

# Spatial ecology of a root parasite – from pattern to process

DAVID M. WATSON,<sup>1\*</sup> DAVID A. ROSHIER<sup>1</sup> AND THORSTEN WIEGAND<sup>2</sup>

<sup>1</sup>*Institute of Land, Water and Society, Charles Sturt University, Albury 2640, Australia (Email: dwatson@csu.edu.au); and* <sup>2</sup>*UFZ Centre for Environmental Research Leipzig-Halle, Department of Ecological Modelling, Leipzig, Germany*

**Abstract** Occurrence patterns of parasitic plants are constrained by the distribution of suitable hosts and movement patterns of seed vectors and, accordingly, represent a simplified system to study many aspects of spatial ecology and determinants of distribution. Previous work has focused on the aerially hemiparasitic mistletoes, and it is unclear whether root parasites are affected by similar factors. Here, we evaluate spatial patterns in the root parasitic *Santalum lanceolatum* in an arid shrubland in north-western New South Wales, central Australia. In this region, the principal host is a long-lived nitrogen fixing shrub *Acacia tetragonophylla* closely associated with ephemeral creek-lines. The location of 765 individuals of both species was mapped along a 250-m section of creek-line using a total survey station, and occurrence patterns of the root parasite related to host distribution and landscape context. We used Ripley's K-function and the O-ring statistic to determine whether the distribution of *S. lanceolatum* was random, aggregated or regular; the spatial scales at which these patterns occurred; and to quantify any spatial associations between the parasite and its host, *A. tetragonophylla*. While acacias were closely associated with the creek-line, *S. lanceolatum* plants were more tightly clustered, displaying significant clustering at two spatial scales (1.2 m and 8.8 m). We suggest that host quality may act as an important constraint, with only those acacias growing in or near the creek-line being physiologically capable of supporting a parasite to maturity. Insights gained from spatial analysis are used to guide ongoing research in this system, and highlight the utility of the O-ring statistic for understanding patterns of distribution affected by multiple processes operating at critical scales.

**Key words:** dispersal, distribution, host quality, O-ring statistic, *Santalum*.

## INTRODUCTION

Distribution patterns of parasitic plants are tightly defined by host distribution and movement patterns of the vector, and have been studied in a range of systems. Most of this research relates to the aerially hemiparasitic mistletoes, and a sound understanding is emerging about the factors affecting their occurrence and spread (Reid *et al.* 1995; Overton 1996; Watson 2001, 2004; Aukema & Martinez del Rio 2002a,b; Bowie & Ward 2004; Shaw *et al.* 2004). Although root parasitism is more widespread, considered to have arisen independently in 10 different Angiosperm divisions (Nickrent 2002), little work has been conducted on distributional ecology of root parasites. Other than documenting host ranges and geographical distribution (Brand & Jones 2002; Woodall & Robinson 2003), research on root parasitic plants has focused on physiological and anatomical aspects of parasitism (Pate *et al.* 1990; Tennakoon *et al.* 1997), horticulture of

commercially valuable species (Sen-Sarma 1977; Radomiljac 1998), and methods to control pest species (Parker & Riches 1993).

Unlike the aerially hemiparasitic mistletoes, root parasites are not wholly dependent on a single host for all mineral and water requirements. In addition to parasitizing multiple hosts simultaneously, root parasites may also access water and nutrients directly from the soil (Ehleringer & Marshall 1995). As such, they may not be as strictly tied to distribution patterns of particular hosts, and it is unclear which factors would be most influential in constraining occurrence patterns. Are they similar to mistletoes and constrained primarily by biotic factors (vector behaviour, availability of suitable hosts), or do abiotic factors that structure autotrophic plant communities (availability of water, nutrient, competition) play a greater role? One way to reveal the factors constraining root parasites is to examine their spatial occurrence relative to these biotic and abiotic factors. By combining detailed spatial mapping of plants and hosts relative to environmental gradients, initial hypotheses about root parasite distribution can be formulated.

Our overall aim in this research is to gain a better understanding of the factors constraining parasitic

\*Corresponding author.

Accepted for publication July 2006.

plant occurrence, comparing root parasites and mistletoes and identifying underlying mechanisms driving observed patterns. In this contribution, we have three objectives: (i) to describe the spatial occurrence of a root parasite relative to its host and position in the landscape; (ii) to reconcile this descriptive information with our understanding of the biotic and abiotic factors affecting parasitic plants, generating a detailed set of priorities for further research; and (iii) to demonstrate the utility of the O-ring statistic for examining spatial patterning in point processes.

## METHODS

### Study site

This research was conducted in the arid rangelands of central Australia, with the study site located in Sturt National Park, north-western New South Wales. The study area has hot summers and mild winters with an annual average of 18.4 days with maximum temperatures above 40°C and 1.5 days with minimum temperatures below freezing (Tibooburra Post Office, 22 km west of study site; Bureau of Meteorology, Melbourne). Rainfall is summer (December–March) dominant and highly variable, averaging 228 mm per annum over a 118 years record, with median (5th decile) monthly rainfall in these months ranging from 12.6 mm to 5.8 mm.

A broad range of habitats occurs in this region (Keith 2004), with open eucalypt woodlands associated with larger creeks and a variety of acacia species growing near smaller drainage lines. Intervening areas support acacia woodlands, grading into Eremophila-dominated shrublands, open grasslands, sparse chenopods and stony downs country largely devoid of woody vegetation. Dead finish *Acacia tetragonophylla* is one of the more common shrubs in the region, growing at low to medium density in close proximity to most of the ephemeral watercourses in the National Park. While it intergrades with mulga *Acacia aneura* and gidgee *A. cambagei* dominated woodlands, it commonly grows in monodominant stands along ephemeral drainage lines, forming narrow meandering networks within large areas otherwise devoid of woody vegetation (Fig. 1, photo).

All data presented here relate to a single site, located 1.3 km NE of Mt Wood homestead (29.480°S, 142.230°E), along an unnamed ephemeral creek-line (referred to hereafter as Mistletoe Creek). Boundaries were determined relative to naturally occurring features – including all plants either side of a 250-m section of creek-line – to define an area of approximately 3.4 ha. Mistletoe Creek continued for approxi-



**Fig. 1.** Photograph of study site showing main drainage line (centre) and associated shrubs. Tallest shrubs are *Santalum lanceolatum* and roundish shrubs *Acacia tetragonophylla*. Person on far right for scale (indicated by arrow).

mately 2.3 km upstream and 1.2 km downstream of the study site, and along its entire length the habitat was dominated by *A. tetragonophylla*.

### Study species

Within the region *A. tetragonophylla* is frequently parasitized by three species of hemiparasite – the root parasitic northern sandalwood (Santalaceae; *Santalum lanceolatum*) and two stem-parasitic mistletoes (Loranthaceae; *Amyema* spp.). While the two *Amyema* species are also being studied on this site, they will not be discussed further here.

