
Analysis of Spatial Point Patterns in Microscopic and Macroscopic Biological Image Data

Frank Fleischer¹, Michael Beil², Marian Kazda³, and Volker Schmidt⁴

¹ Department of Applied Information Processing and Department of Stochastics,
University of Ulm, D-89069 Ulm, Germany
frank.fleischer@mathematik.uni-ulm.de

² Department of Internal Medicine I, University Hospital Ulm, D-89070 Ulm,
Germany

³ Department of Systematic Botany and Ecology, University of Ulm, D-89069
Ulm, Germany

⁴ Department of Stochastics, University of Ulm, D-89069 Ulm, Germany

1 Introduction

The analysis of spatial point patterns by means of estimated point process characteristics like for example Ripley's K-function ([37]), the L-function or Baddeley's J-function ([7], [24]) has proven to be a very useful tool in Stochastic Geometry during the last years (see e.g. [10], [38], [43], and [44]). They offer the possibility to get not only qualitative knowledge about the observed spatial structures of such point patterns, but to quantify them for specific regions of point-pair distances. Other advantages of these methods compared to alternative techniques of spatial analysis like e.g. Voronoi tessellations ([26], [31]) or minimum spanning trees ([13], [14]) is their independence of underlying point process intensities, in other words of the average number of points per unit square. Therefore they are especially useful for cases of data with significant differences with respect to intensities. We will discuss two examples in the fields of cell biology and ecology where these methods can be applied. They have been chosen, because they demonstrate the wide range of applications for the described techniques and because both examples have specific interesting properties. While the point patterns regarded in the first application are three dimensional, the second application reveals planar point patterns having a vertically inhomogeneous structure. For a more extensive description of the studied cases, the reader is referred to [4], [5], and [16], [41], [42], respectively.

1.1 Analysis of Centromeric Heterochromatin Structures

The first example deals with the structure of the cell nucleus, notably the distribution of centromeres, during differentiation (maturation) of myeloid cells. These are the precursors of white blood cells and are normally found in the bone marrow. During differentiation, myeloid cells acquire specialized functions by activating a strictly defined set of genes and producing new proteins. In addition, other genes whose function are not needed in matured cells become silenced. The mechanisms regulating the activity of genes during differentiation remain to be defined in detail.

The architecture of the cell nucleus during the interphase, i.e. the time between consecutive cell divisions, is determined by the packaging of the DNA molecule at various levels of organisation (chromatin structure). The open state of DNA is referred to as euchromatin, whereas heterochromatin is the condensed form of DNA. The production of gene transcripts (mRNA) requires the molecules of the transcriptional machinery to access the DNA molecule. Thus, the regulation of DNA packaging represents an important process for controlling gene activity, i.e. the synthesis of mRNA ([8]). In general, transcriptional activity appears to be impeded by a restrictive (compacted) packaging of DNA ([35]). This way, the heterochromatin compartment is an important regulator of gene transcription and, hence, influences the biological function of cells.

There has been a great interest in investigating the processes governing the organisation of chromatin. Whereas the regulation at the level of nucleosomes, e.g. through biochemical modifications of histones, is now becoming elucidated, the long-range remodelling of large portions of the DNA molecule (higher-order chromatin structure) proceeds by yet unknown rules. The condensed form of DNA, i.e. heterochromatin, is generally associated with the telomeres and centromeres of chromosomes ([36]). These regions can also induce transcriptional repression of nearby genes ([35]). Consequently, the distribution of these regions should be changed during cellular differentiation that is associated with a marked alteration of the profile of activated genes. In fact, previous studies described a progressive clustering of interphase centromeres during cellular differentiation of lymphocytes and Purkinje neurons ([1], [32]). However, the overall structural characteristics of the centromeric heterochromatin compartment, e.g. with respect to spatial randomness, remain to be determined.

Leukaemias are malignant neoplasias of white blood cells. They represent an interesting biological model to study cellular differentiation since they can develop at every level of myeloid differentiation. A particular type, acute promyelocytic leukaemia (APL), is characterized by a unique chromosomal translocation fusing the PML gene on chromosome 15 with the gene of the retinoic acid receptor alpha on chromosome 17 ([47]). Due to the function

of the resulting fusion protein, cellular differentiation is arrested at the level of promyelocytes. However, pharmacological doses of all-trans retinoic acid (ATRA) can induce further differentiation of these cells along the neutrophil pathway ([15]).

In a recent study, the three-dimensional (3D) structure of the centromeric heterochromatin was studied in the NB4 cell line which was established from a patient with APL ([23]). The 3D positions of centromeres served as a surrogate marker for the structure of centromeric heterochromatin. Due to the diffraction-limited resolution of optical microscopy, the notion chromocenter was used to define clusters of centromeres with a distance below the limit of optical resolution. During differentiation of NB4 cells as induced by ATRA, a progressive clustering of centromeres was implied from a decreased number of detectable chromocenters. The 3D distribution of chromocenters was evaluated by analysing the minimal spanning tree (MST) constructed from the 3D coordinates of the chromocenters. The results obtained by this method suggested that a large-scale remodelling of higher order chromatin structure occurs during differentiation of NB4 cells.

1.2 Planar Sections of Root Systems in Tree Stands

In our second example we examine the spatial distribution of tree root patterns in pure stands of Norway Spruce (*Picea Abies*) and European Beech (*Fagus sylvatica* (L.) Karst.). While there exists knowledge about the vertical root distribution which can usually be described by one-dimensional depth functions ([20], [34]), early studies assumed that rooting zones are completely and almost homogeneously exploited by roots ([22]). In recent studies it is however shown that fine roots concentrate in distinct soil patches and that they proliferate into zones of nutrient enrichment and water availability ([6], [39]). Hence a horizontal heterogeneity of the spatial distribution of fine roots might be expected. Trench soil profile walls can be used for the assessment of two-dimensional root distributions, regarding the roots on the wall as points of different diameter. A new method provided (x,y) -coordinates of each root greater than 2 mm ([41]). Using this method, small roots with a diameter between 2 mm and 5 mm were examined in 19 pits on altogether 72 m² of soil profiles on monospecific stands of European Beech and of Norway Spruce.

1.3 Detection of Structural Differences

So to summarize, our aims in both studied applications were quite similar from a mathematical point of view. First, as it is generally the case in such a spatial analysis, to compare our data sets with the *null* model of the homogeneous Poisson process, otherwise described as 'complete spatial randomness' (CSR) ([10]). If such a hypothesis can be rejected, a goal is to detect structural differences between the two regarded groups in the data, namely nuclei from

non-differentiated and differentiated NB4 cells in the first application and beech roots and spruce roots in the second one. For the cell nuclei example we were especially interested in an explanation for the decreasing number of chromocenters during differentiation, while for the tree roots a main aim was to quantify the degree of root aggregation, i.e. the degree of intensity of the exploitation of the soil resources by each tree species. Another question of interest is to provide a suitable and not too complicated mathematical model for underlying generating point processes. Finally, of course it is a necessity to obtain an interpretation of the results from the biological standpoint.

2 Image Data

As it has been mentioned in Section 1, there were two different data sets considered, three dimensional point patterns in cell nuclei of a NB4 cell line and two dimensional point patterns in profile walls of European Beech and Norway Spruce.

2.1 NB4 Cell Nuclei

The procedures for cell culture of NB4 cells, specimen preparation, immunofluorescence confocal microscopy and image analysis are described in detail in [4]. In the following two paragraphs the applied techniques are summarized.

Sample Preparation and Image Acquisition

Differentiation of NB4 cells was induced by incubating cells with 5 *mol/l* ATRA (Sigma, St.Louis, MO) for 4 days. Visualization of centromeres was based on immunofluorescence staining of centromere-associated proteins with CREST serum (Euroimmun Corp., Gross Groenau, Germany). Nuclear DNA was stained with YoPro-3 (Molecular Probes). Two channel acquisition of 3D images was performed by confocal scanning laser microscopy (voxel size: 98 *nm* in lateral and 168 *nm* in axial direction).

