An additive edge correction based on the refined Campbell theorem

Sharon Kühlmann-Berenzon[a]
Juha Heikkinen[b]
Aila Särkkä[a]

[a]Department of Mathematical Statistics
Chalmers University of Technology and Göteborg University
Gothenburg, Sweden
sharon@math.chalmers.se

[b]Finnish Forest Research Institute and Rolf Nevanlinna Institute
Helsinki, Finland

December 4, 2003

Summary: The influence potential is an index for measuring the effect of trees on understory vegetation observed, in this study, in a quadrat of a plot, and is defined as the sum of the effect of all trees in the plot. Since only the trees on the plot have been observed and not the trees outside the plot, the influence potential on a quadrat (IPQ) may be underestimated. Existing edge-corrections are not appropriate for this case. We propose a correction that consists of adding the expected IPQ outside the plot to the observed IPQ on the plot. The expectation is obtained by applying the refined Campbell theorem for stationary marked point processes. If a completely spatially random process is assumed, then the correction is based only on the size and location of the trees. Data from 1985-86 National Forest Inventory of Finland was used to calculate IPQ for six quadrats systematically allocated in 1240 plots. The implementation of the correction for this data is described. The estimates of IPQ with and without the correction proved the existence of edge-effects and the effectiveness of the correction to eliminate the bias. This method has the potential to be applied to other additive functions.

Keywords: edge-effects, forest ecology, influence potential, marked point process, National Forest Inventory, refined Campbell theorem.

1 Introduction

The influence potential of trees on a quadrat (IPQ) is an ecological index used to summarize the effect that trees have on a particular location (e.g. a quadrat in a plot). The effect of a tree is understood in terms of the resources that it may add or subtract, and therefore facilitate or hinder the growth of other plants in that location. IPQ has been previously applied to data from the permanent sample plots (PSP) of the 1985-86 National Forest Inventory (NFI) of Finland to model the presence of a species of understory vegetation (Kühlmann-Berenzon and Hjorth, 2003). The PSP consist of circular plots with quadrats assigned at specific positions inside the plot and are used for measuring the understory vegetation. We here consider the definition introduced by Kühlmann et al. (2001), where IPQ was obtained as the sum of the effects of the trees in the plot; and the effect of a tree was described as a function of the size of the tree and the relative position of the tree with respect to the quadrat. In the calculations, only the trees inside the plot are included in the sum. This means that the trees outside the plot are ignored, which may underestimate the true influence on the quadrat.

This type of problem is also encountered in the analysis of spatial point processes, when the process is observed through a delimited window. Then the information gathered on the events (i.e. the points of the point pattern, for example, trees) is censored by the boundaries of the window; this problem is commonly known as edge-effects. Several corrections have been developed for the G-function, F-function, K-function, and other statistics used in the field; see surveys on edge corrections in Ripley, 1982, 1988; Cressie, 1993; Stoyan, Kendall, and Mecke, 1995; and Kühlmann-Berenzon, 2002. Most of the methods correct for distances between an event and its nearest neighbor, or between an arbitrary point and its nearest event. Neither of these apply to IPQ where we require the distances between every event (tree) and an arbitrary point (quadrat); for these reasons an edge correction specific for IPQ was developed.

The edge correction is based on tools from the theory of marked point processes. The method is in an application of the refined Campbell theorem for stationary marked point processes, which takes advantage of the additive nature of IPQ. Although the correction conditions on the type of process, it is also shown that it is possible to ignore that condition, and that this is equivalent to assuming a completely spatially random (CSR) process.

In section 2 of this paper we define IPQ and illustrate the nature of the edge-effect problem, as well as introduce notations. Section 3 summarizes some useful concepts in the study of spatial point processes and marked point processes. The Campbell theorem and the refined Campbell theorem are stated in section 4. The proposed edge correction for IPQ is developed in section 5. Sections 6 and 7 describe, respectively, the implementation of the correction when applied to the PSP of the 1985-86 NFI of Finland, and compare the results obtained before and after the correction.

