 1; \\
0 & \text{if } z > 1.
\end{cases}$$  (8)

Here, $r$ represents the distance of offspring points from a parent point. Assume that the number of offspring points of each parent point is a Poisson random variable with mean $\mu$. The theoretical intensity of a Matérn process is:

$$\lambda = \kappa \mu.$$  (9)

For a Thomas cluster process, the offspring points of one parent are independent and isotropically normally distributed around the parent point with standard deviation $\sigma$. The theoretical pair correlation function of a Thomas process is as follows:

$$g(r) = 1 + \frac{\lambda}{4\pi\sigma^2} \exp \left( -\frac{r^2}{4\sigma^2} \right).$$  (10)

The theoretical intensity of a Thomas process is similar to that of a Matérn process: $\lambda = \kappa \mu$. Below, we mainly examine the effect of parameter $R$ in the Matérn process and parameter $\sigma$ in the Thomas process on the exponent $b$ of TPL. Both of these parameters can reflect the dispersal distance of parent point, e.g., the dispersal distance of seeds in herbs, the spread distance of underground rhizomes of pygmy bamboos.

2.3. Simulation

In a simulated unit square (i.e., $[0, 1] \times [0, 1]$), we used the Dirichlet tessellation (Baddeley et al., 2015; Fig. 1) to divide the unit square into 50 tiles with different areas. We calculated the theoretical mean and variance for every tile based on Eqs. (5) and (6). We can then fit these data of variance versus mean to test the TPL. By suppose the intensity of parent points to be a constant $k$, we then set different values of $R$ in a Matérn process and different values of $\sigma$ in a Thomas process to generate different realizations (i.e., planar points) in the unit square. For a given $R$ or $\sigma$, we can obtain 50 data point pairs of variance versus mean from 50 tiles in every simulation. We fitted the data to obtain the estimates of $a$ and $b$ of TPL. For the given range of $R$ or $\sigma$, we obtained different estimates.
of $a$ and $b$ of TPL. Thus, we checked the effect of $R$ or $\sigma$ in representing the dispersal ability of plants on $a$ and $b$ of TPL. R (version 3.2.2) was used to perform simulation and parametric estimate (R Core Team, 2015). The package ‘spatstat’ (version 1.43-0.025) was used to calculate the mean and variance of point number in a given window in the Matérn and Thomas processes. The R scripts of these two simulations are shown in online Supplementary data 1.

2.4. Example: dispersal of a pygmy bamboo

$Sasa$ pygmaea (Miq.) E.-G. Camus is a pygmy bamboo. We recorded the spatial locations of all individuals in a 1 x 1-m plot on Jan 23, 2015 (Fig. 1). The study plot was located on the Nanjing Forestry University campus, Jiangsu Province, China (32°5’5.571”N, 118°48’39.315’E). The plot was well protected, and the distributional pattern at that time reflected a natural state rather than influence of human activities. The ‘kppm’ function of the ‘spatstat’ package was used to fit the parameters of the Matérn and Thomas processes. Let $X$ and $Y$ represent the horizontal and longitudinal coordinates of bamboos, respectively. In fitting the intensity of points in spatial point pattern analyzing methods, the log-linear function is usually used (Baddeley et al., 2015). Thus, the log intensity functions of $\sim X$, $\sim Y$, and $\sim X+Y$ in these two cluster points processes were compared. The log intensity with the lowest AIC value was selected, and the target log intensity was then used to check the effects of dispersal distance on the TPL exponent.

**Table 1** Comparison of goodness of fit using the Matérn and Thomas processes.

<table>
<thead>
<tr>
<th>Point process type</th>
<th>Log intensity function</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matérn</td>
<td>$\sim X$</td>
<td>−137,448</td>
</tr>
<tr>
<td>Matérn</td>
<td>$\sim Y$</td>
<td>−138,382</td>
</tr>
<tr>
<td>Matérn</td>
<td>$\sim X+Y$</td>
<td>−137,980</td>
</tr>
<tr>
<td>Thomas</td>
<td>$\sim X$</td>
<td>−137,430</td>
</tr>
<tr>
<td>Thomas</td>
<td>$\sim Y$</td>
<td>−138,365</td>
</tr>
<tr>
<td>Thomas</td>
<td>$\sim X+Y$</td>
<td>−137,964</td>
</tr>
</tbody>
</table>

The bold value indicates the smallest value of AIC.