The root parasitic *S. lanceolatum* is a long-lived slow growing shrub that commonly parasitizes leguminous shrubs and trees, especially *Acacia* species (Johnson 1996). It has small, white, insect pollinated flowers, and fleshy red/black drupes containing a single seed 5–7 mm long. Little is known about the dispersal ecology of *S. lanceolatum*. Forde (1986) lists emu *Dromaius novaehollandiae* and several honeyeaters (spiney-cheeked honeyeater *Acathagenys rufogularis*, striped honeyeater *Plectorhyncha lanceolata*) as presumed vectors, but this is drawn solely from dietary records and there have been no specific studies of seed dispersal per se. Recent research on the congeneric *Santalum spicatum* in south-western Australia has discovered that brush-tailed bettongs or woylies *Bettongia penicillata* (Potoroidae) act as the sole seed disperser (Murphy *et al.* 2005), caching seeds in small scrapes near roots of potential host plants. This 1.3 kg marsupial historically occurred throughout central Australia and may have acted as a dispersal agent for *S. lanceolatum* before becoming regionally extinct 100 years ago (Dickman *et al.* 1993).

## Mapping

All woody plants were identified, measured (maximum height and breadth for the multistemmed *A. tetragonophylla*; maximum height and diameter of the trunk or largest stem for *S. lanceolatum*) and marked using permanent metal tags with unique alphanumeric codes. The location of every woody plant was mapped with tripod-mounted electronic total survey stations (Topcon, model no. GTS-226), using lasers to measure distances from a fixed point to the base of each plant to an accuracy of <50 mm.

## Spatial analysis

In order to describe the position of individual plants in the landscape relative to the main drainage feature running the length of the site, we fitted a global trend model to the topographic (height) data and subtracted the trend from the height data. This produced a surface model in which locations with a similar height above the thalweg had the same elevation.

We used second-order statistics, Ripley's K-Function (Ripley 1976, 1981) and the O-ring statistic (Wiegand *et al.* 1999; Condit *et al.* 2000; Wiegand & Moloney 2004), to determine whether the distribution of *S. lanceolatum* was random, aggregated or regular; the spatial scales at which these patterns occurred; and the nature of any spatial associations between *S. lanceolatum* and its host in this system, *A. tetragonophylla*. Ripley's K is a cumulative frequency distribution of observations at a given point-to-point distance and the function  $K(r)$  is the expected number of points in a circle of radius  $r$  centred on an arbitrary point, divided by the intensity of the pattern. Ripley's (1976) estimator of the K-function is:

$$\hat{K}(r) = n^{-2} |A| \sum_{i \neq j} \omega_{ij}^{-1} I_d(d_{ij})$$

where  $r$  is the radius of a circle centred on a point in the pattern,  $n$  is the number of points in region  $A$  with an area  $|A|$ ,  $d_{ij}$  is the distance between the  $i^{\text{th}}$  and  $j^{\text{th}}$  points, the weight  $\omega_{ij}$  corrects for edge effects and is the proportion of the area of a circle centred at the  $i^{\text{th}}$  point with radius  $d_{ij}$  that lies within the study region, and  $I_d$  is an indicator function which is 1 if the distance  $d_{ij}$  between points  $i$  and  $j$  is  $\leq r$ , else  $I_d = 0$ . This formula can be applied to univariate and bivariate point patterns. Ripley's K is most often visualized as a transformed linear L-function (Besag 1977):

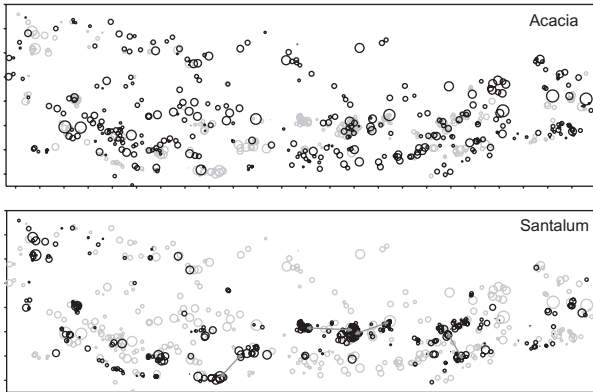
$$L(r) = \sqrt{\frac{K(r)}{\pi}} - r$$

For a univariate point pattern, values of  $L(r) > 0$  indicate that the points are clumped at distances up to

$r$ , while values of  $L(r) < 0$  indicate they are regularly dispersed. For a bivariate point pattern, positive values of  $L(r)$  indicate that there are more points of one pattern within a given distance  $r$  than of points in the other pattern than would be expected if the patterns were independent (attraction). Similarly, negative values of  $L(r)$  indicate repulsion between the two patterns at distance  $r$ . The significance of any observed pattern that deviates from the distribution expected under a null model, often complete spatial randomness, is assessed by comparing the observed distribution function to the confidence envelope generated by Monte Carlo simulations of a null model (Besag & Diggle 1977; Diggle 2003).

The O-ring statistic  $O(r)$  (Wiegand *et al.* 1999; Condit *et al.* 2000; Wiegand & Moloney 2004) is an analogue of Ripley's K that, for isotropic patterns, gives the expected density of points within a ring or annulus a specified distance from an arbitrary point in the pattern. Sometimes it is more convenient to use the density normalized pair-correlation function  $g(r) = (1/\lambda) O(r)$  instead of the O-ring statistic;  $\lambda$  is the intensity of the pattern (Stoyan & Stoyan 1994). Like Ripley's K, the O-ring statistic can be formulated for univariate and bivariate point patterns. For univariate point patterns,  $O(r) > \lambda$  indicates clumping, while values of  $O(r) < \lambda$  indicates regularity. For a bivariate point pattern,  $O(r) > \lambda$  indicates attraction whereas  $O(r) < \lambda$  indicates repulsion. The significance of the observed pattern is assessed by comparing the observed distribution function to the confidence envelope generated by Monte Carlo simulations of a null model (see Wiegand & Moloney 2004 for an extended discussion of the formulation and estimators of Ripley's K and the O-ring statistic).

The form of the O-ring statistic contains direct information on the frequency of point-to-point distances and can be used to assess the nature of the underlying process (Stoyan & Stoyan 1994). For example, if competitive interactions are so strong that plants do not come near to each other (or if plants have a finite size), the O-ring statistic will show a typical hard-core distance  $r_H$  where  $O(r) = 0$  for  $r < r_H$  because no plant-plant distance below  $r_H$  exists. On the other hand, if competitive interactions are relatively unimportant compared with clustering mechanisms, the O-ring statistic will show a peak at small distances as small plant-plant distance occur frequently. If the spatial structure of the pattern is ordered (such as in crystals) the O-ring statistic shows several peaks indicating the first, second, third, etc. nearest neighbour distance (Stoyan & Stoyan 1994). The O-ring statistic is not a cumulative statistic and therefore does not integrate the 'memory' of small scale second-order effects to larger scales, as does Ripley's K (Wiegand & Moloney 2004).