2 IPQ and edge-effects

The relationship between the trees and the understory vegetation has been studied using IPQ to summarize the effect that trees have on the vegetation (Kühlmann et al., 2001; Kühlmann-Berenzon and Hjorth, 2003); see Kuuluvainen and Pukkala (1989) who proposed the original version of this index. The version used here takes into consideration what information was collected from the PSP. The PSP consist of circular plots with a radius of 9.77m (area=300m^2) and distributed on a grid over Finland. In each plot, six quadrats were assigned systematically and utilized for measuring the vegetation. The quadrats were located at 3m (quadrats 2 and 3) and 8m (quadrats 1 and 4) north and south of the center of the plot and at 6m (quadrats 5 and 6) east and west of the center. Additionally all trees with diameter at breast height (DBH) greater than 10.5cm were measured and their location registered. Given this information, IPQ was defined as

$$\mathrm{IPQ}(q) = \sum_{z} d_z \ \exp\left(-\frac{t_z(q)^2}{c}\right); \tag{1}$$

here z indexes the trees; d is the DBH of the tree, $t_z(q)$ is the Euclidean distance between the tree z and the quadrat q; and c is the influence parameter which defines the distance from the quadrat at which a tree has a negligible effect on the quadrat. Figure 1 shows the configuration of the plot and the idea behind IPQ.

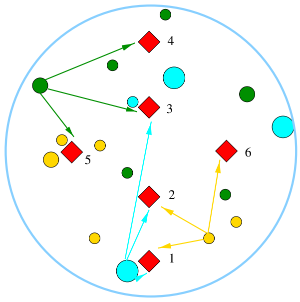


Figure 1: Tree effect in a plot of the PSP. The circles represent trees of different diameters, and the diamonds represent the quadrats with their identification numbers next to them. The arrows show schematically how every tree influences every quadrat; the effect will depend on the size of the tree and the distance to the quadrat. The effects of all trees are added up to produce a value of IPQ for each quadrat.

When IPQ is calculated from the PSP data, it is biased since only the trees observed inside the plot are included. This means that trees outside the plot, which might also affect the vegetation in the quadrat, are ignored. The result is an edge-effect and a possible underestimation of the true influence. Furthermore, the effect of a tree is weighted by its proximity to the quadrat: nearby trees weigh more heavily than those further away. Therefore the edge-effect problem is more pronounced the closer a quadrat is to the border of the plot, as more of the trees at shorter distances are missing.

Possible solutions from the field of spatial point processes include using a guard area or applying the border or toroidal methods (see e.g. Ripley, 1982). During the measuring campaign, however, no guard area was considered; additionally the plots are too small and do not have enough trees to be able to use the border method; and the toroidal principal is difficult to implement in circular plots. The case of the F-function is similar to the problem experienced in IPQ, since it measures the distance between an arbitrary point and the nearest event. In IPQ, however, we are not interested in only the nearest tree from the quadrat, but also other trees further away. The K-function takes into account events at other distances, but is only concerned on event to event distances, so corrections for this case are not appropriate for IPQ either. Moreover we wish to correct influence potential and not only the distance.

We therefore decided to develop a correction specifically for this problem, which would adjust the observed IPQ. The basic idea behind the proposed correction is to calculate the expected value of IPQ outside the plot where the trees were not measured. This expectation is then added to the observed IPQ to obtain the corrected IPQ; i.e.

$$\mathrm{IPQ}(q)^c = \mathrm{IPQ}(q)\{\text{observed}\} + E\big[\mathrm{IPQ}(q)\{\text{unobserved}\}\big]. \tag{2}$$

In theory IPQ should be calculated for an infinite area around the quadrat. For the purpose of the correction, however, we define a circle of influence $b(q,s)$ centered at the quadrat q and with radius s. We assume that the most important contributions to IPQ come from trees located in this circle. The radius s is obtained by defining first a minimum significant effect $\exp(-t_z(q)^2/c) = 0.01$. This means that only those trees with an effect larger than 0.01 are considered to be relevant. Assuming that the parameter c has been previously determined, then the radius of influence is

$$s = \sqrt{c\log(100)}.$$

Depending both on the position of the quadrat with respect to the border of the plot and on s, $b(q,s)$ might be completely contained inside the plot or not. If $b(q,s)$ is inside the plot, then we have all the information necessary for determining IPQ for that quadrat. If $b(q,s)$ is partly inside and partly outside the plot, we further denote those two areas as $I(q,s)$ and $O(q,s)$. More formally,

$$\begin{aligned} I(q,s) &= b(q,s) \cap b(o,r) \\ O(q,s) &= b(q,s) \setminus b(o,r), \end{aligned}$$

where $b(o,r)$ represents the plot centered at the origin and with radius r. Figure 2 illustrates all the necessary concepts. The problem of calculating a corrected IPQ in (2) can be defined now more precisely as

$$\mathrm{IPQ}(q)^c = \mathrm{IPQ}\big(q : z \in b(o,r)\big) + E\big[\mathrm{IPQ}\big(q : z \in O(q,s)\big)\big].$$

For simplicity, the term inside the expectation will be sometimes be written as $\mathrm{IPQ}\left(O(q,s)\right)$.