3. Results

Fig. 2 shows the effects of scale parameter $R$ in the Matérn process and the scale parameter $\sigma$ in the Thomas process on the constant $a$ and exponent $b$ of TPL. When the scale parameter (that exceeded a critical value) increased, the constant $a$ of TPL first dropped rapidly below 1 and then increased very slowly toward 1. The exponent $b$ of TPL first increased to a peak and then dropped towards 1. In the current simulation, both $k$ and $\mu$ are fixed to be 20. The aforementioned conclusions still hold when we choose other values for these two parameters. Fig. S1 in online Supplementary data 2 exhibited the effects of different combinations of $k$ and $\mu$ values on the TPL exponent. Although the responses of TPL exponent to these combinations are slightly different, there is a general trend of its change that the TPL exponent increases at a ‘small’ scale relative to the whole response curve and then drops toward to 1. The mean of offspring points of every parent point (i.e., $\mu$) also can affect the response curve of TPL exponent to dispersal distance. Actually a larger $\mu$ can make the maximal TPL exponent be larger and the corresponding dispersal distance longer (see Fig. S1A and B in online Supplementary data 2). Scale parameters in both Matérn and Thomas processes represent the ability of parental points to cause the dispersal of their offspring. Thus, the exponent $b$ is related to the dispersal distance of the target species. Only if the dispersal distance of the species maintains a low proportion relative to the whole response curve, can we observe a large value of $b$. For the given range of scale parameters in either the Matérn process or the Thomas process, all the coefficients of determination obtained from fitting the nonlinear TPL model exceeded 0.995, indicating a good fit.

For the actual spatial distribution data of $S. pygmaea$, the Matérn cluster point process with log intensity function of $\sim Y$ was demonstrated to be the best because the corresponding AIC was the lowest (Table 1). Based on this model, the estimates of parameters were $\hat{k} = 35.97$, $\hat{R} = 0.099688$, and the estimated intensity function of the Matérn cluster point process was:

$$\hat{\lambda}(X, Y) = \exp(6.3883 - 0.6674 Y).$$

(11)

From this we calculated the reference intensity of mean number of offspring points per cluster using Eq. (9).

Based on the above estimates, using the general formulae of mean and variance of $N(B)$, Fig. 3 exhibits the relationship between variance and its mean for $S. pygmaea$, further demonstrating the existence of spatial TPL with an estimate on the exponent $\hat{b} = 1.503 \pm 0.016$. In addition, we can draw another conclusion that a low proportion of dispersal distance relative to the whole response curve can produce a large exponent $b$ of TPL.
4. Discussion