**Fig. 2.** The spatial pattern of *Acacia tetragonophylla* plants (top) and *Santalum lanceolatum* plants (bottom) indicated by black circles; the grey circles show the respective other pattern. Circle diameter corresponds to plant dimensions, and the position of these circles indicates the relative position of the plants as mapped, with the creek-line flowing from left to right. The ticks are placed every 10 m. Note that there are some individual-rich *S. lanceolatum* clusters approximately 14 m apart (indicated by arrows).

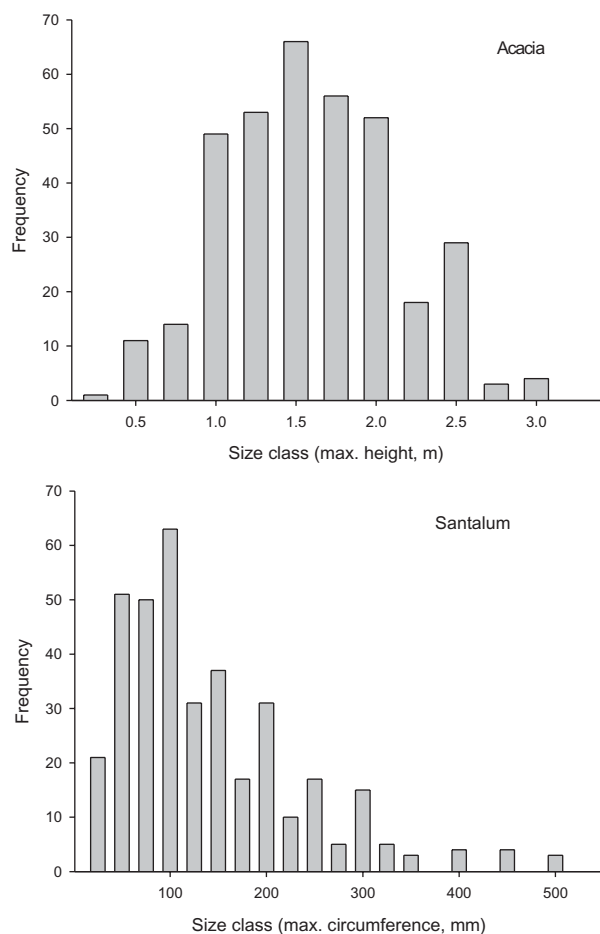
If the intensity of the distribution of points within a region is not homogenous, apparent aggregation or attraction within a point pattern may be due to first-order effects in the environment such as soil properties or drainage patterns. In such circumstances, the null model of complete spatial randomness (CSR) is inappropriate for exploring second-order effects, and the variation in intensity must be accounted for to reveal any point-to-point interactions within or between point patterns (Pelissier & Goreaud 2001; Diggle 2003; Wiegand & Moloney 2004). Visualization of the distributions of *A. tetragonophylla* and *S. lanceolatum* (Fig. 2) suggested that the distributions may show two nested scales of clustering: a large scale clustering due to properties related with the creek-line, and small-scale clustering due to plant-plant interactions. From a statistical viewpoint, clustering and heterogeneity cannot be distinguished and yield indistinguishable second-order statistics (Diggle 2003). We used this property to determine the two critical scales of clustering by fitting the second-order statistics of a nested double-cluster point process to those observed (see Appendix i). In this process, points of a pattern that shows large scale clustering with a critical scale  $\sigma_L$  (described by a Thomas process, see Appendix i) are replaced by smaller clusters of points with a critical scale  $\sigma_S$ , resulting in a pattern of nested clustering. The critical scales  $\sigma_L$  and  $\sigma_S$  are the standard derivations of bivariate Gaussian distributions describing the distances of the points of a cluster relative to the cluster centre.

We hypothesized that the environmental heterogeneity induced by the creek-line should be approxi-

mately the same for the host and parasite. This corresponds to the null hypothesis that the larger-scale distribution of the parasite follows directly that of its host. On a smaller scale, however, processes such as localized seed dispersal may cause additional clustering of the parasite. We performed several analyses to test this hypothesis. First, we tested whether the second-order properties of the patterns of both the host and the parasite can be approximated by a nested double-cluster process and if the critical scales of scales  $\sigma_L$  of the large scale clustering and the number of clusters are approximately the same. Second, we tested if the larger-scale second-order properties of both patterns can be approximated by heterogeneous Poisson processes where the constant intensity of the homogenous Poisson process (CSR) is replaced by an intensity function that varies with location. This requires the selection of an 'appropriate' radius of a moving window estimate of the intensity function (see Wiegand & Moloney 2004 for discussion of use of appropriate null models). This approach avoided the need for arbitrarily defining bounded polygons to exclude areas away from the main drainage feature (where no woody plants occurred). We used the critical scale  $\sigma_L$  of the large-scale clumping to determine the radius of the moving window. As two-thirds of all points of a radially symmetric bivariate Gaussian distribution fall within distance  $1.5 \sigma_L$  from the cluster centre we selected  $1.5 \sigma_L$  as radius of the moving window. Third, to analyse whether and how the spatial distribution of *S. lanceolatum* is related to the distribution of the host *A. tetragonophylla*, we tested random labelling of a bivariate point process to find out whether host plants are more frequently surrounded by conspecifics than by parasite plants (and vice versa). We treated the distribution of the host as fixed (an antecedent condition) and randomized the distribution of *S. lanceolatum* using a heterogeneous Poisson point process as the null model. All analyses were done using the grid-based estimators in the Programita software package (Wiegand & Moloney 2004).

## RESULTS

A total of 370 *A. tetragonophylla* plants occurred within the study area and the population exhibited a normal size class distribution (Fig. 3), dominated by shrubs 1.0–2.5 m in overall dimensions, with few small and large individuals and no very large individuals (i.e. no plants exceeded three standard deviations of the mode). However, the 395 mapped *S. lanceolatum* plants displayed a different pattern, with a strongly left-skewed size class distribution (Fig. 3) because of a



**Fig. 3.** Size class distributions for *Acacia tetragonophylla* plants (top) and *Santalum lanceolatum* plants (bottom).

large number of small plants and several very large individuals (more than four standard deviations greater than the mode).

The spatial distribution of both *A. tetragonophylla* and *S. lanceolatum* was clumped, with all individuals of both species growing within 1 m above the thalweg. The form of the O-ring statistic of the *A. tetragonophylla* spatial pattern (Fig. 4A) is typical for a clustered pattern without hard core distance; the maximum occurs at a small scale of 1 m where plants are about four times as frequent as expected under their overall density in the study window. The absence of a hard core distance indicates that it may be favourable to grow close to each other. The absence of local maxima and minima indicates that there was no order with preferred plant-to-plant distances in the pattern. The form of the O-ring statistic of *S. lanceolatum* (Fig. 4B) indicates an extremely clustered pattern without hard core distance. Pairs of *S. lanceolatum* plants occurred, for example, at a distance of 1 m 30 times as frequently as expected under their overall density, and at a distance of 5 m

were still three times as frequent. The two local maxima at 14 m and 20 m (Fig. 4D) indicate a certain order in the pattern.