3 Spatial point processes

The spatial point processes $\Psi = \{\mathbf{x}_n\}$ we discuss here are restricted to $\mathbb{R}^2$ and assumed to have been observed in a window A. The location of an event $\mathbf{x} \in \Psi$ is determined by its coordinates (x_1, x_2). For the particular case of a stationary process, the intensity or mean number of points per unit area, λ, is constant throughout the area and can be estimated by $n/|A|$, i.e. the total number of events divided by the area of the observation window (see e.g. Diggle, 2003). In a CSR pattern, the number of events in A has a Poisson distribution

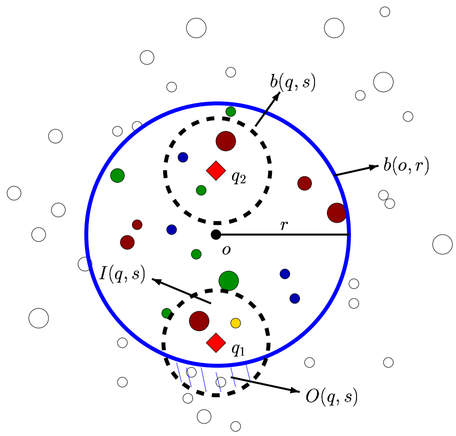


Figure 2: Edge correction concepts: The small circles represent trees: the filled ones have been observed, and the empty ones have not. The diamonds are the quadrats. The circle of influence $b(q, s)$ (dashed line) is of the same size for both quadrats. Since q_1 is closer to the plot border, its circle of influence extends beyond the borders and therefore a correction for IPQ is needed. The correction will estimate the expected value of IPQ in $O(q, s)$ (dashed area). The upper quadrat q_2 has a circle of influence completely inside the plot, thus all the information necessary for calculating IPQ has been measured and no correction is required.

with mean $\lambda|A|$, and the locations of the n events are distributed independently and uniformly in A.

A marked point process can be considered as a spatial point process with an additional dimension that contains information regarding each event (Stoyan et al., 1995). Typically the information may be continuous or discrete, e.g. diameter or species of a tree. The marked process is expressed as $\Psi = \{[\mathbf{x}_n; m_n]\}$, where $\mathbf{x}$ are the locations of the events and m are the corresponding marks. Stationarity is defined as for the unmarked case, i.e. it does not depend on the marks.

One of the statistics of interest in a marked point process is the mean mark $\bar{m}$,

$$\bar{m} = \int_{-\infty}^{\infty} m \, dF_M(m); \tag{3}$$

it depends on the mark distribution function F_M of the mark M, where F_M is estimated as

$$\hat{F_M}(v) = \frac{I\{m \le v : \mathbf{x} \in A\}}{n},$$

for a continuous variable v, and $I\{\cdot\}$ is the indicator function. The mark sum measure

$$S_m(A) = \sum_{[\mathbf{x};m]\in\Psi} m\, I\{\mathbf{x} \in A\}, \tag{4}$$

represents the sum of the marks for all events in A, and its expected value in the stationary case is $\lambda\bar{m}|A|$. It then follows that the estimate of the mean mark for a stationary marked point process observed in the window A is computed

$$\hat{\bar{m}} = \frac{S_m(A)}{n}. \tag{5}$$

The estimated mean mark $\hat{\bar{m}}$ is not always unbiased, but $\hat{\lambda}\hat{\bar{m}}$ is unbiased for $\lambda\bar{m}$ (Stoyan and Stoyan, 1994, p. 278).

4 Campbell theorem for spatial point processes

The *Campbell theorem* gives a way to calculate the expected sum of the values of a non-negative measurable function of a point process; see Stoyan and Stoyan (1994) and Stoyan et al. (1995) for the application to spatial point processes, and Kingman (1993) for the one-dimensional case and detailed proofs. The theorem for a stationary spatial point process can be seen as

$$E\Big[\sum_{\mathbf{x}\in\Psi} f(\mathbf{x})\Big] = \lambda \int f(\mathbf{x})\ d\mathbf{x}.$$

With a parallel formulation, the theorem can also be obtained for a function f of the events and marks in a stationary marked point process as (Stoyan and Stoyan, 1994; Stoyan et al., 1995)

$$E\Big[\sum_{[\mathbf{x};m]\in\Psi} f(\mathbf{x};m)\Big] = \lambda \int\int f(\mathbf{x},m)\ dF_M(m)\ d\mathbf{x}. \tag{6}$$