The negative binomial is frequently used by ecologists to describe the spatial distributions of arthropods and plants because the random variable is assumed to be discrete like the Poisson but its variance is larger than the mean (Bolker, 2008). Let $M$ denote the mean of the negative binomial. The corresponding variance equals $M + M^2/k$, where $k$ is a positive real number that is called an overdispersion parameter, and it is used to measure the amount of clustering. A smaller value of $k$ indicates more aggregation. When a population exhibits a complete spatial randomness in space and $k$ is infinitely great, then using the Poisson to describe the point process is more suitable. Taylor (1961) found that $k$ is not always independent of $M$. The negative binomial cannot account for the variation in $k$ for different sampling scales. Thus, the author built an empirical formula that variance is proportional to $M^p$, and found that it could fit well the observations of variance versus mean of insect and plant population densities. However, what has determined the size of exponent $b$ is always a puzzle (Xiao et al., 2015). Our study provides a theoretical explanation through the Neyman–Scott cluster point process. The greatest advantage of the Neyman–Scott cluster point process is that its parameters have explicit biological meanings. The offspring are usually generated from their parents with a limited dispersal distance (Nathan and Muller-Landau, 2000), thus it applies to many of the spatial distributions in animals and plants. Let’s assume that parents have a large dispersal distance for their offspring relative to the whole response curve of TPL exponent when $\kappa = \mu = 20$ (Fig. S2 in online Supplementary data 2), the spatial distribution consequently appears to be more homogeneous that will render $b$ to be approximately equal to 1. In this case, the offspring will have enough space to grow without strong competition. However, the evolution of such a long-distance dispersal ability might be attained at a great cost to investment in propagation energy. As a result, the parents might have fewer offspring. In addition, the separate individuals in spatial distribution will have lower survival rate than those in cluster when there are herbivores or predators (Nathan and Casagrandi, 2004). Conversely, if the dispersal distance is extremely short, the offspring will exhibit...
extreme clustering (Fig. S2 in online Supplementary data 2), leading to strong competition for light and nutrients (Cournède et al., 2008). In clustering regions, the self-thinning law will function to decrease the extent of aggregation by the density dependence. For instance, according to our recent observations, 70% bamboo shoots of Phyllostachys iridescens C.Y. Yao et S.Y. Chen could not grow into adults in the Nanjing Forestry University campus in the spring of 2016 because of severe light competition (unpublished data). For a species, it is not advantageous to have an extremely short dispersal distance, which needs to be long enough to reduce intraspecific competition. However, a long dispersal distance will require an additional investment in performance energy at the cost of propagation energy. In addition, for long-distance dispersal, seeds need some dispersal media such as wind and animals. Thus, some seeds of long-distance dispersed plants have evolved to be light than the short-distance dispersed seeds. For instance, the air lifted yellow poplar (Liriodendron tulipifera L.) seeds are averagely lighter than seeds at the forest floor in the eastern United States (Nathan et al., 2002). Plants likely experience a trade-off in benefits between investing in performance (i.e., the dispersal distance or evolution of special structures or light seeds that can facilitate dispersal by an external medium such as wind) and investing in propagation (the sheer number of offspring produced). It has been suggested that the mediate overlap of the distribution of offspring with that of their parents is beneficial for population recruitment (Nathan and Muller-Landau, 2000; Nathan and Casagrandi, 2004). The total energy investment of plants caused by environmental stresses that is used to produce offspring and generate special performance structures is hypothesized to be constant. The two types of energy investments (propagation and performance) can then be considered as essential resources for population recruitment. Using the resource-based competition theory of Tilman (1980), we can determine the optimal combination of these two investments to make the best use of the limited total energy (Fig. S3 in online Supplementary data 2). This trade-off combination determines the optimal dispersal distance. For example, with herbs our study suggests that short-distance dispersal of seeds is the optimal use of resources, and this conforms with actual observations on seed dispersal (Kanarek and Kao, 2011). They reported relatively short-distance dispersal of Bromus tectorum L. in Rocky Mountain National Park and the seeds fell in close proximity to parent plants. Even for many ground-dwelling arthropods, when food supplies are sufficient, their activity distances are usually very short. For instance, the effective trapping distances are less than 3 m for six species of common arthropods (Argiope bruennichi Scopoli, Blaps femoralis Fischer-Waldheim, Blaps murocana Latreille, Chlaenius bioculatus Chaudoir, Pardosa astrigera L. Koch, Telogryllus mitratus Burmeister) in the northern China (Zhao et al., 2013). Ground-dwelling arthropods usually lay eggs in cluster around food sources. In this case, we expect the TPL exponent to be around 1 because the dispersal distance is so short. In the original paper of Taylor (1961), there were many insect species for which the estimates of b were approximately equal to 1. We assumed that the trade-off between propagation energy investment and performance energy investment not only determines the optimal dispersal distance, but also the TPL exponent. He and Gaston (2003) found a link between TPL and the occupancy-abundance pattern. Our study provides an understanding of TPL that makes it applicable for considering such a link between the patterns of occupancy-abundance from the angle of dispersal distance. In fact, species dispersal ability can significantly affect the proportion of occupied sites in a self-organized community (Tilman et al., 1997). This study suggests that further investigation of TPL from the perspective of a species’ dispersal ability and optimal population size in a given environment is needed. We considered that a range from 1 to 2 for the exponent $b$ of TPL represents an optimal dispersal of seeds (Fig. S4 in online Supplementary data 2). The spatial distribution of soil nutrients and interspecific competition also can affect the spatial distribution pattern of plants to an extent. The influence of interspecific competition is highly dependent on the availability of soil nutrients (Aerts, 1999). Therefore, the most crucial determinants of the exponent of TPL are the spatial dispersal ability of plants and spatial distribution of soil nutrients. However, some competitor can win by adjusting the spatial level of soil nutrients. If the above assumption holds, we can make deductions about the inter-specific competition of plants. If a plant can cause another plant to change its spatial distribution pattern, as for example with leguminous plants and their competitors, then the changed species might decline or even die out in a given area. Leguminous plants can increase the nitrogen concentration in soil disturbing the typical spatial distribution pattern of competitors, leading to a decrease in the competitors’ fitness. Presumably, leguminous plants also undergo change in the exponent of TPL, but these changes should be smaller than that of the competitors. Future studies should compare the exponent of TPL in non-competitive and that in a competitive environment.

Acknowledgements

We thank Peter J. Diggle, Adrian Baddeley, Rolf Turner, Cang Hui, and Zi-Hua Zhao for their invaluable help during the preparation of this manuscript. We also thank the editor-in-chief, Brian D. Fath, and two anonymous reviewers very much for their invaluable comments. This work was supported by the Key Project of National Science & Technology Ministry (2012BAD23B05), the National Natural Science Foundation of China (grant number 31400348), and the PAPD of Jiangsu Province.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ecolmodel.2016.05.008.

References