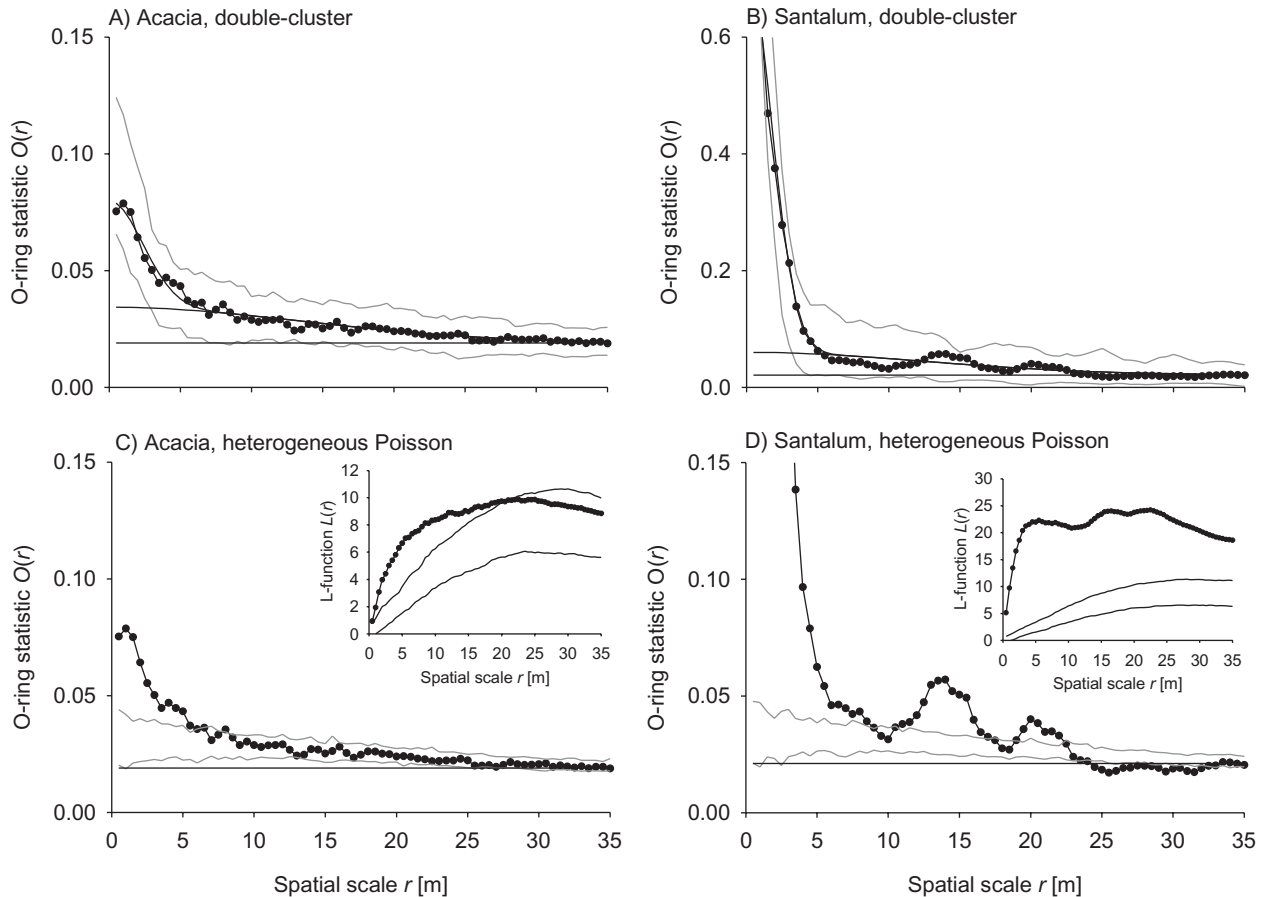
### Fit with double-cluster model

For *A. tetragonophylla* we found a critical scale of large-scale clustering of  $\sigma_L = 9.7$  m and a parents intensity of  $\rho_L = 0.0011$  parent per square metre which yields some 20 parent events. The double-cluster model yielded a good fit of the data (Fig. 4A), and the critical scale of the small-scale clustering was  $\sigma_S = 1.7$  m and the parents intensity  $\rho_S = 0.012$  parent per square metre (224 parents, on average 1.6 plants per cluster). *Acacia tetragonophylla* plants at the <3 m neighbourhood were two times as frequent as expected under the large-scale clustering (Fig. 4A, compare solid curves showing fit with single and double-cluster models). Monte Carlo simulation of the double cluster process showed that the spatial pattern of *A. tetragonophylla* plants is well described by this process.

For *S. lanceolatum* we found a critical scale of the large scale clustering of  $\sigma_L = 8.8$  m which is slightly smaller than the 9.7 m found for *A. tetragonophylla*. This finding supports our hypothesis of a common environmental heterogeneity. However, there were only 10 parent events (i.e.  $\rho_L = 0.0006$  parent per square metre) which indicates that *S. lanceolatum* used less habitat clusters than *A. tetragonophylla* and was thus more strongly clustered at the larger scale. Despite not showing the oscillatory behaviour, the double-cluster model approximates the data well (Fig. 4B). The parameters of the small-scale clustering are  $\sigma_S = 1.2$  m (*vs.* 1.7 m for the host) and  $\rho_S = 0.0017$  parent per square metre (32 parent events, on average 12 plants per cluster). *Santalum lanceolatum* plants at the 1.5 m neighbourhood were about eight times as frequent as expected under large-scale clustering (Fig. 4B, compare solid curves showing fit with single and double-cluster models). Monte Carlo simulation of the double cluster process showed that the data are well described by this process (Fig. 4B).

### Heterogeneous Poisson model

In the previous analyses we found that both host and parasite follow large-scale clustering with similar patch sizes (captured by the parameter  $\sigma_L$  of the double-cluster model). This, however, does not tell us whether the large scale clusters actually overlap. Therefore, we contrasted the patterns of the host and the parasite to heterogeneous Poisson null models where the intensity functions were constructed from the host pattern using a 15 m moving window radius.



**Fig. 4.** For the univariate analyses, the O-ring statistic (black dots) is presented, together with confidence envelopes for the lowest and highest  $O(r)$  taken from 99 simulations of the null model (grey lines). The solid horizontal lines give the overall intensity of the patterns in the study window. All analyses were performed with a cell size of  $0.5 \times 0.5$  m and a ring width of 1 m. Panel A illustrates the host pattern and the double-cluster null, with the solid curves illustrating the fitted single and double-cluster models. Panel B is for the parasite pattern and the double-cluster null, with solid curves showing the fitted single and double-cluster models. Panel C illustrates the host pattern and the heterogeneous Poisson null model with a 15-m moving window radius, and the inset figure illustrates the result using the  $L$ -function. Panel D illustrates the parasite pattern and the heterogeneous Poisson null model with a 15-m moving window radius using the pattern of the host, and the inset shows the results for the  $L$ -function.