The *refined Cambell theorem* or *Campbell-Mecke theorem* computes expectations where the function depends not only on $\mathbf{x}$ but also on the whole realization; thus, for any non-negative measurable function h and a stationary point process,

$$\begin{aligned} E\Big[\sum_{\mathbf{x}\in\Psi} h(\mathbf{x};\Psi)\Big] &= \int \sum_{\mathbf{x}\in\psi} h(\mathbf{x},\psi)P(d\psi) \\ &= \lambda \int\int h(\mathbf{x},\psi_{\mathbf{x}})P_o(d\psi)d\mathbf{x}, \end{aligned} \tag{7}$$

where $P_o(d\psi)$ is the Palm distribution at the origin for the realization ψ (Stoyan et al., 1995). The Palm distribution is defined as

$$P_o(Y) = \int \sum_{x\in\psi\cap A} \frac{I\{\psi_{-x}\in Y\}}{\lambda|A|}\ P(d\psi),$$

where Y is a property of the process Ψ, and ψ_{-x} is the realization translated by x. The Palm distribution provides the probability that for a typical event of the process, ψ_{-x} holds the property Y. Then (7) can be interpreted as the expectation of the sum of a function given a typical point from the realization ψ.

The analogous refined Campbell theorem for a stationary marked point process is (Stoyan et al., 1995, p. 125)

$$\begin{aligned} E\Big[\sum_{[\mathbf{x};m]\in\Psi} h(\mathbf{x},m,\Psi)\Big] &= \\ &= \int \sum_{[\mathbf{x};m]\in\psi} h(\mathbf{x},m,\psi)P(d\psi) \\ &= \lambda \int\int\int h(\mathbf{x},m,\psi_{\mathbf{x}})\ P_o^m(d\psi)dF_M(m)d\mathbf{x}. \end{aligned} \tag{8}$$

Here P_o^m is the Palm distribution for a marked point process that allows conditioning on a point at the origin with mark m.

5 Edge correction for IPQ

To correct the edge-effects, we are interested in finding the expectation of IPQ in $O(q,s)$. We can consider the trees as being a realization of a marked point process. Furthermore, IPQ is defined as the sum of the effect of individual trees, and so IPQ can be thought of as the sum of a function of marked events. Then it follows that the Campbell theorem for marked point processes is an appropriate method for estimating the expectation of IPQ in $O(q,s)$. By plugging directly into (6) and assuming a stationary process,

$$E\big[\mathrm{IPQ}\big(O(q,s)\big)\big] = \lambda \int_{z\in O(q,s)} \int_{d} d\ \exp\bigg(-\frac{t_z(q)^2}{c}\bigg)\ dF_M(d)\ dz; \tag{9}$$

and F_M is the mark distribution function of DBH. This correction only takes into account the individual trees and not the relationship between the trees. It is likely, however, that the position of a tree will depend in part on the location of other nearby trees. This implies that information on the process that generated the tree pattern would be beneficial for the correction. In an indirect way, this is already taken into account in IPQ through the distance between the trees and the quadrat; the distribution of the distances between an event and an arbitrary point is known to differ among different types of spatial process; (see e.g. Diggle, 2003 and the F-function). But if the process can be identified, and thus more information is available, then the expectation of IPQ can be obtained with the refined Campbell theorem (8), such that

$$E\big[\mathrm{IPQ}\big(O(q,s)\big)\big] = \lambda \int_{z\in O(q,s)} \int_{d} \int_{\Psi} d\ \exp\bigg(-\frac{t_z(q)^2}{c}\bigg)\ P_z^d(d\psi)\, dF_M(d)\, dz. \tag{10}$$

Here the expectation is conditioned on a typical point of the marked point process located at z and with mark d. Even when the process is known, it may still be difficult to calculate the expectation. Nevertheless, in the particular case of a CSR process with independent marks, the Palm distribution coincides with the unconditional distribution, so $P_z^d(d\psi) = P(d\psi)$ (Stoyan et al., 1995). As a result, the integral for the Palm distribution can be separated from the rest of the formula, and

$$E\big[\mathrm{IPQ}\big(O(q,s)\big)\big] = \lambda \int\limits_{z\in O(q,s)} \int\limits_{d} d\exp\left(-\frac{t_z(q)^2}{c}\right) dF_M(d)\, dz \int\limits_{\Psi} P(d\psi).$$

The integral over Ψ is equal to one, and so the expectation simplifies to

$$E\big[\mathrm{IPQ}\big(O(q,s)\big)\big] = \lambda \int\limits_{z\in O(q,s)} \int\limits_{d} d\exp\left(-\frac{t_z(q)^2}{c}\right) dF_M(d)\, dz.$$

This equation is identical to (9) where no process was assumed. In other words, if the process is CSR, then the expectation does not use any information on the process itself. In other words, when applying the Campbell theorem, we indirectly assume that there are no interactions between the trees, i.e. that they area a realization of a CSR process.

