

Numerical Methods for characterizing Structure and Diversity applied to a natural Tropical Forest and an even aged Teak stand

Peter Biber and Horst Weyerhaeuser

1. Introduction

1.1. Why dealing with numerical methods for describing forest structure and diversity in forest management ?

There are at least two good reasons for considering aspects of structure and biodiversity in forest management. First, there is a strong feedback relation between the spatial structure, e.g. the horizontal arrangement of trees and tree species in a stand and the ongoing growth processes. That means, that forest growth will be strongly determined by stand structure, but as growth is a continuous process, it will itself change the structure of the stand (PRETZSCH, 1996). As knowledge about forest growth is essential for sustainable forest management, we should always incorporate structural aspects. The second important reason for considering structure is the present worldwide development of forest management, moving away from pure timber production towards multiple purpose forestry, regarding forests as managed ecosystems. Such an approach is concerned by aspects of forest utilization in the same way as it is by nature conservation and protection. As we can see from recent research work (AMMER et al., 1995; BLAB, 1986; ELLENBERG et al., 1985; HABER, 1982) the ecological value of forest stands as a habitat for many plants and animals and its capabilities to withstand disturbances like insect attacks or windthrow increases considerably with increasing structural and biological diversity. We can therefore consider structure and diversity as multi-dimensional stand characteristics determined by many single aspects. Some main aspects could be

- the horizontal pattern of tree distribution, with and without regard to tree dimensions,
- stand density,
- differentiation in height and diameter structure,
- species richness (diversity),
- the mixture of tree species.

As we try to characterize some of these structural properties of a forest stand, we could do it as a verbal description as many experienced foresters in fact do. For example some descriptions for different horizontal tree distribution patterns could be terms like 'randomly distributed', 'clustered' or 'regularly distributed'. The advantage of a linguistical approach like this is, that for performing structural characterisation there is only experience needed, but not an expensive data assessment. On the other hand the essential disadvantage of the linguistical approach is evident, as the achievable classification of stand structure is very vague and strongly depends on the subjective opinion of the person who performs the classification. So if we want to compare structural properties of different stands, characterized by different persons, they may be not comparable at all. From this we can conclude for scientific studies of stand structure and its relationship to stand development, management or ecological value, that there is need for numerical methods which allow to express aspects of structure in a quantitative and therefore in an objective and comparable manner. Various research work on this problem has been done

(CLARK AND EVANS, 1954; PIELOU, 1975, 1977; RIPLEY, 1977, 1981; UPTON AND FINGLETON, 1985, 1989; PENTTINEN et al., 1992; STOYAN and STOYAN, 1992; DEGENHARDT, 1993; QUEDNAN and FRÖHLICH, 1994; PRETZSCH, 1997). So, we can rely on some well-tried methods to characterize structure and diversity of forest stands.

Besides many other purposes those methods give us the possibility to find out by scientific analysis, how stable forests are structured and which management concepts are suitable to achieve stable structures in managed stands.

1.2. Aim of this case study

The aim of this case study is to show how structure and diversity in forest stands can be characterized by numerical methods. Therefore five methods are chosen exemplarily. For the calculations required, the single tree based stand growth simulator SILVA, developed at the CHAIR OF FOREST YIELD SCIENCE of the MUNICH UNIVERSITY (PRETZSCH, 1992) is used in its latest version. This computer program cannot be used only as a tool for forest growth projection but also for providing information of timber amounts, the economic value and structural aspects of a given stand. By name, the methods we use are the aggregation index R by CLARK AND EVANS (1954), the segregation index S by PIELOU (1977) and the species profile index A by PRETZSCH (1996), which is a modification of the well known SHANNON-index (SHANNON 1948). Additionally, we estimate pair- and mark correlation functions as proposed by STOYAN and STOYAN (1992), PENTTINEN et al. (1992) and DEGENHARDT (1993).

To show how these methods perform when applied to totally different types of forests, they are calculated for a complex natural tropical forest stand on the one hand and one even aged, uniform Teak stand on the other hand. As SILVA could be parameterized for growth prognosis in pure Teak stands (KAHN, 1998) there was no sufficient database to calibrate the simulator for growth prognosis in natural tropical forest stands, however. In this study therefore, SILVA cannot be used for stand growth prognosis, but only to calculate numerical information about the status quo of the two examined forest stands.

2 The Data

2.1. The Natural Forest Stand

The natural forest data used in this study has been collected by Mr. H. WEYERHAEUSER, who is responsible for silviculture research in a National Park in Thailand. The research area is adjacent to the western border of Thailand with Myanmar. The forests in the project area consist mainly of two types, the dry dipterocarp forest (DDF) in the lower elevations and the mixed deciduous forest (MDF) in the higher and mostly more fertile locations. For this case study a forest stand of the MDF type has been chosen to represent the high structural diversity of such stands. The research plot has the size of 2 ha. 70 different tree species are represented in this plot out of 180 species which can be found in this forest type.