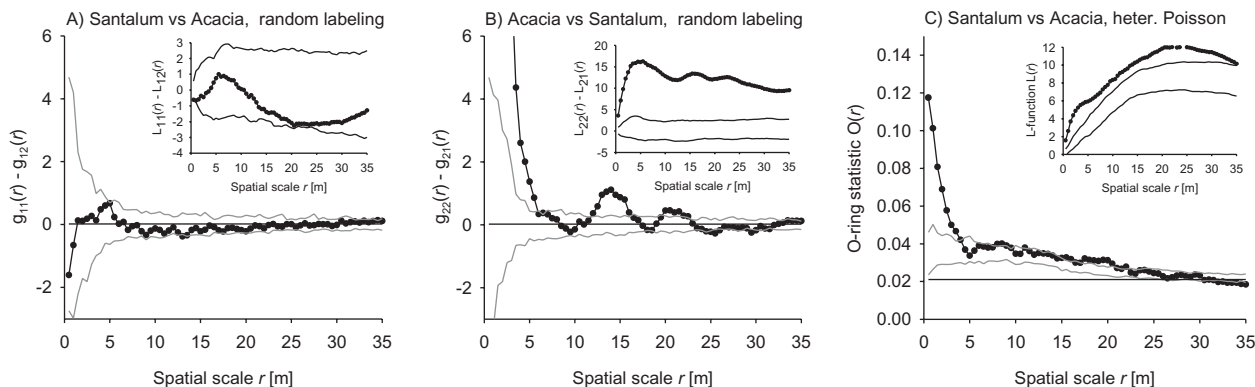
The distribution of *A. tetragonophylla* diverged only from the null model at distances of  $<5$  m due to small-scale clumping (O-ring statistic; Fig. 4C). Likewise, *S. lanceolatum* plants were also highly clumped, but the fit of the distribution to a heterogeneous Poisson process null model based on the host distribution was poor. In general, the large scale clustering of *S. lanceolatum* was stronger than that of the null model. The reason for this is, as already indicated by the analysis using a double cluster process, that the root parasite used less habitat clusters than the host. Additionally, the order in the pattern induced by individual-rich clusters at some 14 m and 20 m distance and regular dispersion at distances 25–30 m was not captured by the null model. This suggests that the two species were not constrained by environmental heterogeneity in the same

manner and therefore the processes determining the observed distribution must have an alternative explanation.

Analyses using Ripley's  $K$  suggested significant clumping in the distribution of both species at distances less than 20 m ( $L(r)$ , Figure 4C,D), but the apparent clumping at larger distances than that suggested by the O-ring statistic is probably due to 'virtual aggregation' – the observed strong aggregation at small distances influencing a cumulative distance function at larger distances (Wiegand & Moloney 2004).

### Bivariate analyses

Analysing the difference  $g_{11}(r) - g_{12}(r)$  that compares the relative densities of points of the two patterns in rings



**Fig. 5.** For the bivariate analyses, Panel A illustrates results for random labelling with test statistic  $g_{11}(r) - g_{12}(r)$ , which tests whether host plants are more frequently surrounded by conspecifics than by parasites. Panel B is the same as A, but patterns are exchanged. In panel C, results are for an antecedent condition null model treating the distribution of the host as fixed and randomizing the distribution of the parasite using a heterogenous Poisson point process based on the host distribution and a 15-m moving window. Confidence envelopes were based on 99 simulations and the insets show results for the  $L$ -function.

around the host (relative because they are divided by the respective intensities  $g_{i,i}(r) = O(r)/\lambda_i^{-1}$ ,  $i = 1, 2$ ) reveals that the host plants were surrounded on average at all distances by the same number of host and parasite plants (Fig. 5A). This is an interesting result that points to a neighbour balance of the hosts where one parasite plant is equivalent to roughly one host plant ( $395/356 = 1.1$ ). In other words, available resources may allow only for a certain density of neighbour plants around the host plants. However, as anticipated by the univariate analyses, the parasites are surrounded at distances  $< 5$  m (and around 14 m and 20 m) by many more parasites than hosts (Fig. 5B). This apparent asymmetry between host and parasite suggests a strong clustering mechanism of the parasite, which is limited by the capacity of the host to support or tolerate neighbours.

In relating parasitic plant occurrence to host distribution, the intensity of the host's distribution was used to distribute points in a Monte Carlo simulation, to reveal whether there was attraction or repulsion between host and parasite. Analyses using the O-ring statistic showed a clear attraction which was limited to distances less than 3 m (Fig. 5C), whereas Ripley's  $K$  as  $L(r)$  suggests that *S. lanceolatum* was significantly attracted to *A. tetragonophylla* at all distances up to 35 m.

## DISCUSSION

To our knowledge, this is the first quantitative study of the spatial ecology of a root parasitic plant worldwide, hence it is difficult to evaluate the generality of our findings or determine whether the results for our one site are applicable to other systems. It is noteworthy that

our findings are supported by parallel research on aerially hemiparasitic mistletoes (Overton 1996; Norton & Stafford Smith 1999; Watson 2001, 2004; Aukema & Martinez del Rio 2002a). Despite the ability of *S. lanceolatum* to parasitize multiple hosts, and access water and nutrients directly, we found that plants were strongly clumped and closely associated with their host at distances of  $< 3$  m. Both *A. tetragonophylla* and *S. lanceolatum* exhibited clumped distribution patterns at two critical scales: a large scale of about  $\sigma = 10$  m (a disk with radius  $2\sigma$  contains 86% of the plants in a cluster) and a tighter clustering of about 1.7 m and 1.2 m respectively. The large-scale clustering corresponds to their restriction to less than 1 m above the thalweg. Despite having a smaller small-scale cluster size, *S. lanceolatum* showed much stronger small-scale clustering than *A. tetragonophylla*, due to a lower number of clusters which comprised, on average, eight times more plants. Unlike *A. tetragonophylla*, the distribution of *S. lanceolatum* was not described well by the simple model, suggesting that the underlying process determining clustering in this species was not the same as for the host. A conspicuous feature of the parasite pattern was the existence of individual-rich clusters some 14 m away which imprinted a strong signal on the O-ring statistic (Fig. 2).