Furthermore, all the terms in the previous equation are only functions of the events or of the marks, and therefore it is possible to separate the integrals as

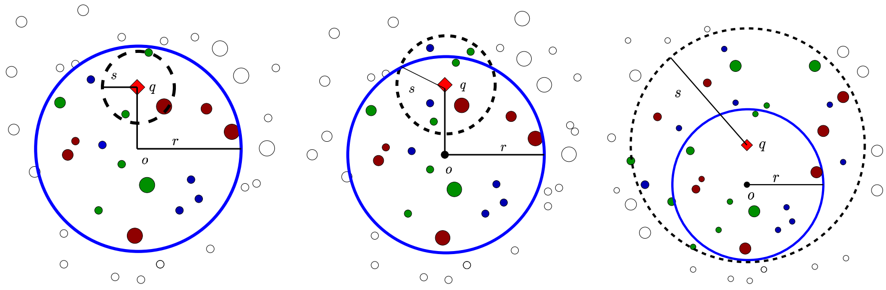
$$E\big[\mathrm{IPQ}\big(O(q,s)\big)\big] = \lambda \int\limits_{d} d\, dF_M(d) \int\limits_{z\in O(q,s)} \exp\left(-\frac{t_z(q)^2}{c}\right) dz.$$

The parameter λ can be estimated in the usual way with $\hat{\lambda} = n/|O(q,s)|$, where n is the number of trees in $O(q,s)$. The first integral represents $\bar{d}$, the mean mark of the DBH as in (3), and can be approximated with (5), namely $S_d(O(q,s))/n$. Since $O(q,s)$ has not been observed, n and $S_d(O(q,s))$ are unknown, but $\hat{\lambda}\bar{d}$ can be estimated from the observed trees in the plot. The estimate of the expected value of IPQ in $O(q,s)$ then becomes

$$\hat{E}\big[\mathrm{IPQ}\big(O(q,s)\big)\big] = \frac{\sum_{z\in b(o,r)} d_z}{|b(o,r)|} \int\limits_{z\in O(q,s)} \exp\left(-\frac{t_z(q)^2}{c}\right) dz. \tag{11}$$

Several advantages can be pointed out for the correction (10). It only requires that the process inside and outside of the plot be stationary. If information on the distribution of DBH is available, then a better estimate of $\bar{d}$ can be included in the correction. Furthermore, IPQ may be interesting for a specific group of events, e.g. tree species, and this method allows for the correction to be calculated for each group separately, using their individual $\hat{\lambda}$ and $\bar{d}$; this may be especially significant if the intensity or mean diameter differ greatly among the groups.

If the CSR assumption is not reasonable, then the marked point process needs to be modeled; studying groups of similar plots might facilitate this step. Once the process



(a) Case I (b) Case II (c) Case III

Figure 3: Cases I-III: The plot is represented by the large circle with solid line (center o and radius r), and the circle of influence by a dashed line (center q and radius s). The small circles represent trees: the filled ones are relevant trees, i.e. they are inside the circle of influence or the plot, and the empty ones are of no interest.

has been estimated, the expectation in (10) can also be estimated. Typically, this would be done by simulating the estimated marked point process a large number of times and calculating $\mathrm{IPQ}(O(q,s))$ for each simulation; then the estimate of the expected correction would be obtained from the average of the simulations.

6 Implementation

Three possible situations can occur in terms of $O(q,s)$ and $b(o,r)$, depending on the value of c that defines s:

I. $s < (r - \overline{oq})$;

II. $(r - \overline{oq}) < s < (r + \overline{oq})$;

III. $(r + \overline{oq}) < s$.

Figure 3 illustrates the three cases. Case I does not require any correction, because all the trees in $b(q,s)$ needed for calculating IPQ have been observed. In case II $O(q,s)$ has the shape of a crescent, and in case III that of an annulus of variable width. The main difference between cases II and III is how to determine the boundaries of $O(q,s)$. Here we present explicitly the calculations for case II, and the formulas for case III are included in the Appendix.