For this case study tree species were aggregated by botanic family and comparable ecology to ten groups, which were treated like single species in the further proceeding. Although SILVA can deal with at maximum 20 species, this clustering was done to provide for every species group a

tree number high enough for getting interpretable results. Anyway, the tree numbers within these groups differ in a wide range: while the *Terminalia* group counts for 215 trees, the *Azalia* group consists of only 18 trees (figure 1). The total amount of trees available for the natural forest is 787. Based on these data, a SILVA input file was constructed for further analysis. An impression of the horizontal tree distribution by species groups is given by figure 2.

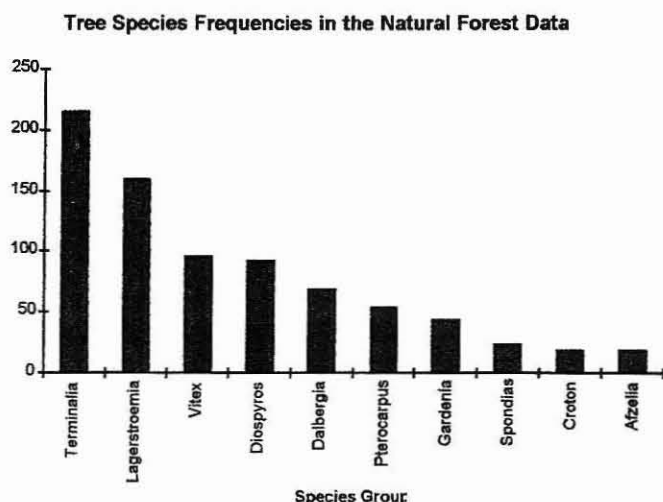


Figure 1: The ten tree species groups constructed and the tree numbers available for analysis of the natural forest data.

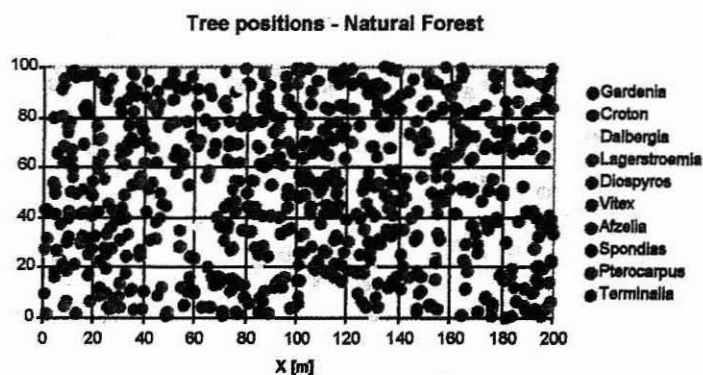


Figure 2: Positions of the single trees in the natural forest plot, marked by species group.

2.2. The Teak stand

During the DSE training course, three sample plots in Teak plantations were measured by the course participants. The stands are located in the Mae Moh region in Northern Thailand and grow under typical site conditions for this region. Stand ages were 12, 30 and 100 years at the time of measurement. For our analysis the 12 year old plot was chosen because it provided with 55 trees the greatest sample size of all measured plots. The parameters measured were dbh, total

height, height of the crown base, crown radii and x-y-coordinates of every tree. For more details see the articles by KAHN in this book. Like for the natural forest data, a SILVA input file was constructed to make the capabilities of this program usable for our analysis. Figure 3 gives a visual impression of the horizontal tree distribution in the 12 year old Teak stand. As can be seen, the plant rows are still quite visible.

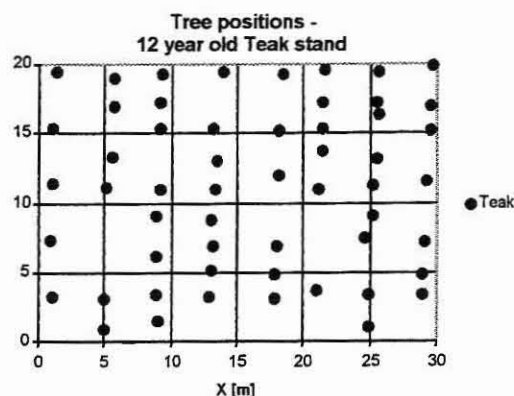


Figure 3: Positions of the single trees in the 12 year old Teak stand.

3. Methods applied

3.1. Single-Number-Indices

The term 'single-number-indices' stands for methods which create one single number to characterize a certain structural aspect. Such indices provide information at a highly condensed degree and are, in general, easy to interpret. Their disadvantage is, that this aggregated view may conceal some relevant structural information. As examples for single-number-indices we apply the aggregation index R by CLARK AND EVANS (1954), the segregation index S by PIELOU (1977) and the species profile index A by PRETZSCH (1996).