Image Segmentation

Segmentation of chromocenters as stained by CREST serum was performed in two steps. First, objects at each confocal section were segmented by edge detection followed by a conglomerate cutting procedure. In a second step, 3D chromocenters were reconstructed by analyzing series of 2D profiles. The center of gravity was used to define the 3D coordinates for each chromocenter. The final analysis included 28 cell nuclei from untreated controls with 68 chromocenters on average and 27 cell nuclei from ATRA-differentiated NB4 cells with 57 chromocenters on average (see Figure 1).

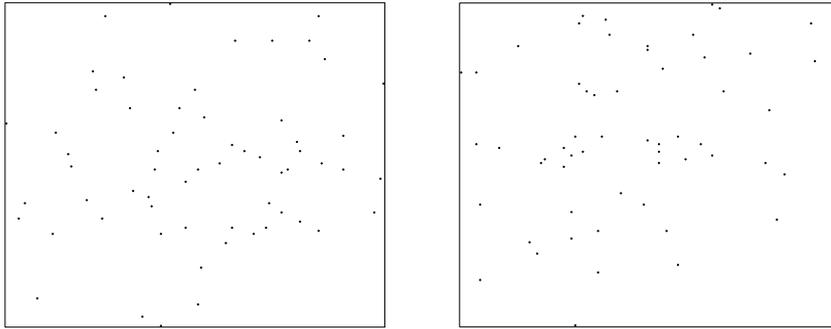


Fig. 1. Projections of the three dimensional chromocenter location patterns of an undifferentiated NB4 cell (left) and a differentiated NB4 cell (right) onto the xy -plane

2.2 Profile Walls

For details of site description, pit excavation and root mapping and already attained results, see [41]. Our investigations are based upon this article and thus only a short summary of the most important facts is given.

Site Description

Data collection took place near Wilhelmsburg, Austria ($4805'51''N$, $1539'48''E$) in adjoining pure stands of *Fagus sylvatica* and of planted *Picea abies*. One experimental plot of about 0.5 ha was selected within each stand. The sites were similar in aspect (NNE), inclination (10%) and altitude (480 m). The characteristics of the spruce and beech stands, e.g. the age (55 and 65 years), the dominant tree height (27 m and 28 m) and the stand density (57.3 and 46.6 trees /ha), also were similar to each other. The soils with only thin organic layer (about 4 cm) can be classified as Stagnic Cambisols developed from Flysch sediments. Annual rainfall in Wilhelmsburg averages 843 mm with a mean summer precipitation from May to September of 433 mm . The mean annual temperature is 8.4° C , and the mean summer temperature is 15.7° C .

Pit Excavation and Root Mapping

In every stand 10 soil pits with a size of $2 \times 1 \text{ m}$ were excavated. Thus up to 20 profile walls could be obtained in each stand. In most cases 13 – 19 trees were within a radius of 10 m around the pit centre. The minimum distance from the pit centre to the nearest tree ranged from 0.5 m to 2.8 m . On each wall all coarse roots were identified and divided into living and dead. All living small roots ($2 - 5 \text{ mm}$) were marked with pins and digitally photographed. These pictures were evaluated and a coordinate plane was drawn over each profile wall W , so that every root corresponds to a point x_n in the plane. Thus, for each profile wall W , a point pattern $\{x_n\} \subset W$ of root locations

was determined.

Data Description

Root mapping was performed on 20 profile walls of *Fagus sylvatica* and on 16 profile walls of *Picea abies* (see Figure 3). The profile walls B with area $\nu_d(B) = 200 \text{ cm} \times 100 \text{ cm}$ are regarded as sampling windows of stochastic point processes in \mathbb{R}^2 .

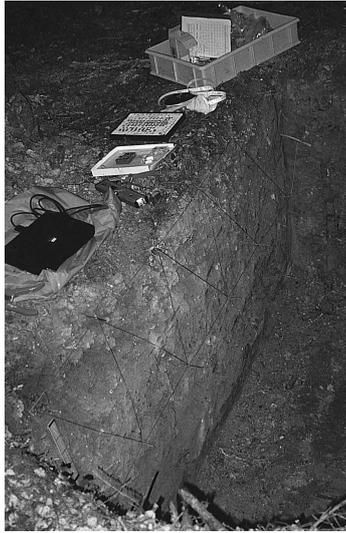


Fig. 2. Data collection

3 Statistical Methods and Results

Data analysis for all data groups was performed using the GeoStoch library system. GeoStoch is a Java-based open-library system developed by the Department of Applied Information Processing and the Department of Stochastics of the University of Ulm which can be used for stochastic-geometric modelling and spatial statistical analysis of image data ([27],[28], <http://www.geostoch.de>).

For both study cases it is important to notice that, considering estimated point process characteristics, means for each group were regarded. This is due to the fact that variability inside a single group (European Beech and Norway Spruce or NB4 cell nuclei, respectively) was large compared to the differences between the two groups for each studied case (tree roots and cell nuclei). For functions, these means were taken in a pointwise sense.

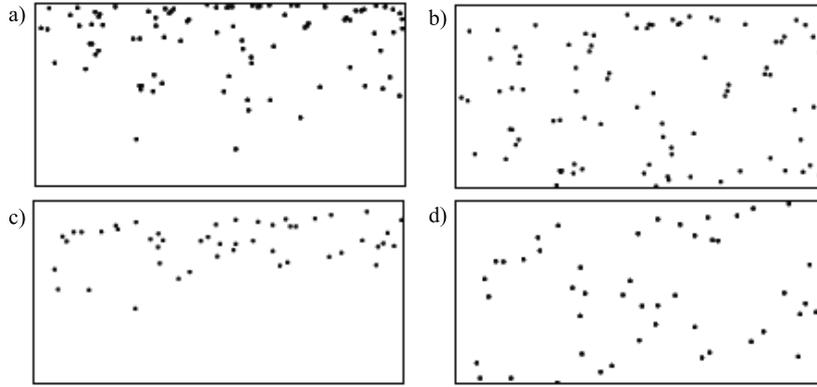


Fig. 3. a) an original sample of roots for *Picea abies* b) the transformed sample for *Picea abies* c) an original sample of roots for *Fagus sylvatica* d) the transformed sample for *Fagus sylvatica*

3.1 3D Point Patterns of Chromocenters in NB4 Cells

The real sampling regions for the cell nuclei are not known, therefore assumed sampling regions were constructed as follows: For all three coordinates the smallest and largest values appearing in a sample were determined and denoted as x_{min} , x_{max} , y_{min} , y_{max} , z_{min} and z_{max} respectively. Then the 8 vertices of the assumed sampling cuboid were given by all possible combinations of the three coordinate pairs $\{x_{min}, x_{max}\}$, $\{y_{min}, y_{max}\}$ and $\{z_{min}, z_{max}\}$. Although we performed statistical tests on stationarity and isotropy of the regarded point fields which showed results in favor of such assumptions, we would like to consider stationarity and isotropy as prior assumptions that are not under investigation. This is due to the fact that the numbers of points per sample do not seem to be large enough to provide reliable information on both properties, especially with regard to the three dimensionality of the data. Therefore formal tests can be only hints that such assumptions might not be badly chosen.

Intensities and Volumes

It is already known from [4] that the average number of detected chromocenters was significantly decreased during differentiation. Regarding the volumes of the assumed sampling cuboids, the hypothesis of having the same volume before and after differentiation could not be rejected ($\alpha = 0.05$), observing mean volumes of $429.507 \mu m^3$ before differentiation and $470.929 \mu m^3$ afterwards. Hence the intensity estimate $\hat{\lambda}$ of detected chromocenters of NB4 cells, that means the average number per unit volume (see Formula (3) in Appendix A), is significantly decreased as well.