The correction in (11) expressed in polar coordinates is

$$\hat{E}[\mathrm{IPQ}\big(O(q,s)\big)] = \\ = \frac{\sum_{z\in b(o,r)} d_z}{|b(o,r)|} \int_{\theta(O(q,s))} \int_{u(O(q,s))} u \,\exp\left(-\frac{u^2}{c}\right) du\, d\theta,$$

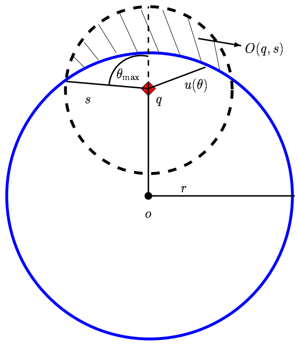


Figure 4: Geometry of the edge correction for case II. The angle $\theta_{\max}$ measures the intersection between the plot (centered at o and with radius r) and the circle of influence (centered at q and with radius s); $u(\theta)$ is the distance between the quadrat and the border of the plot at angle θ.

which makes it more convenient for defining the boundaries of $O(q, s)$. The limits of the integrals for case II are shown in detailed manner in figure 4. The angle $\theta_{\max}$ represents the angle of intersection between the circle of influence and the plot, and $u(\theta)$ is the distance between the quadrat and the edge of the plot at angle θ. It follows that the integral term of the correction is

$$\int_{-\theta_{\max}}^{\theta_{\max}} \int_{u(\theta)}^{s} u \exp\left(-\frac{u^2}{c}\right) du\, d\theta =$$
$$= \frac{c}{2} \int_{-\theta_{\max}}^{+\theta_{\max}} \exp\left(-\frac{u(\theta)^2}{c}\right) d\theta - c\,\theta_{\max} \exp\left(-\frac{s^2}{c}\right),$$

and the expectation of IPQ in $O(q, s)$ is therefore

$$\hat{E}\big[\mathrm{IPQ}\big(O(q,s)\big)\big] =$$
$$= \frac{\sum_{z \in b(o,r)} d_z}{300} \left\{ \frac{c}{2} \int_{-\theta_{\max}}^{+\theta_{\max}} \exp\left(-\frac{u(\theta)^2}{c}\right) d\theta - c\,\theta_{\max} \exp\left(-\frac{s^2}{c}\right) \right\}. \tag{12}$$

In this equation the sum of the diameters is computed over the trees in the plot, and $|b(o, r)|$, the area of the plot, is equal to $300\mathrm{m}^2$. The remaining integral must be solved numerically. Its limits are obtained by applying the law of cosines:

$$r^2 = u(\theta)^2 + \overline{oq}^2 - 2\,u(\theta)\,\overline{oq}\,\cos(\pi - \theta)$$

giving

$$u(\theta) = \sqrt{r^2 - \overline{oq}^2 \sin^2\theta} - \overline{oq}\cos(\theta); \tag{13}$$

and

$$\theta_{\max} = \arccos\left(\frac{r^2 - \overline{oq}^2 - s^2}{2\,\overline{oq}\;s}\right). \tag{14}$$

With (13) and (14), (12) is now expressed in terms of the known quantities $|b(o,r)|$, c, r, s, $\overline{oq}$, and $\sum_z d_z$. These are the same for all quadrats and plots, expect for $\overline{oq}$ that is specific for each quadrat, and the sum of the diameters that is different for each plot.

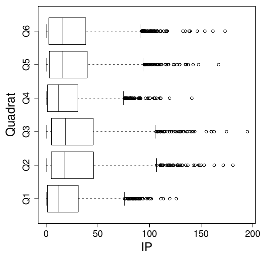
7 Application

IPQ was calculated for a subset of plots from the PSP comprising 1240 plots that contained one single tree stand and were located on mineral soils, and for each of the dominating tree species, Scots pine (*Pinus sylvestris*), Norway spruce (*Picea abies*), and birch (including *Betula pubescens* and *B. pendula*). By assuming that the influence potential of the species were independent, the observed IPQ (1) and the correction (11) were first calculated separately for each species T, and then added together to obtain one IPQ per quadrat. So the biased and the corrected IPQ were obtained from

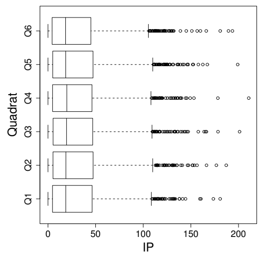
$$\begin{aligned}\mathrm{IPQ}(q) &= \sum_T \mathrm{IPQ}_T\big(q; z \in b(o,r)\big)\\ \mathrm{IPQ}(q)^c &= \sum_T \Big\{\mathrm{IPQ}_T\big(q : z \in b(o,r)\big) + E\big[\mathrm{IPQ}_T\big(q : z \in O(q,s)\big)\big]\Big\}.\end{aligned}$$