3.1.1. Aggregation Index by CLARK and EVANS (1954)

The aggregation index R by CLARK and EVANS (1954) is used to characterize the horizontal distribution pattern of trees in a stand by the so-called method of the nearest neighbour. The calculation of R starts with an estimation of the average distance of a tree to its nearest neighbour, $\bar{r}_{expected}$, which would be expected, if the trees were randomly (poisson-) distributed over the stand area:

$$\bar{r}_{expected} = \frac{1}{2 \cdot \sqrt{\frac{N}{A}}}$$

As this formula does not take into account the finiteness of experimental plots, $\bar{r}_{expected}$ can be slightly corrected for this purpose, using a proposal by DONELLY (1978) as it is also performed by the program SILVA:

$$\bar{r}_{expected} = \frac{1}{2 \cdot \sqrt{\frac{N}{A}}} + 0.051368 \cdot \frac{P}{N} + 0.041 \cdot \frac{P}{N^{\frac{3}{2}}}$$

For more details see DONELLY (1978), PRETZSCH (1993), QUEDNAU and FRÖHLICH (1994).

The second input parameter needed for the calculation of R is the observed mean distance of a tree to its nearest neighbour, $\bar{r}_{observed}$, which can easily be found by:

$$\bar{r}_{observed} = \frac{\sum_{i=1}^N r_i}{N}$$

The aggregation index R is now calculated as:

$$R = \frac{\bar{r}_{observed}}{\bar{r}_{expected}}$$

The variable names used above are:

- r_i = distance of tree i to its nearest neighbour in m,
- N = Number of trees found in the observed experimental plot,
- A = Size of the experimental plot in m^2 ,
- P = Perimeter of the experimental plot in m.

The theoretical Range of R lies between 0, which represents maximum aggregation, and 2.1491, which indicates a hexagonal distribution pattern. Values around 1.0 indicate random distribution. So, if a value of R is less than 1.0 we observe a tendency away from random distribution towards aggregation. If we find Values greater than 1.0 there is a tendency towards a regular spatial pattern. These tendencies can be tested for statistical significance, for details hereon see PRETZSCH (1993), QUEDNAU AND FRÖHLICH (1994).

3.1.2. Segregation Index S by PIELOU (1977)

Another important aspect of stand structure, the segregation of species, can be quantified using the segregation index S by PIELOU (1977). Like the aggregation index R, it relies on the method of the nearest neighbour. The Index S was originally designed by PIELOU for being applied to a two-species mixture. The idea is to estimate the number of mixed-species next-neighbour pairs $N_{expected}$, which we would expect, if the two tree species would be distributed

independently of each other. This number can be compared to the number of mixed-species pairs observed in fact, which we call $n_{observed}$. So, the Index S is calculated as follows:

$$S = 1 - \frac{n_{observed}}{n_{expected}}$$

All information necessary to find S for a given stand can be drawn out of a common fourfold table (Table 1).

Table 1: Fourfold table providing the information needed for calculation of the segregation index S. The variables a,b,c,d,m,n,v,w,N indicate the numbers of next-neighbour tree-pairs which have the properties shown by the row and column heads.

		Next neighbour		
Tree of interest	Species 1	Species 1	Species 2	Sum of Row
	Species 2	c	d	n
	Sum of column	v	w	N

This table can be used to calculate S in the following manner:

$$S = 1 - \frac{N \cdot (b + c)}{v \cdot n + w \cdot m}$$

The range of values which can be taken by S is $[-1; +1]$. If S is less than zero, we observe a tendency towards an association of the two examined species. If S is greater than zero it indicates a tendency towards a spatial segregation of the tree species. Values of S around 0 indicate that the two tree species are distributed independently of each other on the sample plot. The tendencies found by calculation of S can be tested statistically. For details on test statistics see PRETZSCH (1993) and also QUEDNAU AND FRÖHLICH (1994).

When S is applied to the natural forest data, we faced the problem, that we had to deal with more than two tree species. The Program SILVA is therefore designed to calculate S for every tree species in a stand, in a way, that the segregation of this species is calculated in relation to all trees, which belong to another species. So the segregation index given by SILVA for one species in a mixed forest stand with more than two species expresses the segregation of this species alone and all other species on the other hand. As it is not possible per definition to calculate S for a pure stand like our teak stand examined in this study, we define S to be 1 for pure stands, assuming that a pure stand is at maximum segregation.

3.1.3. Species Profile Index A by PRETZSCH (1996)

The third index applied in this study was developed by PRETZSCH (1996) to characterize the species profile of a forest stand. The more species are found in a stand and the more they are distributed over three height layers of the stand as defined below, the greater the value of the index, called A, will be. A is therefore an index which takes into account biodiversity and the vertical structure of the examined forest stand. The calculation of A is based on the well known

diversity index H of SHANNON and WEAVER (1948), which has been successfully applied in biological and ecological science. To give a better understanding of the index A it seems to be reasonable to first discuss the original SHANNON-index H , which is calculated as follows:

$$H = - \sum_{i=1}^S p_i \cdot \ln p_i$$

The variable names used are:

- S = number of Species found in the examined stand
- p_i = frequency of Species i in the whole stand; $p_i = \frac{n_i}{N}$
- n_i = number of trees belonging to species i
- N = total number of trees in the examined stand.