Averaged Estimated Pair Correlation Function

Figure 4 shows estimations $\hat{g}(r)$ of the pair correlation function $g(r)$ for $c = 0.06$, where c determines the bandwidth of the Epanechnikov kernel used in the definition of $\hat{g}(r)$; see Formula (11) in Appendix A. It is clearly visible that the frequency of point-pair distances for a distance between 350 nm and 800 nm is higher before than after differentiation of NB4 cells. Also a hardcore distance r_0 of about 350 nm can be recognized which is determined by the diffraction-limited spatial resolution of the microscopic imaging method. This means that all point pairs have a distance bigger than r_0 . Note that the smaller hardcore values for larger values of c are due to the increased bandwidths of the Epanechnikov kernel in these cases. The results for the estimated pair correlation functions do not depend on the fact that the two groups have different numbers of detectable chromocenters.

Averaged Estimated L-Function

We consider the estimator $\hat{L}(r)$ for $L(r)$ given in Formula (17) of Appendix A. Figure 5 shows the estimated averaged L-function $\hat{L}(r)$ while Figure 6 shows $\hat{L}(r) - r$ where the theoretical value r for Poisson point processes has been subtracted; see Formula (16) in Appendix A.

A similar scenario as for the pair correlation function is observed. Especially for small point-pair distances between 350 nm and 500 nm, there is a higher

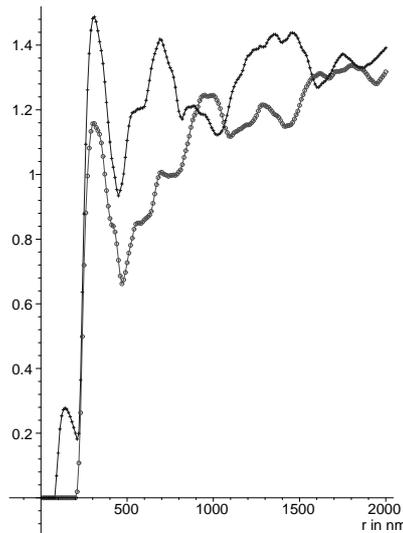


Fig. 4. Averaged estimated pair correlation functions $\hat{g}(r)$ using Epanechnikov kernel and parameter $c = 0.06$. The group of undifferentiated NB4 cells is denoted by +, while the group of differentiated NB4 cells is denoted by o

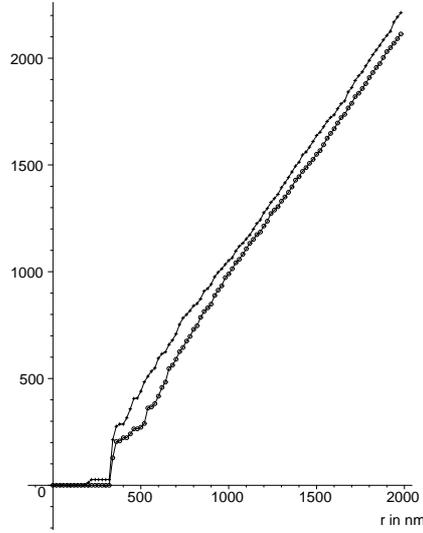


Fig. 5. Averaged estimated functions $\widehat{L}(r)$, where + denotes the group of undifferentiated NB4 cells and o denotes the group of NB4 differentiated cells

percentage of point pairs before than after ATRA-induced differentiation of NB4 cells. While for the group of undifferentiated cells the graph $\widehat{L}(r) - r$ has a mostly positive slope in this region, which is an indicator for attraction, the group of differentiated NB4 cells shows a negative slope which is a sign for rejection. The same hardcore distance $r_0 \approx 350 \text{ nm}$ is visible. Again the results do not depend on the different numbers of detectable chromocenters. Performing a Wilcoxon-Mann-Whitney test for the two group samples for fixed radii shows a significant difference in the values of L-functions before and after differentiation for all radii between 350 nm and 1300 nm , especially for the region between 500 nm and 700 nm ($\alpha = 0.05$).

Averaged Estimated Nearest-Neighbor Distance Distribution and Averaged Estimated J-Function

The structural conclusions obtained from the results for the estimated point field characteristics $\widehat{D}_H(r)$ and $\widehat{J}(r)$ were very similar compared to the averaged estimated pair correlation function $\widehat{g}(r)$ and the averaged estimated L-Function $\widehat{L}(r)$, where $\widehat{D}_H(r)$ and $\widehat{J}(r)$ are given by Formulae (18) and (22) in Appendix A. Therefore the averaged estimates $\widehat{D}_H(r)$ and $\widehat{J}(r)$ are not displayed here. Again a hardcore distance of 350 nm can be recognized and the two different groups show strong differences in their behavior especially for a range between 350 nm and about 800 nm .

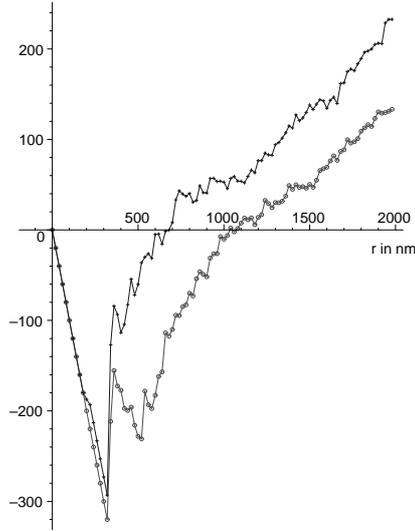


Fig. 6. Averaged estimated functions $\widehat{L}(r) - r$, where + denotes the group of NB4 undifferentiated cells and o denotes the group of NB4 differentiated cells

3.2 2D Point Patterns in Planar Sections of Root Systems

From [41] it was already known that the depth densities of the roots of Norwegian Spruce and European Beech can be approximated by exponential and gamma distributions respectively. The data has been homogenized with respect to the vertical axis in order to allow the assumptions of stationarity and isotropy for models of generating point processes. A suitable homogenization can be based on the well-known fact that each random variable Y with a continuous distribution function F_Y can be transformed to a uniformly distributed random variable U on the interval $[0, 1]$ by

$$U = F_Y(Y). \quad (1)$$

Therefore, by denoting the original depths, the total depth of the sampling window and the transformed depths as h_{orig} , h_{tot} and h_{tran} respectively, we get

$$h_{tran} = \frac{F^*(h_{orig})}{F^*(h_{tot})} h_{tot}, \quad (2)$$

where $F^*(x)$ symbolizes the suitable distribution function, i.e. the exponential distribution function in the case of Norway Spruce and the gamma distribution function in the case of European Beech. The total depth was given as $h_{tot} = 100 \text{ cm}$. For each sampling window parameters of the distribution functions $F^*(x)$ are estimated individually using maximum-likelihood estimators.

Notice that in the following, first only vertically homogenized data is regarded (see Figure 3), that means considering the vertical coordinate a uniform distribution on $[0, h_{tot}]$ can be assumed. Later on, an inverse transformation is applied to obtain inference for the original data.

Intensities

The average number of points for the samples of *Picea abies* is significantly higher than for the samples of *Fagus sylvatica* ($\alpha = 0.05$). Since sampling windows have the same sizes, the same result is obtained regarding the estimated intensities per cm^2 ($\hat{\lambda}^{spruce} = 0.00403$ vs. $\hat{\lambda}^{beech} = 0.00262$). Notice that the following results for the considered point process characteristics are independent of this fact since the functions are scaled with respect to the intensities.