The results can be seen in figure 5, where boxplots for each of the six quadrats were drawn for IPQ and IPQ^c. Case II is illustrated with $c = 20$ and case III with $c = 75$. In figures 5(a) and 5(c), without the correction, the distributions of quadrats 1 and 4 concentrate on a smaller scale compared to the distributions of other quadrats (see figure 1 for location of quadrats in the plot). The reason is that those quadrats are positioned closest to the edge of the plot, so fewer trees are taken into account when calculating the influence potential. This bias is more subtle for quadrats 5 and 6 situated further inside the plot. Quadrats 2 and 3, close to the center, also suffer of edge-effects with these values of c, because the $b(q,s)$ extends beyond the plot borders. In this case, however, $O(q,s)$ is smaller than for the other quadrats, so the bias is smaller.

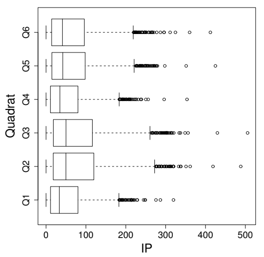
Figures 5(b) and 5(d) show the distributions when IPQ has been corrected by adding the expectation over $O(q,s)$. Although the correction cannot reproduce observations on the tail, it adds an important proportion of tree effects. The boxes are now almost perfectly aligned; this was anticipated, since the IPQ should be in the same range for all quadrats. Again we observe that the correction needed is larger the closer the quadrat is to the border. Furthermore the correction is also larger as c, and therefore $O(q,s)$, increases. These results confirm the presence of edge-effects and proves the effectiveness of the proposed correction to correct the bias.



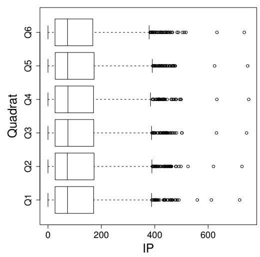
(a) IPQ: $c = 20$ ($s = 9.5$ m)



(b) IPQ^c: $c = 20$ ($s = 9.5$ m)



(c) IPQ: $c = 75$ ($s = 18.6$ m)



(d) IPQ^c: $c = 75$ ($s = 18.6$ m)

Figure 5: IPQ and IPQ^c by quadrat with c at 20 and 75 for case II (a and b) and case III (c and d).

8 Conclusions

The initial measurements of IPQ are biased due to edge-effects, because they only consider the trees inside the plot. Trees outside the borders, however, are likely to affect the vegetation in the quadrats, too, which means that the true IPQ is underestimated. Quadrats closer to the border were expected to have a larger bias as more trees at short distance are missing and therefore do not contribute to IPQ. This was confirmed by comparing the distributions of IPQ for the six quadrats, where the quadrats closer to the edge had a smaller range than those located towards the center. Edge corrections for spatial point patterns are available, but these do not take into account the mathematical definition of the effect of a tree.

The proposed correction consisted in estimating the expected IPQ outside the plot, which was then added to the observed IPQ. By taking the trees and their DBH as a stationary marked point process, the correction was based on the refined Campbell theorem. This means that the expectation of IPQ was conditioned on the realization, and therefore required the Palm distribution of the process. When a CSR process was assumed, the correction could be calculated by using only the locations and diameters of the trees. It is reasonable to believe, however, that the correction will be more precise if information on the process is included; e.g. Tomppo (1996) describes the spatial pattern of trees in different type of forests. Nevertheless, we allow the possibility that the process is not known, but in some sense we then assume CSR.

The correction integrated the exponential part of IPQ over $O(q,s)$, i.e. over the area of the circle of influence that lies outside the plot. The limits of the integral were defined by polar coordinates and simple geometrical properties such as the law of cosines. The quadrat as reference point for the integral limits was important, since the tree effect was measured according to the distance to the quadrat. This step was only required to be calculated once for each of the quadrats. When applied to the PSP data, the method was satisfactory in the sense that the bias observed in the original calculations of IPQ was eliminated.

The procedure we have introduced is based on a stationary process, and thus a different method should be applied in the non-stationary case. Although the assumption of CSR seems to be appropriate in our example, it does not always hold. It remains to be studied how much of the edge-effect remains or new bias is introduced when the process is not CSR, and on the other hand, how well the expectation can be estimated by simulations of the estimated model. The method, however, has the potential to be adapted to other plot shapes and to other definitions of influence potential as long as they can be described as a sum (cf. Kuuluvainen and Pukkala, 1989; Økland, Rydgren, and Økland, 1999).