The idea of H is, that rare species contribute to a greater extent to biodiversity of an ecosystem than dominant species. This concept is realized by weighting the frequencies of the species by their negative logarithms. H will be the greater, the more species we find in a population and the more equal their frequencies are. If we find only one species, H will be zero, indicating the least diversity possible.

Understanding the concept of the SHANNON-index H , we can proceed introducing the species-profile index A . The first step to do is to divide the examined stand into three height layers, called 'zones' further on. If we assume the height of the highest tree in the stand to be 100%, zone 1 extends from 100% down to 80%, zone 2 from 80% down to 50% and zone 3 from 50% down to the forest ground (figure 4). If the tip of a tree is located in one of these zones, we consider the tree belonging to this zone. When the SHANNON-index H is applied to species and height zones at the same time, we get the species profile index A :

$$A = - \sum_{i=1}^S \sum_{j=1}^Z p_{ij} \cdot \ln p_{ij}$$

The variable names used are:

- S = number of species in the examined stand
- Z = number of height zones
- N = total number of trees in the examined stand,
- p_{ij} = frequency of Species i in the height zone j stand; $p_{ij} = \frac{n_{ij}}{N}$
- n_{ij} = number of trees of species i in height zone j .

If p_{ij} for one height zone j will be zero, the logarithm of p_{ij} is not defined. So, if this occurs, we define the contribution of Species i in height zone j to the value of index A being zero. For a pure species, single-layered stand A will be zero. A will increase considerably, if trees spread over more than one height zone and if additional species will be introduced. As PRETZSCH (1996) states, the index A quantifies, what is often understood as 'minglement structure' in forest management practice. Figure 4 illustrates the preceeding explanations.

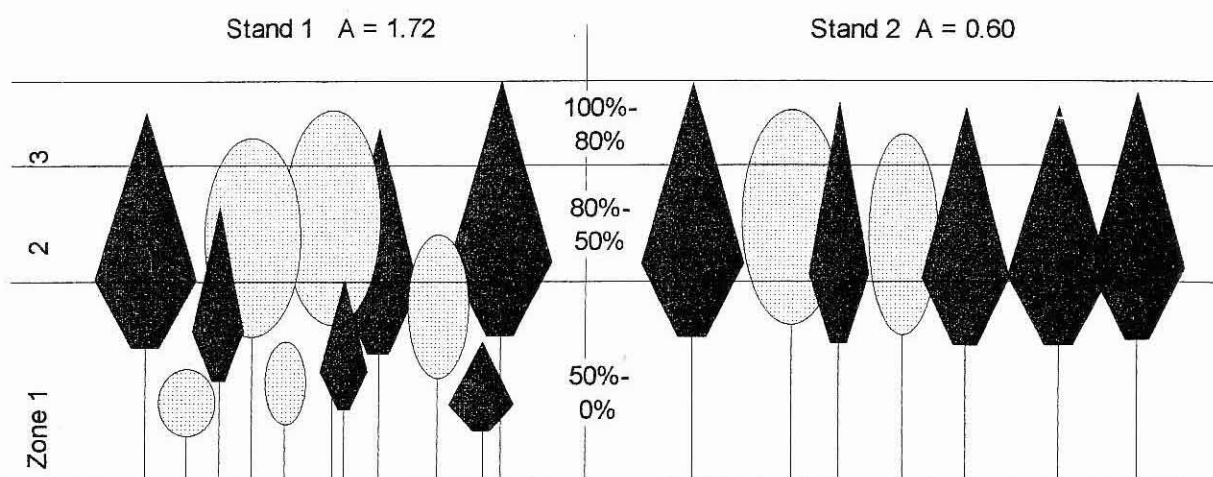


Figure 4: Behaviour of species profile index A (PRETZSCH, 1996) for one multi-layered (stand 1) and one single-layered two-species stand (stand 2).

3.2. Pair- and Markcorrelation-Functions

In contrast to the single-number-indices we dealt with above, the methods shown below do not aggregate structural information to one single number. We get a series of numbers which can be interpreted as estimated values of functions with certain properties. These methods give a much more detailed view on structural aspects than single-number-indices do, their theoretical background and their handling is much more complex, however. In the following text we introduce the so called pair- and mark correlation functions as proposed by STOYAN and STOYAN (1992) and DEGENHARDT (1993).

3.2.1. Pair Correlation Functions

The task of pair correlation functions is similar to the single-number aggregation index R by CLARK and EVANS (1954). Like this index, pair correlation functions regard the horizontal distribution of trees in relation to a random distribution without considering their dimensions. Pair correlation functions enable one to decide, whether pairs of trees at a certain horizontal distance occur more or less frequently than we would expect for a random distribution. These pairs of trees are not automatically next neighbour pairs as we looked for calculating the indices R and S. They are just the pairs of trees which have a certain distance, no matter if they are first, second, third or nth neighbours.