Isotropy and Complete Spatial Randomness

Isotropy was tested by determining the directional distribution of the angles of point pairs to the axes in a quadratic sampling window and testing them for uniform distribution. The hypothesis of isotropy could not be rejected ($\alpha = 0.05$), hence in the following isotropy is assumed. The quadrat count method ([44]) was used to test on complete spatial randomness. Here, using a 4×4 grid, the hypothesis that the given point patterns are extracts of realizations of homogeneous Poisson processes was rejected ($\alpha = 0.05$).

Averaged Estimated J-Function

In Figure 7 the averaged estimated J-functions $\hat{J}(r)$ for both groups are displayed. There is a clear indication for attraction between point pairs of a distance less than 12 *cm*, since both functions are below 1 in this region and have a negative slope. A second observation is that the graph of *Picea abies* lies beneath the graph of *Fagus sylvatica*, which means that the point pairs of spruces are more attracted to each other than the point pairs of beeches for such distances. Notice that for radii larger than 20 *cm* the estimator becomes numerically unstable, and therefore should not be taken into further consideration. Also one should keep in mind that the J-function is a cumulative quantity.

Averaged Estimated L-Function

In Figure 8 the graphs for the averaged estimated values of $\hat{L}(r) - r$ are shown. Since in the Poisson case $L(r) \equiv r$ a positive slope means that there is an attraction, while a negative slope indicates repulsion. Again there are signs of attraction for small point-pair distances, less than 9.5 *cm* and less than 13.5 *cm* respectively, and the attraction seems to be stronger for *Picea abies* compared to *Fagus sylvatica* since the slope of $\hat{L}(r) - r$ is bigger. The negative values for very small distances might indicate a slight hardcore effect between the points.

Averaged Estimated Pair Correlation Function

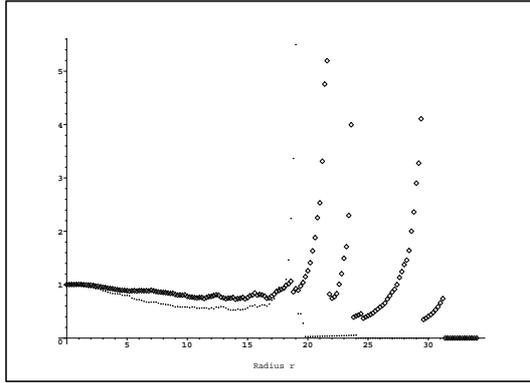


Fig. 7. Averaged estimated J-functions for *Picea abies* (·) and *Fagus sylvatica* (◇)

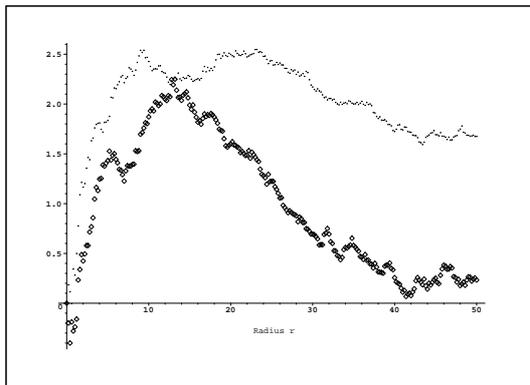


Fig. 8. Averaged estimated functions $\widehat{L}(r) - r$ for *Picea abies* (·) and *Fagus sylvatica* (◇)

Further indication for an attraction between point pairs of distances less than 14 cm is provided by the averaged estimated pair correlation functions $\widehat{g}(r)$ displayed in Figure 9. Again a stronger attraction is observed for the spruces since the function runs above the function for beeches for r less than 9 cm. Both functions are above 1 for $r < 14$ cm.

3.3 Model Fitting for the Root Data

Homogeneous Matern-Cluster Model

Regarding the results of the estimated point process characteristics described before and because of its simplicity, Matern-cluster processes are chosen as a model for the underlying point processes; see Appendix B for a definition. Using once more the natural intensity estimator given in Formula (3) of Appendix A, the parameter λ_{mc} was estimated as $\widehat{\lambda}_{mc}^{spruce} = 0.00403$ and

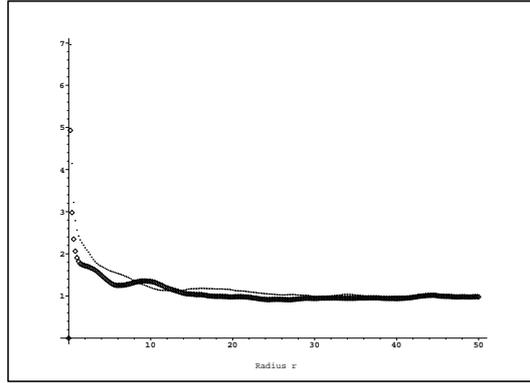


Fig. 9. Averaged estimated pair correlation functions $\widehat{g}(r)$ for *Picea abies* (\cdot) and *Fagus sylvatica* (\diamond), estimated using Epanechnikov kernel with parameter $c = 0.15$

$\widehat{\lambda}_{mc}^{beech} = 0.00262$, respectively. Concerning the regarded point-pair distances a range from 0 cm to $r_{max} = 50\text{ cm}$ has been chosen, where r_{max} equals half the minimum of the given depth and width. The parameters R and λ_e are estimated by minimum-contrast estimators, which are computed by numerical minimization of the integral

$$\int_0^{r_{max}} (\widehat{g}(r) - g_{theo}(r))^2 dr, \quad (3)$$

where $\widehat{g}(r)$ is the averaged estimated pair correlation function given in Formula (11) of Appendix A for the Epanechnikov kernel with $c = 0.15$, and $g_{theo}(r)$ is the theoretical value for the pair correlation function of the Matern-cluster process with parameters λ_e , R and λ_{mc} . Since the parameter λ_{mc} has already been estimated, the minimization of the integral in (3) yields an estimation for the pair of parameters R and λ_e . The obtained estimates are $\widehat{R}^{spruce} = 4.9\text{ cm}$ and $\widehat{\lambda}_e^{spruce} = 0.00690$ for spruce roots, while for beech roots $\widehat{R}^{beech} = 7.4\text{ cm}$ and $\widehat{\lambda}_e^{beech} = 0.00603$ are obtained.

Model Conclusions

The given point patterns are modeled as extracts of realizations of stationary Matern-cluster processes with estimated intensities $\widehat{\lambda}_{mc}^{spruce} = 0.00403$ and $\widehat{\lambda}_{mc}^{beech} = 0.00262$, with cluster radii $\widehat{R}^{spruce} = 4.9\text{ cm}$ and $\widehat{R}^{beech} = 7.4\text{ cm}$, and with parent-process intensities $\widehat{\lambda}_e^{spruce} = 0.00690$ and $\widehat{\lambda}_e^{beech} = 0.00603$. In order to get an idea for the degree of clustering, the quantity

$$\widehat{\lambda}_t = \frac{\widehat{\lambda}_{mc}}{\widehat{\lambda}_e \pi R^2} \quad (4)$$

was evaluated. For *Picea abies* one gets $\widehat{\lambda}_t^{spruce} = 0.00774$, while for *Fagus sylvatica* $\widehat{\lambda}_t^{beech} = 0.00253$ is obtained. From the estimated parameters one can conclude that there is stronger clustering within a smaller cluster radius for spruce roots, while for the beech roots the clustering is weaker, but the cluster radius is slightly larger.