In evaluating the processes leading to the strong aggregation in *S. lanceolatum*, we offer two potential mechanisms. In the absence of dispersers (Murphy *et al.* 2005), *S. lanceolatum* seeds may simply germinate where they fall, leading to a halo of young seedlings around a parent plant. Alternatively, this clustering may reflect suckering – vegetative shoots growing from the stem base or from near-surface lateral roots – a strategy that has been recorded in other populations of this species (Warburton *et al.* 2000). This suckering is more

likely for plants growing in disturbed locations, such as areas adjacent to high velocity flows along ephemeral creek-lines. These two processes would be expected to yield clustering at different scales – one approximating the extent of the crown (<2 m), the other corresponding to the extent of the lateral roots.

In treating the univariate distribution of *S. lanceolatum* as a clustering process, aggregation was revealed at two critical scales – 8.8 and 1.2 m. Thus, rather than being driven exclusively by either seed rain or suckering, we suggest that the tightly clumped occurrence of young sandalwoods may be due to both processes operating in concert. The 14 m distance between individual-rich clusters may be the product of suckering from the previous generation, with the clusters corresponding to the past location of parent plants. Further work will be needed to establish whether animals act as seed dispersers and, if so, how far they move seeds. Physically excavating roots (after Woodall & Robinson 2003) would establish upper limits on the extent of lateral roots and may also reveal whether plants in this area are clonal, but given that other species of *Santalum* are known to graft on to each other, this would not be conclusive. Rather, genotyping individuals would be definitive (e.g. Warburton *et al.* 2000), and this is a longer-term goal of our ongoing research.

Neither host-parasite dependence, seed-rain, nor suckering can explain the clustering of sandalwoods at greater distances (i.e. their restriction to areas adjacent to the creek-line) and, given their parasitic habit, direct access to water cannot explain this clustering. Rather, we suggest variation in host quality is involved, consistent with the recently articulated host quality hypothesis (Watson, in press). Even though Mistletoe Creek rarely flows, shrubs growing on or near the creek bed would have access to more water for longer periods. Hence, regardless of where seeds are deposited, or the potential range of suckers emerging from lateral roots, we suggest *S. lanceolatum* plants only establish and grow near hosts of sufficient vigour.

While access to sufficient water is just one aspect of host quality, we suggest that it is especially important in arid habitats. We propose that there is a close association between available soil moisture, the vigour of the potential host, and its susceptibility to infection by a root parasite (Watson, in press). This relationship was previously noted by Norton and Stafford Smith (1999) who studied mistletoe occurrence in mulga woodlands in central Australia, with roadside mulgas having higher water content in their leaves, and supporting greater numbers of mistletoes (see also Bowie & Ward 2004).

Using an alternative approach, Ehleringer (1989) measured carbon isotope composition to estimate long-term water stress of creosote *Larrea tridentata* in the Sonoran Desert. Those shrubs with lower water stress than their neighbours supported more mistletoes, and rather than relating to differences in dis-

perser behaviour, this pattern was interpreted as a reflection of differential mortality on hosts of different vigour (Ehleringer 1989). We suggest that many other parasitic plant systems may also reveal this sort of pattern, and we urge researchers first to estimate which factors may be limiting in their system (water, nutrients, light, etc.) and then measure this, both indirectly (via the host) and directly, to evaluate whether it plays a role in explaining parasite occurrence.

In this study, the O-ring statistic enabled us to differentiate two scales of clustering in *S. lanceolatum* and the scale of association between this parasite and its host *A. tetragonophylla* where Ripley's K could not. Analyses using Ripley's K suggested that clustering in *A. tetragonophylla* and *S. lanceolatum* occurred at all spatial scales examined, as did associations between host and parasite. This is not the case and comes about because small-scale second-order effects are translated to larger scales when using a cumulative statistic such as Ripley's K – 'virtual aggregation'. In systems where a realized distribution is the result of clustering processes operating at critical scales, a cumulative statistic may obscure clustering at those critical scales (Wiegand & Moloney 2004). Moreover, our results suggest using the non-accumulative O-ring statistic (or pair-correlation function) instead of Ripley's K for exploratory analysis because non-accumulative second order statistics have a direct interpretation in terms of the underlying processes and are not plagued by a possible bias introduced by effects at smaller scales. Ripley's K, however, is a valuable tool for testing goodness-of-fit to decide whether a given null model fits the data.

## ACKNOWLEDGEMENTS

We are grateful to the team of students who helped us tag, measure and map all 765 plants, to Martin Asmus for technical support and to Ingrid Witte and NSW NPWS for facilitating our research (Scientific Licence S10906). Detailed comments from the referees and associate editor clarified an earlier version of this manuscript. This study was supported by the Australian Research Council (DP0449417).

## REFERENCES

- Aukema J. E. & Martinez del Rio C. (2002a) Mistletoes as parasites and seed-dispersing birds as disease vectors: current understanding, challenges and opportunities. In: *Seed Dispersal and Frugivory; Ecology, Evolution and Conservation* (eds D. J. Levey, W. R. Silva & M. Galetti) pp. 99–110. CABI Publishing, Wallingford.
- Aukema J. E. & Martinez del Rio C. (2002b) Variation in mistletoes seed deposition: effects of intra and interspecific host characteristics. *Ecography* **25**, 139–44.