Acknowledgements: The study was partially funded by the Academy of Finland (projects 47980 and 52303). Aila Särkkä has partly been supported by The Swedish Research Council and Swedish Foundation for Strategic Research.

9 Appendix: Case III

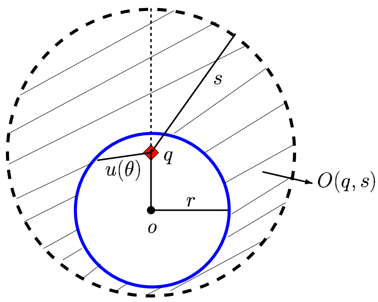


Figure 6: Case III: Geometry of the edge correction

$$\int_{z\in O(q,s)} \exp\left(-\frac{t_z(q)^2}{c}\right)dz =$$
$$= \int_{z\in b(q,s)} \exp\left(-\frac{t_z(q)^2}{c}\right)dz -$$
$$- \int_{z\in\{b(q,s)\cap b(o,r)\}} \exp\left(-\frac{t_z(q)^2}{c}\right)dz$$

$$\begin{aligned}\int_{z\in b(q,s)} \exp\left(-\frac{t_z(q)^2}{c}\right)dz &= \int_{-\pi}^{+\pi}\int_0^s u\exp\left(-\frac{u^2}{c}\right)dud\theta \\ &= \pi\, c\left[1-\exp\left(-\frac{s^2}{c}\right)\right]\end{aligned}$$

$$\begin{aligned}&\int_{z\in\{b(q,s)\cap b(o,r)\}} \exp\left(-\frac{t_z(q)^2}{c}\right)dz = \\ &= \int_0^{2\pi}\int_0^{u(\theta)} u\exp\left(-\frac{u^2}{c}\right)dud\theta \\ &= \pi\, c - \frac{c}{2}\int_0^{2\pi}\exp\left(-\frac{u(\theta)^2}{c}\right)d\theta\end{aligned}$$

$$\begin{aligned}r^2 &= u(\theta)^2 + \overline{oq}^2 - 2\,u(\theta)\,\overline{oq}\,\cos(\theta) \\ u(\theta) &= \overline{oq}\,\cos(\theta) - \sqrt{r^2 - \overline{oq}^2\,\sin(\theta)^2}\end{aligned}$$

References

[1] Cressie, N.A.C., 1993: Statistics for Spatial Data. Wiley, New York.

[2] Diggle, P.J., 2003: Statistical Analysis of Spatial Point Patterns, 2nd edition. Arnold, London.

[3] Kingman, J .F. C., 1993: Poisson Processes. Clarendon Press, Oxford.

[4] Kühlmann, S., Heikkinen, J., Särkkä, A., and Hjorth, U., 2001: Relating abundance of ground vegetation species and tree patterns using ecological field theory. In: K. Rennollls (ed.): Proc. of IUFRO 4.11 Conference: Forest Biometry, Modelling and Information Science. University of Greenwich, Greenwich, UK, cms1.gre.ac.uk/conferences/iufro/proceedings/.

[5] Kühlmann-Berenzon, S., 2002: Tree influence on understory vegetation: an edge correction and a conditional model, Technical report No. 2002:12. Chalmers University of Technology, Göteborg.

[6] Kühlmann-Berenzon, S., and Hjorth, U., 2003: Accounting for large-scale effects in the study of understory vegetation using a conditional logistic model (submitted).

[7] Kuuluvainen, T., and Pukkala, T., 1989: Effect of Scots pine seed trees on the density of ground vegetation and tree seedlings. Silva Fennica 23, 159-167.

[8] Økland, R., Rydgren, K., and Økland, T., 1999: Single-tree influence of understorey vegetation in a Norwegian boreal spruce forest. OIKOS 87, 488-498.

[9] Ripley, B. D., 1982: Edge effects in spatial stochastic processes. In: B. Ranneby (ed.): Statistics in Theory and Practice. Essays in Honour of Bertil Matérn, Swedish University of Agricultural Sciences, Section of Forest Biometry, Umeå, 247-262.

[10] Ripley, B. D., 1988: Statistical Inference for Spatial Processes. Cambridge University Press, Cambridge.

[11] Stoyan, D., Kendall, W. S., and Mecke, J., 1995: Stochastic Geometry and its Applications, 2nd edition. Wiley, Chichester.

[12] Stoyan, D., and Stoyan, H., 1994: Fractals, Random Shapes and Point Fields. Methods for Geometrical Statistics. Wiley, Chichester.

[13] Tomppo, E.: Models and methods for analysing spatial patterns of trees. Communicationes Instituti Forestalis Fenniae 138, 1-65.