Inhomogeneous Matern-Cluster Model

For the original data, which shows a vertical distribution property, the Matern-cluster model fitted for the homogeneous case has to be retransformed, where the inverse transformation

$$h_{orig} = (F^*)^{-1}\left(\frac{F^*(h_{tot})}{h_{tot}}h_{tran}\right) \quad (5)$$

of the depth is considered, with $(F^*)^{-1}(y)$ representing the generalized inverse function of $F^*(x)$. Then, in the retransformed model, the parent process is given by an inhomogeneous Poisson process with intensity function

$$\lambda_e(x, y) = \lambda_e(y) = \lambda_e \frac{f^*(y)}{F^*(h_{tot})} h_{tot}, \quad (6)$$

where x and y represent the horizontal and vertical coordinate, $f^*(x)$ is the density function of the suitable distribution function $F^*(x)$ (exponential distribution for spruces and gamma distribution for beeches), λ_e is the intensity of the parent process of the homogeneous model and h_{tot} represents the total depth of the sampling window. The cluster regions are no longer circles, but the images of these circles under the mapping given in (5). They can be written as

$$\{(x, y) : (x - x_p)^2 + (F^*(y) - F^*(y_p))^2 \left(\frac{h_{tot}}{F^*(h_{tot})}\right)^2 \leq R^2\}, \quad (7)$$

where the corresponding parent point is denoted as (x_p, y_p) . Since the mean total number of points in the given window as well as the mean total number of points in a cluster stay the same compared to the homogeneous model, the intensity function for the inhomogeneous matern cluster point process is given as

$$\lambda_{mc}(x, y) = \lambda_{mc}(y) = \lambda_{mc} \frac{f^*(y)}{F^*(h_{tot})} h_{tot}, \quad (8)$$

where λ_{mc} is the corresponding intensity of the homogeneous model. Figure 11 shows a realization of the inhomogeneous Matern-cluster model, which corresponds to the homogeneous realization displayed in Figure 10 using an exponential depth distribution. Note that only those simulated data shown in the upper part of Figure 11 should be used for interpretation purposes. In the lower part of Figure 11 the influence of transformation and retransformation of data clearly dominates the original spatial structure of those (sparse) root data with larger vertical depths.

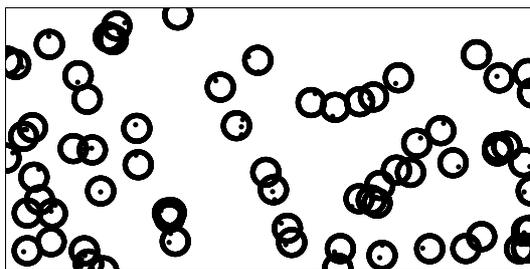


Fig. 10. Realization of the homogeneous Matern-cluster model fitted in the case of *Picea abies*

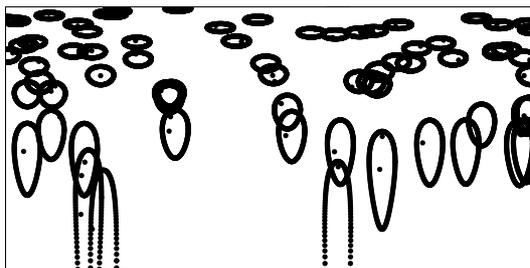


Fig. 11. Realization of the retransformed inhomogeneous Matern-cluster model

4 Discussion

In both studied cases clear differences between two biologically distinct groups can be recognized using estimated point process characteristics. Apart from rejecting in all cases the null hypotheses of having homogeneous Poisson process as generating processes for the observed point patterns, it has been possible in the first example to detect a distance region where the number of chromocenters differ strongly between the non-differentiated and the differentiated state of NB4 cell nuclei. In the second example differences in the clustering behavior of the fine roots for European Beech compared to Norway Spruce have become visible. Apart from that a simple point process model has been fitted to the tree root data.

4.1 NB4 Cell Nuclei

The centromeric regions of chromosomes represent an important part of the heterochromatin compartment in interphase nuclei. A previous study was focused on the quantitative description of three dimensional distribution patterns of centromeric heterochromatin in NB4 cells using features of the MST ([4]). From a mathematical point of view, this approach has several disadvantages. Quantities like the MST edge lengths or their variance are strongly dependent on the mean number of points per volume unit. Apart from that,

the methods applied in the present study allow to get inference about different specific regions of point pair distances. Thus, this approach provides the opportunity for a more detailed analysis of three dimensional centromere distributions.

Notice that, although the observed point patterns are finite and bounded, it can be assumed that they are realizations of stationary point processes restricted to a bounded sampling region. This notional step is supported by the fact that tests for isotropy and stationarity do not show any significant rejections and that the volumes of the assumed sampling regions before and after differentiation are of comparable sizes. The method of assuming unbounded stationary point processes as sources for observed realizations restricted to a bounded sampling region is a quite common practice since very often data is given in finite sampling regions and behave in a rather different non-stationary way outside of these regions ([11], [40]).

Due to the assumption that the observed point samples in the bounded sampling regions are extracts of unbounded realizations of stationary point processes it is necessary, although having only bounded sampling regions, to perform edge-corrections in order to insure compatibility with the applied methods. We want to emphasize that this procedure has its statistical justification in the facts that tests for isotropy and stationarity do not show any significant rejections and that the sampling regions have similar volumes.

Other types of estimators apart from spatial Horvitz-Thompson style estimators, e.g. of Kaplan-Meier type ([3]) and other techniques of edge corrections might also be applicable.

Clustering of chromosomal regions in interphase cell nuclei is supposed to be an important mechanism regulating the functional architecture of chromatin. In our previous study, we observed a progressive clustering of centromeric heterochromatin after differentiation of NB4 cells with ATRA ([4]). These clusters (chromocenters) represent groups of centromeres with a distance below the limit of spatial resolution of optical microscopy. In the present study, we have analyzed the distance of these chromocenters and found a higher frequency of distances between 350 *nm* and 800 *nm* for undifferentiated cells in comparison to ATRA-differentiated NB4 cells (Figures 4 and 6). These new data imply the existence of heterochromatin regions with a range of 350 *nm* to 800 *nm* containing functionally related centromeric zones. The centromeres in these regions cluster during ATRA-induced differentiation of NB4 cells as demonstrated by the decreased number of detectable chromocenters, i.e. groups of centromeres within a sphere with a diameter of less than 350 *nm*. The existence of heterochromatin regions containing centromeres of specific chromosomes would imply that the restructuring of these chromosome territories has to proceed in a coordinated nonrandom way during the differentiation-induced "collapse" of these heterochromatin zones. This model is in accordance with a topological model for gene regulation based on the structural remodeling of chromosome territories during modulation of transcription ([8], [33]).

Another important result of the present study is the finding that the 3D distribution of chromocenters is not completely random in undifferentiated as well as in ATRA-differentiated NB4 cells. These findings, thus, rejects a previous hypothesis which was based on the comparison of centromere distributions in NB4 cells with simulated completely random patterns using MST features ([4]). The result of the present study is in accordance with other studies, which suggested that interphase centromeres are not arranged in a completely random way ([17], [21]). Importantly, investigations of interphase chromosome positions indicate that a strictly maintained structure of chromatin appears to be necessary for a normal function of cells even in tumours ([9]).

4.2 Planar Sections of Root Systems

The point process characteristics using the transformed data described the two dimensional distribution of small roots in pure stands of *Picea abies* and *Fagus sylvatica*. The results for such homogenised data using the averaged estimated pair correlation function (Fig. 9) show that attraction can be observed for point pairs with distances less than approximately 14 cm. It means that roots of both species tend to cluster in areas up to this diameter. As roots react to nutrient enriched soil patches by enhanced growth and greater biomass in these areas ([12], [30]), this attraction of roots within this diameter could also be a direct link to a local occurrence of soil resources. On the other hand, for very small distances of less than 0.5 cm there is a hardcore property in the homogeneous case, which can possibly be explained by the thickness of the regarded roots. Also, as the small roots are associated to the uptake-oriented fine roots, concentration of the small roots in clusters of smaller diameters (i.e. less than 0.5 cm) is not reasonable. In this context it is important to notice that these structural differences are independent of the observed significant difference in the average number of detected points (roots) for the samples of *Picea abies* and *Fagus sylvatica*.