Definition and Properties of Paircorrelation-Functions

We imagine the horizontal distribution of n trees in an area W as one realisation of an isotropic and homogenous stochastic point process (see STOYAN and STOYAN, 1992). The intensity of this point process is called λ and can be estimated as follows:

$$\hat{\lambda} = \frac{n}{A}$$

where n is the number of trees found in the area W of the size A . Now consider two circles C_1 and C_2 with infinitesimally small areas dF_1, dF_2 . The distance of their centers is called r . So, the probability $P(r)$ to find one stem-center-point in each one of these two circles can be denoted as:

$$P(r) = g(r) \cdot \lambda \cdot dF_1 \cdot \lambda \cdot dF_2$$

The function $g(r)$ is called pair correlation function.

The paircorrelation-function has some typical properties:

if $g(r) > 1$, then the number of tree-pairs having the distance r is greater than expected for a random (Poisson-) distribution. If $g(r) = 1$, then the number of tree-pairs with the distance r equals the expected number for a random distribution. If $g(r) < 1$, then we observe a lesser number of tree-pairs at the distance r as we would expect for a random distribution. So, if the distribution of all trees in the area W would be a realisation of a Poisson-process, $g(r)$ would be 1 for every r . For big values of r , $g(r)$ comes close to the value 1 in any case.

Estimation of Paircorrelation-Functions

If we want to estimate $g(r)$, we apply the following formula for every $r > 0$ we are interested in:

$$\hat{g}(r) = \frac{1}{2\pi \cdot r \cdot \hat{\lambda}^2 \cdot \gamma(r)} \cdot \sum_{i=1}^n \sum_{\substack{j=1 \\ (j \neq i)}}^n k_h(r - |q_i - q_j|) \quad , \quad r > 0 \quad (1)$$

The symbols used are:

- $\hat{g}(r)$ = estimator for $g(r)$,
- q_i, q_j = Stem center points of tree i and j . They have the coordinates x_i, y_i and x_j, y_j , respectively,
- $|q_i - q_j|$ = Euclidian distance of q_i and q_j , calculated by

$$|q_i - q_j| = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2},$$
- n = number of trees observed in the area of interest,
- k_h = core-function as explained below,
- $\square(r)$ = edge bias control function as explained below.

The core function k_h is used to smooth the estimated paircorrelation-function. The idea of k_h is to include these pairs of points whose distance is more or less equal to r , but to weight them the more, the nearer their distance is to r . STOYAN and STOYAN (1992) recommend the so-called EPANECNIKOV-core:

$$k_h(t) = \begin{cases} \frac{3}{4h} \cdot \left(1 - \frac{t^2}{h^2}\right) & \text{if } -h \leq t \leq h \\ 0 & \text{else} \end{cases}$$

The parameter h symbolizes the band-width of the core-function. The choice of h is considered to be more important than the choice of the type of core-function itself. In any case, h should be

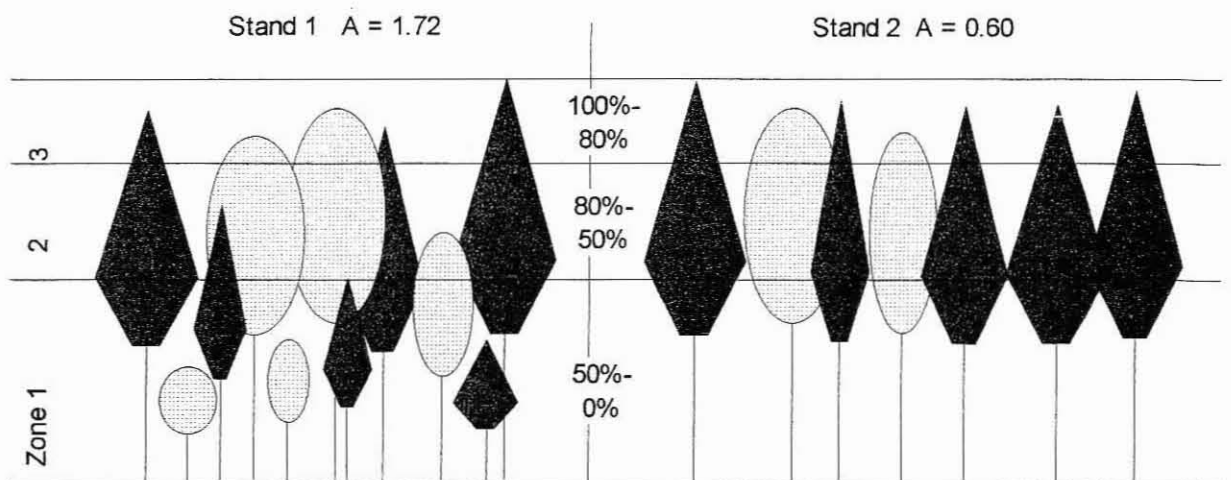


Figure 4: Behaviour of species profile index A (PRETZSCH, 1996) for one multi-layered (stand 1) and one single-layered two-species stand (stand 2).