- Batista J. L. F. & Maguire D. A. (1998) Modelling the spatial structure of tropical forests. *For. Ecol. Manage.* **110**, 293–314.
- Besag J. (1977) Contribution to the discussion of Dr. Ripley's paper. *J. R. Stat. Soc. B* **39**, 193–5.
- Besag J. & Diggle P. J. (1977) Simple Monte Carlo tests for spatial pattern. *Appl. Stat.* **26**, 327–33.
- Bowie M. & Ward D. (2004) Water and nutrient status of the mistletoe *Phicosepalus acaciae* parasitic on isolated Negev Desert populations of *Acacia raddiana* differing in level of mortality. *J. Arid Environ.* **56**, 487–508.
- Brand J. E. & Jones P. J. (2002) The influence of landforms on sandalwood (*Santalum spicatum* (R.Br.) A.DC.) size structure and density in the north-eastern goldfields, Western Australia. *Rangel. J.* **24**, 219–26.
- Condit R., Ashton P. S., Baker P. *et al.* (2000) Spatial patterns in the distribution of tropical tree species. *Science* **288**, 1414–18.
- Cressie N. (1991) *Statistics for Spatial Data*. Wiley, New York.
- Dickman C. R., Pressey R. L., Lim L. & Parnaby H. E. (1993) Mammals of particular concern in the Western Division of New South Wales. *Biol. Conserv.* **65**, 219–48.
- Diggle P. J. (2003) *Statistical Analysis of Spatial Point Patterns*, 2nd edn. Academic Press, New York.
- Dixon P. M. (2002) Ripley's *K* function. *Encyclopedia Environ.* **3**, 1796–803.
- Ehleringer J. R. (1989) Carbon isotope ratios and physiological processes in aridland plants. In: *Stable Isotopes in Ecological Research* (eds P. W. Rundel, J. R. Ehleringer & K. A. Nagy) pp. 41–54. Springer-Verlag, New York.
- Ehleringer J. R. & Marshall J. D. (1995) Water relations. In: *Parasitic Plants* (eds J. D. Press & M. C. Graves) pp. 125–40. Chapman and Hall, London.
- Forde N. (1986) Relationships between birds and fruits in temperate Australia. In: *The Dynamic Partnership: Birds and Plants in Southern Australia* (eds H. A. Ford & D. C. Paton) pp. 42–58. Government Printer, Adelaide.
- Johnson G. (1996) *Action Statement No. 75. Northern Sandalwood, Santalum lanceolatum. Flora and Fauna Guarantee*. Department of Natural Resources and Environment, Melbourne.
- Keith D. A. (2004) *Ocean Shores to Desert Dunes: The Native Vegetation of NSW and the ACT*. DEC, Sydney.
- Murphy M. T., Garkaklis M. J., Hardy G. E. St. J. (2005) Seed caching by woylies *Bettongia penicillata* can increase sandalwood *Santalum spicatum* regeneration in Western Australia. *Austral Ecol.* **30**, 747–55.
- Nickrent D. L. (2002) Phylogenetic origins of parasitic plants. In: *Parasitic Plants of the Iberian Peninsula and Balearic Islands* (eds J. A. López-Sáez, P. Catalán & L. Sáez) pp. 29–56. Mundi-Prese, Madrid.
- Norton D. A. & Stafford Smith D. M. (1999) Why might roadside mulgas be better mistletoe hosts? *Aust. J. Ecol.* **24**, 193–8.
- Overton J. M. (1996) Spatial autocorrelation and dispersal in mistletoes: field and simulation results. *Vegetatio* **125**, 83–98.
- Parker C. & Riches C. R. (1993) *Parasitic Weeds of the World: Biology and Control*. CABI Publishing, Wallingford.
- Pate J. S., Davidson N. J., Kuo J. & Milburn J. A. (1990) Water relations of the root hemiparasite *Olax phyllanthi* (Labill) R. Br. (Olacaceae) and its multiple hosts. *Oecologia* **84**, 186–93.
- Pelissier R. & Goreaud F. (2001) A practical approach to the study of spatial structure in simple cases of heterogeneous vegetation. *J. Vég. Sci.* **12**, 99–108.
- Potts M. D., Davies S. J., Bossert W. H., Tan S. & Nur Supardi M. N. (2004) Habitat heterogeneity and niche structure of trees in two tropical rain forests. *Oecologia* **139**, 446–53.
- Radomiljac A. M. (1998) *Santalum Album L. (Indian sandalwood) plantation development: a complex interaction between parasite and host* (PhD Thesis). Murdoch University, Perth.
- Reid N., Stafford Smith D. M. & Yan Z. (1995) Ecology and population biology of mistletoes. In: *Forest Canopies* (eds M. D. Lowman & N. M. Nadkarni) pp. 285–310. Academic Press, San Diego.
- Ripley B. D. (1976) The second-order analysis of stationary point processes. *J. Appl. Probabil.* **13**, 255–66.
- Ripley B. D. (1981) *Spatial Statistics*. Wiley, New York.
- Sen-Sarma P. K. (1977) Sandalwood – its cultivation and utilization. In: *Cultivation and Utilization of Medicinal and Aromatic Plants* (eds C. K. Attal & B. M. Kapoor) pp. 287–97. RRL, Jammu.
- Shaw D. C., Watson D. M. & Mathiasen R. L. (2004) Comparison of dwarf mistletoes (*Arceuthobium* spp., Viscaceae) in the western United States with mistletoes (*Amyema* spp., Loranthaceae) in Australia – ecological analogs and reciprocal models for ecosystem management. *Aust. J. Bot.* **52**, 481–98.
- Stoyan D. & Stoyan H. (1994) *Fractals, Random Shapes and Point Fields: Methods of Geometrical Statistics*. Wiley, New York.
- Stoyan D. & Stoyan H. (1996) Estimating pair correlation functions of planar cluster processes. *Biom. J.* **38**, 259–71.
- Tennakoon K. U., Pate J. S. & Stewart G. R. (1997) Haustorium-related uptake and metabolism of host xylem solutes by the root hemiparasitic shrub *Santalum acuminatum* (R. Br.) A. DC. (Santalaceae). *Ann. Bot.* **80**, 257–64.
- Warburton C. L., James E. A., Frupp Y. J., Trueman S. J. & Wallace H. M. (2000) Clonality and sexual reproductive failure in remnant populations of *Santalum lanceolatum* (Santalaceae). *Biol. Conserv.* **96**, 45–54.
- Watson D. M. (2001) Mistletoe – a keystone resource in forests and woodlands worldwide. *Ann. Rev. Ecol. Syst.* **32**, 219–49.
- Watson D. M. (2004) Mistletoe: a unique constituent of canopies worldwide. In: *Forest Canopies*, edition 2nd edn (eds M. D. Lowman & H. B. Rinker) pp. 212–23. Elsevier, San Diego.
- Watson D. M. (in press) Determinants of parasitic plant distribution: the role of host quality. *J. For. Ecol. Manage.*
- Wiegand T. & Moloney K. A. (2004) Rings, circles, and null-models for point pattern analysis in ecology. *Oikos* **104**, 209–29.
- Wiegand T., Moloney K. A., Naves J. & Knauer F. (1999) Finding the missing link between landscape structure and population dynamics: a spatially explicit perspective. *Am. Nat.* **154**, 605–27.
- Woodall G. S. & Robinson C. J. (2003) Natural diversity of *Santalum spicatum* host species in south-coast river systems and their incorporation into profitable and biodiverse revegetation. *Aust. J. Bot.* **51**, 741–53.

## Appendix I

Our specific hypothesis required use of univariate point-processes with multiple and nested clustering. Examples for point processes which include an explicit clustering mechanism are Poisson cluster processes, Cox processes, or Gibbs processes (Diggle 2003), however, only a few of them have the advantage that the second-order statistics can be calculated analytically. We used the simplest family of cluster

processes which can be solved analytically – so-called Thomas processes – as the basic module. While univariate cluster processes have been used sporadically in ecological applications (e.g. Cressie 1991; Batista & Maguire 1998; Dixon 2002; Diggle 2003; Potts *et al.* 2004), double-cluster processes are rarely used (but see Diggle 2003; Stoyan & Stoyan 1996).

The univariate Thomas process assumes that: (i) parents follows a CSR process with intensity  $\rho$ ; (ii) each parent produces a random number of offspring following a Poisson distribution; and (iii) the locations of the offspring, relative to the parents, have a bivariate Gaussian distribution  $h(r, \sigma)$  with variance  $\sigma^2$ . The pair-correlation function  $g(r)$  of the Thomas process yields:

$$g(r, \sigma, \rho) = 1 + \frac{1}{\rho} \frac{\exp(-r^2/4\sigma^2)}{4\pi\sigma^2} \tag{A1}$$

The unknown parameters  $\rho$  and  $s$  can be fitted by comparing the empirical  $\hat{g}(r)$  and  $\hat{K}(r)$  with the theoretical functions (Diggle 2003).