The homogenized point patterns were modeled as Matern-cluster processes with estimated parameters described in Section 3.2. The Matern-cluster model chosen has some serious advantages. First the model is of a certain simplicity and theoretical values for point process characteristics are known. Even more important is that the sample data is fitted well by this model. The estimated point process characteristics using the Matern-cluster processes further differentiated between the species. The results show for spruces a stronger clustering in a smaller range of attraction ($\hat{R}_{spruce} = 4.9$ cm), while the clustering is weaker for beeches, but the range of attraction ($\hat{R}_{beech} = 7.4$ cm) seems to be slightly larger. This finding is in accordance with another investigation ([42]) calculating influence areas for each root. Their results indicated that the root system of spruce requires more roots to achieve a similar degree of space acquisition and thus beech exploits patchily distributed soil resources at lower root numbers. In summary there is a combination of two effects, the depth distribution already described in [41] and the cluster effects analyzed in the present

paper. Structural differences between spruce and beech indicated in [42] have been mathematically described. Stronger clustering in the case of spruce than in the case of beech can be seen also regarding the characteristics mentioned above as well as by a comparison of the estimated parameters for the homogeneous model $\hat{\lambda}_t^{spruce} = 0.00774$ vs. $\hat{\lambda}_t^{beech} = 0.00253$, keeping in mind that the estimated parent intensities are almost equal ($\hat{\lambda}_p^{spruce} = 0.00690$ vs. $\hat{\lambda}_p^{beech} = 0.00603$). The previous GIS-based investigation of root distribution ([42]) was not able to quantify the differences of clustering between the two species so precisely as the applied modelling by point processes.

Finally a non-homogeneous Matern-cluster model has been constructed by a retransformation of the homogeneous model, thereby reflecting the observed depth distribution of the tree roots. The visualisation of the retransformed data suggests a depth-dependent size and shape of root clusters. Close to the soil surface, roots form clusters along the horizontal axis. This shape agrees also with the horizontally distributed root points in the original samples (c.f. Fig. 3). Horizontally distributed roots as well as the shape of generated clusters may reflect the attractive soil patches in the nutrient-rich topsoil layers. Deeper, the real size of clusters is larger and more circular. However, because the transformation and retransformation of root data at low intensities in the deep parts of the soil profile makes the results unstable, the lower third of Figure 11 is not really useful for interpretation of spatial structures. The investigated small roots were also described regarding water and nutrient uptake ([25]) and mediates to the most active fine roots ($< 2\text{ mm}$). Thus, clusters of small roots reflect the presence of nutrient patches or zones of better water availability ([19], [34], [39]). As the number of small roots and their clustering was independent of the distance to the surrounding trees and of their diameter ([42]), the root clusters are suggested as an inherent property of below-ground space acquisition.

5 Acknowledgement

Data collection of the profile walls was financially supported by the Austrian Science Foundation within the Special Research Program "Restoration of Forest Ecosystems", F008-08.

References

1. Alcobia, I., Quina, A.S., Neves, H., Clode, N., Parreira, L.: The spatial organization of centromeric heterochromatin during normal human lymphopoiesis: evidence for ontogenically determined spatial patterns. *Exp Cell Res.* **290**, 358-369 (2003)
2. Baddeley, A. J.: Spatial sampling and sensing. In: Kendall, W. S., Lieshout, M. N. M. van, Barndorff-Nielsen, O. E (eds) *Current Trends*

- in Stochastic Geometry and its Applications. Chapman and Hall, London, New York (1998)
3. Baddeley, A. J., Gill, R. D.: Kaplan-Meier Estimators of Distance Distribution for Spatial Point Processes. *The Annals of Statistics* **25**, No. 1, 263-292 (1997)
 4. Beil, M., Durschmied, D., Paschke, S., Schreiner, B., Nolte, U., Bruel, A., Irinopoulou, T. : Spatial distribution patterns of interphase centromeres during retinoic acid-induced differentiation of promyelocytic leukemia cells. *Cytometry*. **47**, 217-225 (2002)
 5. Beil, M., Fleischer, F., Paschke, S., Schmidt, V.: Statistical analysis of 3D centromeric heterochromatin structure in interphase nuclei. *Journal of Microscopy* **217**, 60-68 (2005)
 6. Caldwell, M. M., Manwaring, J. H., Durham, S. L.: Species interaction at the level of fineroots in the field: influence of soil nutrient heterogeneity and plant size. *Oecologia* **106** 440-447 (1996)
 7. Chiu, S. N., Stoyan D.: Estimators of distance distributions for spatial patterns. *Statistica Neerlandica* **52/2**, 239-246 (1998)
 8. Cremer, T., Cremer, C.: Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat Rev Genet.* **2**, 292-301 (2001)
 9. Cremer, M., Kupper, K., Wagler, B., Wizelman, L., von Hase, J., Weiland, Y., Kreja, L., Diebold, J., Speicher, M. R., Cremer, T.: Inheritance of gene density-related higher order chromatin arrangements in normal and tumor cell nuclei. *Journal Cell Biol.* **162**, 809-820 (2003)
 10. Diggle, P. J.: *Statistical Analysis of Spatial Point Patterns*. Academic Press, London (1989)
 11. Diggle, P. J., Mateu, J., Clough, H. E.: A comparison between parametric and non-parametric approaches to the analysis of replicated spatial point patterns. *Advances in Applied Probability* **32**, 331-343 (2000)
 12. Drew M. C.: Comparison of the effects of a localized supply of phosphate, nitrate, ammonium and potassium on the growth of the seminal root system, and the shoot, of barley. *New Phytol.* **75**, 479-490 (1975)
 13. Dussert C., Rasigni, G., Palmari, J., Rasigni, M. Llebaria, A., Marty, F.: Minimal spanning tree analysis of biological structures. *Journal of Theoretical Biology* **125**, 317-323 (1987)
 14. Dussert, C., Rasigni, G., Rasigni, M., Palmari, J., Llebaria, A.: Minimal spanning tree: a new approach for studying order and disorder. *Physical Revue B* **34**, 3528-3531 (1986)
 15. Fenaux P., Chomienne C., Degos L.: Acute promyelocytic leukemia: biology and treatment. *Semin Oncol* **24**,92-102 (1997)
 16. Fleischer, F., Eckel, S., Schmid, I., Schmidt, V., Kazda, M.: Statistical analysis of the spatial distribution of tree roots in pure stands of *Fagus sylvatica* and *Picea abies*. Preprint (2005)
 17. Haaf, T., Schmid, M.: Chromosome topology in mammalian interphase nuclei. *Exp Cell Res.* **192**, 325-332 (1991)
 18. Hanisch, K.-H.: Some remarks on estimators of the distribution function of nearest-neighbor distance in stationary spatial point patterns. *Statistics* **15**, 409-412 (1984)
 19. Jackson R. B., Caldwell M. M.: Geostatistical patterns of soil heterogeneity around individual perennial plants. *J. Ecol.* **81**, 683-692 (1993)