3.2. Pair- and Markcorrelation-Functions

In contrast to the single-number-indices we dealt with above, the methods shown below do not aggregate structural information to one single number. We get a series of numbers which can be interpreted as estimated values of functions with certain properties. These methods give a much more detailed view on structural aspects than single-number-indices do, their theoretical background and their handling is much more complex, however. In the following text we introduce the so called pair- and mark correlation functions as proposed by STOYAN and STOYAN (1992) and DEGENHARDT (1993).

3.2.1. Pair Correlation Functions

The task of pair correlation functions is similar to the single-number aggregation index R by CLARK and EVANS (1954). Like this index, pair correlation functions regard the horizontal distribution of trees in relation to a random distribution without considering their dimensions. Pair correlation functions enable one to decide, whether pairs of trees at a certain horizontal distance occur more or less frequently than we would expect for a random distribution. These pairs of trees are not automatically next neighbour pairs as we looked for calculating the indices R and S . They are just the pairs of trees which have a certain distance, no matter if they are first, second, third or n th neighbours.

Definition and Properties of Paircorrelation-Functions

We imagine the horizontal distribution of n trees in an area W as one realisation of an isotropic and homogeneous stochastic point process (see STOYAN and STOYAN, 1992). The intensity of this point process is called λ and can be estimated as follows:

$$\hat{\lambda} = \frac{n}{A}$$

where n is the number of trees found in the area W of the size A . Now consider two circles C_1 and C_2 with infinitesimally small areas dF_1 , dF_2 . The distance of their centers is called r . So, the probability $P(r)$ to find one stem-center-point in each one of these two circles can be denoted as:

$$P(r) = g(r) \cdot \lambda \cdot dF_1 \cdot \lambda \cdot dF_2$$

The function $g(r)$ is called pair correlation function.

The paircorrelation-function has some typical properties:

if $g(r) > 1$, then the number of tree-pairs having the distance r is greater than expected for a random (Poisson-) distribution. If $g(r) = 1$, then the number of tree-pairs with the distance r equals the expected number for a random distribution. If $g(r) < 1$, then we observe a lesser number of tree-pairs at the distance r as we would expect for a random distribution. So, if the distribution of all trees in the area W would be a realisation of a Poisson-process, $g(r)$ would be 1 for every r . For big values of r , $g(r)$ comes close to the value 1 in any case.

Estimation of Paircorrelation-Functions

If we want to estimate $g(r)$, we apply the following formula for every $r > 0$ we are interested in:

$$\hat{g}(r) = \frac{1}{2\pi \cdot r \cdot \hat{\lambda}^2 \cdot \gamma(r)} \cdot \sum_{i=1}^n \sum_{\substack{j=1 \\ (j \neq i)}}^n k_h(r - |q_j - q_i|) \quad , \quad r > 0 \quad (1)$$

The symbols used are:

- $\hat{g}(r)$ = estimator for $g(r)$,
- q_i, q_j = Stem center points of tree i and j . They have the coordinates x_i, y_i and x_j, y_j , respectively,
- $|q_i - q_j|$ = Euclidian distance of q_i and q_j , calculated by

$$|q_i - q_j| = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2},$$
- n = number of trees observed in the area of interest,
- k_h = core-function as explained below,
- $\square(r)$ = edge bias control function as explained below.

The core function k_h is used to smooth the estimated paircorrelation-function. The idea of k_h is to include these pairs of points whose distance is more or less equal to r , but to weight them the more, the nearer their distance is to r . STOYAN and STOYAN (1992) recommend the so-called EPANECNIKOV-core:

$$k_h(t) = \begin{cases} \frac{3}{4h} \cdot \left(1 - \frac{t^2}{h^2}\right) & \text{if } -h \leq t \leq h \\ 0 & \text{else} \end{cases}$$

The parameter h symbolizes the band-width of the core-function. The choice of h is considered to be more important than the choice of the type of core-function itself. In any case, h should be

kept dependent on the intensity of the point process as can be done according to STOYAN and STOYAN (1992):

$$h = c \cdot \lambda^{-1/2} \text{ with a recommended range for } c = 0.1 \dots 0.2.$$

$\square(r)$ is needed to avoid plot edge bias, assuming that the point-process observed will be realized also outside the borders of W . The function $\square(r)$ delivers the expected value for the intersection area of W and a copy of W , which is dislocated by a random vector having the length r . For rectangular areas, this function can be written as follows (STOYAN and STOYAN, 1992):

$$\gamma(r) = \begin{cases} \pi - 2x - \frac{2x}{\beta} + \frac{x^2}{\beta}; & 0 \leq x \leq 1 \\ 2 \cdot \arcsin\left(\frac{1}{x}\right) - \frac{1}{\beta} - 2(x-u); & 1 < x \leq \beta \\ 2 \cdot \arcsin\left(\frac{\beta-uv}{x^2}\right) + 2u + \frac{2v}{\beta} - \beta - \frac{1+x^2}{\beta}; & \beta < x < \sqrt{\beta^2+1} \\ 0; & x \geq \sqrt{\beta^2+1} \end{cases}$$

with $x = \frac{r}{\sqrt{\frac{A}{\beta}}}$, $u = \sqrt{x^2-1}$, $v = \sqrt{x^2-\beta^2}$.