The univariate nested double-cluster process arises if the CSR parents of the Thomas process are replaced by parents which also follow a Thomas process with large-scale clustering. Thus, this process is a multigeneration cluster process in which the offspring become the parents of the next generation (Diggle 2003). We calculated the pair-correlation function of the double-cluster process in two steps. In the first step, we calculated the bivariate pair-correlation function  $g_{12}(r)$  between the second-generation offspring (pattern 2) and its parents (pattern 1) which can be used to calculate the univariate pair-correlation function  $g_{22}(r)$  for the pattern of the second-generation offspring. For the following, we use the abbreviation  $h(r, \sigma)$  for a radial symmetric bivariate normal distribution with variance  $\sigma^2$ .

The calculation of the pair-correlation function  $g_{12}(r)$  is based on the second-order intensity function  $\lambda_2(r) = \lambda_1\lambda_2 g_{12}(r)$  which is the probability to find a point 1 at location  $l_1$  and a type 2 point at locations  $l_2$  which are a distance  $r = |l_1 - l_2|$  away. Because the two points may stem either from the same first-generation parent or from different first-generation parents, we need to sum the contribution of both cases. In the first case, the calculation is simple: we multiplied the intensity of type 1 points ( $=\lambda_1 = \rho_2$ ) with the expected number of offspring ( $\mu_2 = \lambda_2/\lambda_1$ ) with the probability that an offspring is distance  $r$  away from the type 1 point [ $=h(r, \sigma_2)$ ]. To derive the contribution of the second case, we needed to multiply the probability to find a second type 1 point at a location  $l_2$  which is a vector distance  $R$  from a first type 1 point [ $=\lambda_1^2 g_T(R, \sigma_1)$ ] with the expected number of offspring ( $\mu_2 = \lambda_2/\lambda_1$ ) with the probability that the type 2 offspring is a dis-

tance  $|R - r|$  away from the second type 1 point [ $=h(R - r, \sigma_2)$ ] and integrate over all possible vector distances  $R$ :

$$\begin{aligned} \lambda_{12}(r) &= \lambda_1\mu_2 h(r, \sigma_2) + \int \lambda_1\lambda_1 g_T(R, \sigma_1, \rho_1)\mu_2 h(R - r, \sigma_2) dR \\ &= \lambda_1 \left[ \frac{\lambda_2}{\rho_2} (h(r, \sigma_2) + \lambda_2 \int \left[ 1 + \frac{1}{\rho_1} h(R, \sqrt{2}\sigma_1) \right] \right. \\ &\quad \left. h(R - r, \sigma_2) dR \right] \tag{A2} \\ &= \lambda_1\lambda_2 \left[ \frac{1}{\rho_2} (h(r, \sigma_2) + \int h(R - r, \sigma_2) dR + \right. \\ &\quad \left. \int \frac{1}{\rho_1} h(R, \sqrt{2}\sigma_1) h(R - r, \sigma_2) dR \right] \\ &= \lambda_1\lambda_2 \left[ \frac{1}{\rho_2} h(r, \sigma_2) + 1 + \frac{1}{\rho_1} h(r, \sqrt{2\sigma_1^2 + \sigma_2^2}) \right] \end{aligned}$$

The second-order intensity function  $\lambda_{22}(r)$  of the univariate double-cluster follows directly from the formula for  $g_{12}(r)$ :

$$\lambda_{22}(r) = \int \lambda_2 g_{12}(R - r) h(R, \sigma_2) dR. \tag{A3}$$

Integrating and using Equation 2 yields:

$$\begin{aligned} g_{22}(r, \sigma_1, \rho_1, \sigma_2, \rho_2) &= 1 + \frac{1}{\rho_2} h(r, \sqrt{2}\sigma_2) + \\ &\quad \frac{1}{\rho_1} h(r, \sqrt{2\sigma_1^2 + 2\sigma_2^2}) \\ &= 1 + \underbrace{\frac{1}{\rho_2} \frac{\exp(-r^2/4\sigma_2^2)}{4\pi\sigma_2^2}}_{\text{small-scale clustering of offspring}} + \tag{A4} \\ &\quad \underbrace{\frac{1}{\rho_1} \frac{\exp(-r^2/4(\sigma_1^2 + \sigma_2^2))}{4\pi(\sigma_1^2 + \sigma_2^2)}}_{\text{interaction of the two cluster mechanisms}} \end{aligned}$$

Note that Equation 4 collapses back to a Thomas process (Eqn 1) if there is no large scale clustering (i.e.  $\sigma_1 \rightarrow \infty$ , or  $\rho_1 \rightarrow \infty$ ). Ripley's K follows directly from the  $g$ -function by integration:

$$K(r) = 2\pi \int_0^r g(r') r' dr'. \tag{A5}$$

Separation of the scales of clustering (i.e.  $\sigma_2^2 \ll \sigma_1^2$ ) suggested a convenient approach to fit the four parameters of a double-cluster process: as a first step we fitted the corresponding single-cluster model (Eqn 1), but only for scales  $r$  larger than the range of the small-scale clustering. We could do this because the contribution of small-scale clustering to the  $g$ -function (Eqn 4) disappears for larger scales. However, as the

range of the small-scale clustering was unknown *a priori*, we started with a small initial value and repeated the fitting procedure with successively increasing values until the single-cluster process provided a good fit of the data. In the second step, we used the parameters  $\sigma_1^2$  and  $\rho_1$  of the large scale clustering determined in the first step and fitted the two unknown parameters  $\sigma_2^2$  and  $\rho^2$  of the small-scale clustering using the full double-cluster model.

When fitting the formula of a cluster process to the data, we followed the minimal contrast method outlined in Diggle (2003), but used both the  $L$ - and the  $g$ -function simultaneously for the parameter fit by minimizing the geometric mean of the two contrasts.

Because the  $K$ -function is accumulative, it carries a ‘memory’ caused by a possible departure from the theoretical  $K$ -function up to scale  $r_0$  [ $K_0 = K(r_0) \neq \pi r_0^2$ ] to higher scales  $r > r_0$  (Wiegand & Moloney 2004). To prevent this memory from biasing estimation of the parameters, we used the transformed  $K$  and  $L$  functions for the fitting procedure:

$$K_{\text{unbiased}}(r, r \geq r_0) = K_0 - K(r_0) + K(r)$$

$$L_{\text{unbiased}}(r, r \geq r_0) = -r + \sqrt{\frac{K_0 - K(r_0) + K(r)}{\pi}} \quad (\text{A6})$$

$K_0$  is the observed value at scale  $r_0$  (i.e.  $\hat{K}(r_0) = K_0$ ), and  $K(r_0)$  is the theoretical value of the biases  $K$ -function at scale  $r_0$ .