20. Jackson, R. B., Caldwell, M. M.: Integrating resource heterogeneity and plant plasticity: modeling nitrate and phosphate uptake in a patchy soil environment. *Journal of Ecology* **84**, 891-903 (1996)
21. Janevski, J., Park, P. C., De Boni, U.: Organization of centromeric domains in hepatocyte nuclei: rearrangement associated with de novo activation of the vitellogenin gene family in *Xenopus laevis*. *Exp Cell Res.* **217**, 227-239 (1995)
22. Krauss, G., Müller, K., Gärtner, G., Härtel, F., Schanz, H., Blanckmeister, H.: Standortsgemässe Durchführung der Abkehr von der Fichtenwirtschaft im nordwestsächsischen Niederland. *Tharandter Forstl. Jahrbuch* **90**, 481-715 (1939)
23. Lanotte, M., Martin-Thouvenin, V., Najman, S., Balerini, P., Valensi, F., Berger, R.: NB4, a maturation inducible cell line with t(15;17) marker isolated from a human acute promyelocytic leukemia (M3). *Blood.* **77**, 1080-1086 (1991)
24. Lieshout, M. C. van, Baddeley, A. J.: A nonparametric measure of spatial interaction in point patterns. *Statistica Neerlandica* **50**, 344-361 (1996)
25. Lindenmair, J., Matzner, E., Gttlein, A., Kuhn, A. J., Schröder, W. H.: Ion exchange and water uptake of coarse roots of mature Norway spruce trees. In: Horst, W., J. et al. *Plant Nutrition - Food Security and Sustainability of Agro-Ecosystems*, Kluwer Academic Publishers (2001)
26. Marcelpoil, R., Usson, Y.: Methods for the study of cellular sociology: voronoi diagrams and parametrization of the spatial relationships. *Journal of Theoretical Biology* **154**, 359-369 (1992)
27. Mayer, J.: On quality improvement of scientific software: Theory, methods, and application in the GeoStoch development. *Doctoral Dissertation, University of Ulm* (2003)
28. Mayer, J., Schmidt, V., Schweiggert, F.: A unified simulation framework for spatial stochastic models. *Simulation Modelling Practice and Theory* **12**, 307-326 (2004)
29. Molenaar C., Wiesmeijer K., Verwoerd N. P., Khazen S., Eils R., Tanke H. J., Dirks R.W.: Visualizing telomere dynamics in living mammalian cells using PNA probes. *EMBO Journal* **22**, 6631-6641 (2003)
30. Morris E. C.: Effect of localized placement of nutrients on root-thinning in self-thinning populations. *Ann. Bot.* **78**, 353-364 (1996)
31. Okabe, A., Boots, B., Sugihara, K., Chiu, S. N.: *Spatial Tessellations* J. Wiley & Sons, Chichester (2000)
32. Park, P. C., De Boni, U.: Spatial rearrangement and enhanced clustering of kinetochores in interphase nuclei of dorsal root ganglion neurons in vitro: association with nucleolar fusion. *Exp. Cell Res.* **203**, 222-229 (1992)
33. Park, P. C., De Boni, U.: Dynamics of structure-function relationships in interphase nuclei. *Life Sci.* **64**, 1703-1718 (1999)
34. Parker M. M., van Lear D. H.: Soil heterogeneity and root distribution of mature loblolly pine stands in piedmont soils. *Soil Sci. Society Am. J.* **60**, 1920-1925 (1996)
35. Perrod S., Gasser S. M.: Long-range silencing and position effects at telomeres and centromeres: parallels and differences. *Cell Mol Life Sci.* **60**, 2303-2318 (2003)
36. Richards E. J., Elgin S. C. R.: Epigenetic codes for heterochromatin formation and silencing. *Cell.* **108**, 489-500 (2002)

37. Ripley, B. D.: The second-order analysis of stationary point processes. *Journal of Applied Probability* **13**, 255-266 (1976)
38. Ripley, B. D.: *Spatial Statistics*. J. Wiley & Sons, Chichester (1981)
39. Ryel R. J., Caldwell M. M., Manwaring J. H.: Temporal dynamics of soil spatial heterogeneity in sagebrush-wheatgrass steppe during a growing season. *Plant Soil* **184**, 299-309 (1996)
40. Schladitz, K., Särkkä, A., Pavenstädt, I., Haferkamp, O., Mattfeldt, T.: Statistical analysis of intramembranous particles using fracture specimens. *Journal of Microscopy* **211**, 137-153 (2003)
41. Schmid, I., Kazda, M.: Vertical distribution and radial growth of coarse roots in pure and mixed stands of *Fagus sylvatica* and *Picea abies*. *Canadian Journal of Forest Research* **31**, 539-548 (2001)
42. Schmid, I., Kazda, M.: Clustered root distribution in mature stands of *Fagus sylvatica* and *Picea abies*. Preprint (2005)
43. Stoyan, D., Kendall, W. S., Mecke, J.: *Stochastic Geometry and its Applications*. J. Wiley & Sons, Chichester (1995)
44. Stoyan, D., Stoyan, H.: *Fractals, Random Shapes and Point Fields. Methods of geometrical statistics*. J. Wiley & Sons, Chichester (1994)
45. Stoyan, D., Stoyan, H.: Improving ratio estimators of second order point process characteristics. *Scandinavian Journal of Statistics* **27**, 641-656 (2000)
46. Stoyan, D., Stoyan, H., Tscheschel, A., Mattfeldt, T.: On the estimation of distance distribution functions for point processes and random sets. *Image Analysis and Stereology* **20**, 65-69 (2001)
47. The H. de, Chomienne, C., Lanotte, M., Degos, L., Dejean, A. : The t(15;17) translocation of acute promyelocytic leukaemia fuses the retinoic acid receptor alpha gene to a novel transcribed locus. *Nature*. **347**, 558-561 (1990)

A Point Process Characteristics and their Estimators

In the following let $\mathbf{x} = \{\mathbf{x}_n\}$ be a random point process in \mathbb{R}^d , where $d \in \{2, 3, \dots\}$ and let $N(B) = \#\{n : \mathbf{x}_n \in B\}$ denote the number of points \mathbf{x}_n of \mathbf{x} located in a sampling window B

Intensity Measure

The intensity measure Λ is defined as

$$\Lambda(B) = EN(B) \quad (1)$$

for a given set B . Hence $\Lambda(B)$ is the mean number of points in B . In the homogeneous case it suffices to regard an intensity λ since then

$$\Lambda(B) = \lambda\nu_d(B) \quad (2)$$

where $\nu_d(B)$ denotes the volume of B . A natural estimator for λ is given by

$$\hat{\lambda} = \frac{N(B)}{\nu_d(B)}. \quad (3)$$

However, for the estimation of the nearest-neighbor distance distribution a different estimator

$$\hat{\lambda}_H = \sum_{\mathbf{x}_n \in B} \frac{1_{B \ominus b(o, s(\mathbf{x}_n))}(\mathbf{x}_n)}{\nu_d(B \ominus b(o, s(\mathbf{x}_n)))} \quad (4)$$

is recommended ([46]), where $s(\mathbf{x}_n)$ denotes the distance of \mathbf{x}_n to its nearest neighbour and $b(x, r)$ is the ball with radius r and midpoint x .

Notice that, following the recommendation in [45], λ^2 has been estimated by

$$\hat{\lambda}^2 = \frac{N(B)(N(B) - 1)}{(\nu_d(B))^2}, \quad (5)$$

since even in the Poisson case $(\hat{\lambda})^2$ is not an unbiased estimator for λ^2 .