Before starting an estimation algorithmus for $g(r)$ based on formula (1), three values have to be determined. These are the lowest and the greatest r , r_{min} and r_{max} , for which $g(r)$ shall be estimated. Finally, the step-width $\square r$, which is the distance of the estimated values for $g(r)$ on the r -axis, must be known. According to DEGENHARDT (personal communication) good results can be obtained by settings like in the following example:

$$\begin{aligned} r_{min} &= d_{min} - 0.001 \cdot \lambda^{-\frac{1}{2}}, \text{ where } d_{min} \text{ is the lowest tree-distance observed,} \\ r_{max} &= 1.5 \cdot \lambda^{-\frac{1}{2}}, \\ \Delta r &= \frac{r_{max} - r_{min}}{200}. \end{aligned}$$

The estimation algorithmus now has to start by estimating $g(r)$ according to formula (1) beginning with r_{min} as r , after that incrementing r_{min} by $\square r$ estimating $g(r)$ for the new r and so on, until r_{max} is reached. After that, $g(r)$ can be plotted against r . The graph obtained is used to interpret the estimated paircorrelation-function. Two examples for typical shapes of $g(r)$ are depicted in figure 5.

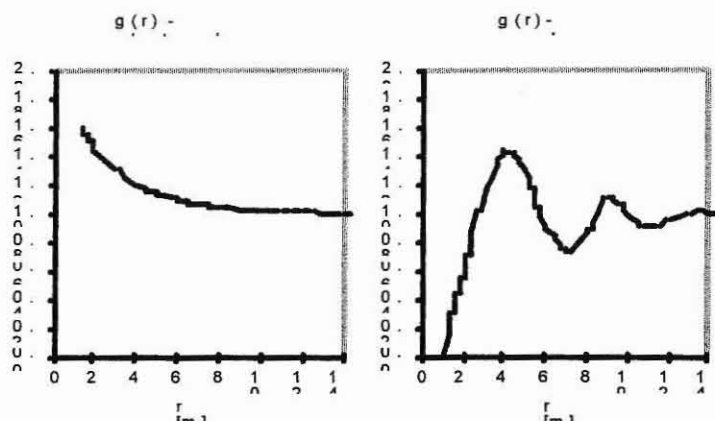


Figure 5: Typical shapes for empirical paircorrelation-functions. Left chart: clustered tree-distribution, right chart: regular tree-distribution.

The left graph in figure 5 shows a typical pair correlation function for a clustered tree distribution as we observe the highest values of $g(r)$ at low tree-distances r . That means, tree pairs with distances around 1 up to 3 m occur far more often than they would do if trees were distributed randomly. The right graph shows $g(r)$ as observed for considerable regular patterns. At distances up to 1 m $g(r)$ is 0, so we do not observe tree-pairs at these distances at all. So we can conclude, that the observed tree distribution is a realisation of a so called "hard-core" process. At distances of approximately 4 m $g(r)$ shows a sharp maximum which can be interpreted in the way, that most next-neighbour-pairs are found at this distance. The following local minimum is due to the space between first and second-next neighbour while the next local maximum at somewhat 9 m occurs as most second-neighbour-pairs have this distance.

3.2.2. Mark Correlation Functions

Mark correlation functions have similar properties to the pair correlation function discussed above. Compared to the latter they regard not only the coordinates of the examined objects, but also a so-called mark which is attributed to every object describing the realisation of a certain property for every object. Applying this concept to a given tree stand, one could consider any tree-dimension as a mark, for example the dbh. A mark correlation function shows for point-pairs of a given distance r , if and how the average dbh-combination of such point pairs deviates from the expected one for an independent spatial distribution of the dbh. For example, considering an older forest stand, we could expect to frequently find pairs of trees which both have big dbh's only at certain distances, considerably greater than the minimum tree distance, as a result of competition processes.

Definition and properties of markcorrelation-functions

The characteristic element of mark correlation functions is a non-negative test function $f(m_1, m_2)$ which is dependent on the marks m_1 and m_2 of the points q_1 and q_2 . There are many possibilities to construct such test functions (see STOYAN and STOYAN 1992), but when regarding object dimensions, as we do here, the function

$$f(m_1, m_2) = m_1 \cdot m_2$$

proves to be suited very well (see PENTTINEN et al. 1992). This function takes highest values, when both marks are big. Consider again two infinitesimally small circles C_1 and C_2 with the areas dF_1 and dF_2 . The distance between their centers is r . Consider $Z(r)$ as a random variable which takes the value $f(m_1, m_2)$ if there is a point in both circles and the value zero otherwise. The expected value for $Z(r)$, $E(Z(r))$ can be denoted then as follows