Moment Measure and Product Density

Let B_1 and B_2 be two sets. The second factorial moment measure $\alpha^{(2)}$ of \mathbf{x} is defined by

$$\alpha^{(2)}(B_1 \times B_2) = E\left(\sum_{\substack{\mathbf{x}_1, \mathbf{x}_2 \in \mathbf{N} \\ \mathbf{x}_1 \neq \mathbf{x}_2}} 1_{B_1}(\mathbf{x}_1) 1_{B_2}(\mathbf{x}_2) \right). \quad (6)$$

Often $\alpha^{(2)}$ can be expressed using a density function $\varrho^{(2)}$ as follows

$$\alpha^{(2)}(B_1 \times B_2) = \int_{B_1} \int_{B_2} \varrho^{(2)}(x_i, x_j) dx_i dx_j. \quad (7)$$

The density function $\varrho^{(2)}$ is called the second product density. If one takes two balls C_1 and C_2 with infinitesimal volumes dV_1 and dV_2 and midpoints x_1 and x_2 respectively, the probability for having in each ball at least one point of \mathbf{x} is approximately equal to $\varrho^{(2)}(x_1, x_2)dV_1dV_2$. In the homogeneous and isotropic case $\varrho^{(2)}(x_1, x_2)$ can be replaced by $\varrho^{(2)}(r)$, where $r = \|x_1 - x_2\|$. As an estimator

$$\widehat{\varrho^{(2)}}(r) = \frac{1}{db_d r^{d-1}} \sum_{\substack{\mathbf{x}_i, \mathbf{x}_j \in B \\ i \neq j}} \frac{k_h(r - \|\mathbf{x}_i - \mathbf{x}_j\|)}{\nu_d(B_i \cap B_j)} \quad (8)$$

has been used ([45]), where $k_h(x)$ denotes the Epanechnikov kernel

$$k_h(x) = \frac{3}{4h} \left(1 - \frac{x^2}{h^2}\right) 1_{(-h, h)}(x), \quad (9)$$

$B_{\mathbf{x}_j} = \{x + \mathbf{x}_j : x \in B\}$ is the set B translated by the point \mathbf{x}_j , and the sum in (8) extends over all pairs of points $\mathbf{x}_i, \mathbf{x}_j \in B$ with $i \neq j$. The bandwidth h has been chosen as $h = c\widehat{\lambda}^{-1/d}$ with a fixed parameter c .

Pair Correlation Function

The product density $\varrho^{(2)}(r)$ is used to obtain the pair correlation function $g(r)$ as

$$g(r) = \frac{\varrho^{(2)}(r)}{\lambda^2}. \quad (10)$$

The pair correlation function at a certain value r can be regarded as the frequency of point pairs with distance r , where $g(r) = 1$ is a base value. The pair correlation function can be estimated by the usage of estimators for $\varrho^{(2)}(r)$ and λ^2 respectively. In particular, we consider the estimator

$$\widehat{g}(r) = \frac{\widehat{\varrho^{(2)}}(r)}{\widehat{\lambda^2}}, \quad (11)$$

where $\widehat{\lambda^2}$ and $\widehat{\varrho^{(2)}}$ are given by (5) and (8), respectively. Note that $g(r) \geq 0$ for all distances r . In the Poisson case $g_{Poi}(r) \equiv 1$, therefore $g(r) > 1$ indicates that there are more point pairs having distance r than in the Poisson case, while $g(r) < 1$ indicates that there are less point pairs of such a distance.

K-Function

Ripley's K-function ([37]) is defined such that $\lambda K(r)$ is the expected number of points of the stationary point process $\mathbf{x} = \{\mathbf{x}_n\}$ within a ball $b(\mathbf{x}_n, r)$ centered at a randomly chosen point \mathbf{x}_n which itself is not counted. Formally

$$\lambda K(r) = E \sum_{\mathbf{x}_n \in B} \frac{N(b(\mathbf{x}_n, r)) - 1}{\lambda \nu_d(B)}. \quad (12)$$

The K-function has been estimated by

$$\widehat{K}(r) = \frac{\widehat{\kappa}(r)}{\widehat{\lambda}^2}, \quad (13)$$

where

$$\kappa(r) = \sum_{\substack{\mathbf{x}_i, \mathbf{x}_j \in B \\ i \neq j}} \frac{1_{b(o,r)}(\mathbf{x}_j - \mathbf{x}_i)}{|B_{\mathbf{x}_j} \cap B_{\mathbf{x}_i}|}, \quad (14)$$

For Poisson processes it is easy to see that $K_{Poi}(r) = b_d r^d$.

L-Function

Often it is more convenient to scale the $K(r)$ in order to get a function equal to r for the Poisson case. Hence $L(r)$ is defined as

$$L(r) = \sqrt[d]{\frac{K(r)}{b_d}}, \quad (15)$$

where b_d denotes the volume of the d -dimensional unit sphere. Thus, in the Poisson case, we have

$$L(r) - r = 0. \quad (16)$$

A natural estimator for $L(r)$ is given by

$$\widehat{L}(r) = \sqrt[d]{\frac{\widehat{K}(r)}{b_d}}. \quad (17)$$

Nearest-Neighbor Distance Distribution

The nearest-neighbor distance distribution D is the distribution function of the distance from a randomly chosen point \mathbf{x}_n of the given stationary point process \mathbf{x} to its nearest neighbor. Hence $D(r)$ is the probability that a randomly chosen point \mathbf{x}_n of \mathbf{x} has a neighbor with a distance less than or equal to r . According to [46] we used the Hanisch estimator $\widehat{D}_H(r) = D_H(r)/\widehat{\lambda}_H$ ([2], [18]) with

$$\widehat{D}_H(r) = \sum_{\mathbf{x}_n \in B} \frac{1_{B \ominus b(o, s(\mathbf{x}_n))}(\mathbf{x}_n) 1_{(0,r]}(s(\mathbf{x}_n))}{\nu_d((B \ominus b(o, s(\mathbf{x}_n))))}. \quad (18)$$

A useful property of the nearest-neighbor distance distribution is that in the case of stationary Poisson processes we have

$$D_{Poi}(r) = 1 - \exp(-\lambda b_d r^d). \quad (19)$$

Therefore one can conclude that $D(r) < D_{Poi}(r)$ indicates rejection between points, on the other hand $D(r) > D_{Poi}(r)$ indicates attraction, keeping in

mind that the nearest-neighbor distance distribution function is a cumulated quantity.

Spherical Contact Distribution Function

The spherical contact distribution function $H_s(r)$ is the distribution function of the distance from an arbitrary point, chosen independently of the point process \mathbf{x} , to the nearest point belonging to \mathbf{x} . Notice that the value $H_s(r)$ can be interpreted as the probability that at least one point \mathbf{x}_n of \mathbf{x} is in the sphere of radius r centered at the origin. As an estimator for $H_s(r)$,

$$\widehat{H}_s(r) = \frac{\nu_d((B \ominus b(0, r)) \cup_{\mathbf{x}_n \in B} b(\mathbf{x}_n, r))}{\nu_d(B \ominus b(0, r))} \quad (20)$$

is used.

J-Function

Based on $H_s(r)$ and on $D(r)$, Baddeley's J-function is defined by

$$J(r) = \frac{1 - H_s(r)}{1 - D(r)}. \quad (21)$$

where

$$\widehat{J}(r) = \frac{1 - \widehat{H}_s(r)}{1 - \widehat{D}(r)}, \quad (22)$$

is a natural estimator for $J(r)$. In the case of Poisson point processes $J_{Poi}(r) \equiv 1$ and therefore if $J(r) > 1$ one can conclude that there is repulsion between point pairs of distance r . On the other hand if $J(r) < 1$ there is attraction between point pairs compared to the case of complete spatial randomness.

B Matern-Cluster Model

The Matern-cluster point process \mathbf{x}_{mc} is based on a Poisson process with intensity λ_e whose points are called parent points. Around each parent point a sphere with radius R is taken in which the points of the Matern-cluster process are scattered uniformly where the number of points in such a sphere is Poisson distributed with parameter $R^d b_d \lambda_t$. Notice that λ_t is the mean number of points per unit area generated by a single parent point in a sphere of radius R . Since the parent points themselves are not part of the Matern-cluster process, its intensity is given as

$$\lambda_{mc} = R^d b_d \lambda_t \lambda_e. \quad (23)$$

Thus, the Matern-cluster point process \mathbf{x}_{mc} is uniquely determined by three of the four parameters λ_e , λ_t , R and λ_{mc} . Obviously, for small distances, points of the Matern-cluster process are attracted to each other, in other words there

is a bigger expected number of points of \mathbf{x}_{mc} in a sphere around an arbitrarily chosen point of \mathbf{x}_{mc} than for Poisson processes of comparable intensity $\lambda_{Poi} = \lambda_{mc}$. For \mathbf{x}_{mc} , closed formulae for the point process characteristics described in Appendix A are known ([44]).