$$E(Z(r)) = k_f(r) \cdot g(r) \cdot \bar{m}^2 \cdot \lambda dF_1 \cdot \lambda dF_2$$

where $k_f(r)$ is called the mark correlation function, $g(r)$ is the paircorrelation function as defined above. Additionally, we have to take into account the expected value for $f(m_1, m_2)$ if the marks would be distributed independently, which is, in our special case, when $f(m_1, m_2) = m_1 \times m_2$, the square of the arithmetic mean of all marks \bar{m}^2 . \square denotes, as showed above, the point density of the observed point process. $k_f(r)$ can be regarded as the expectation for $f(m_1, m_2)$, standardized by \bar{m}^2 , with the constraint, that there is one point located in each one of the two circles C_1 and C_2 . The mark correlation function takes values greater than 1 if the expectation for $f(m_1, m_2)$ at a given distance r is higher than we would expect if the marks were distributed independently. If $k_f(r)$ takes values lesser than 1, we would expect a greater average of $f(m_1, m_2)$ if we had an independent mark distribution. For completely independently distributed marks we get an overall value of 1 for $k_f(r)$. Like the pair correlation function the mark correlation function approaches the value 1 for big values of r . To avoid non-defined arithmetical operations, we define $k_f(r)$ being zero if there are no pairs of points found at a given distance r .

Estimation of Markcorrelation-Functions

$k_f(r)$ can be estimated by applying the following formula for every $r > 0$ which is of interest.

$$k_f(r) = \frac{\sum_{i=1}^n \sum_{j=1, j \neq i}^n m_i \cdot m_j \cdot k_h(r - \|q_i - q_j\|)}{\bar{m}^2 \cdot \sum_{i=1}^n \sum_{j=1, j \neq i}^n k_h(r - \|q_i - q_j\|)}, \quad r > 0 \quad (2)$$

Like above the symbols used are:

$\hat{k}_f(r)$ = Estimator for $k_f(r)$,

q_i, q_j = Stem center points of tree i and j . They have the coordinates x_i, y_i and x_j, y_j , respectively,

$\|q_i - q_j\|$ = Euclidian distance of q_i and q_j , calculated by

$$\|q_i - q_j\| = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2},$$

n = number of trees observed in the area of interest,

k_h = core-function as explained above,

The estimation procedure is identical to the one for the pair correlation function, when formula (2) is used instead of formula (1). The recommendations for the choice of r_{min} , r_{max} and $\square r$ are

also the same. How to interpret mark correlation functions is shown below, when they are applied to our example stands.

4. Results

4.1. Aggregation Index R by CLARK AND EVANS (1954)

Figure 6 shows the results obtained for the aggregation index R. In the natural forest all groups of species regarded isolated from the others show a more or less evident clustered distribution, as all values obtained for R are below 1. The groups *Croton* and *Azelia* seem to be clustered in a most pronounced manner. Also, all trees in the natural forest without dividing into species groups show a considerable tendency to an aggregated distribution, which is, however, not as pronounced as at nearly all species regarded separately. Unexpectedly the index value 0.90 for the 12 year old planted Teak stand does not show a tendency towards a regular distribution.

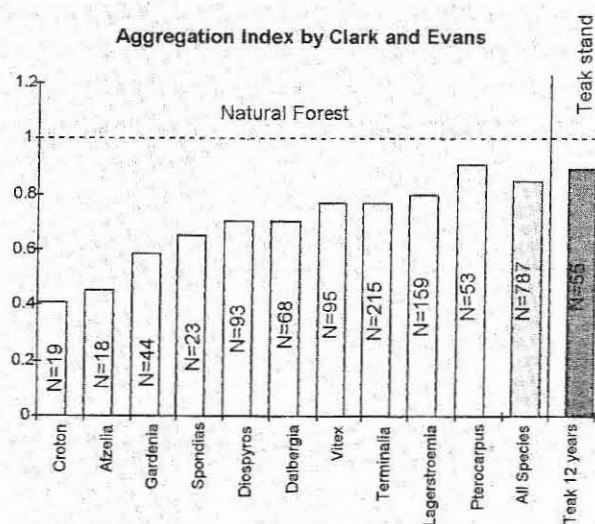


Figure 6: Calculated values of the aggregation index R for the 12 year old teak stand and the tree species groups and all species together in the natural forest. Index values below 1 show a tendency towards an aggregated pattern, index values above 1 indicate a tendency towards a regular tree distribution pattern. N means the number of trees used to calculate the index expressed by the corresponding bar.

4.2. Segregation Index S by PIELOU (1977)

The results obtained by application of the segregation index S can be seen in figure 7. For S can theoretically reach values between -1 and +1 the tendencies indicated seem not to be very pronounced. Anyway, some differences between the single species groups can be observed. S values for most groups are more or less close to zero, but except for *Pterocarpus*, *Spondias* and *Dalbergia* they show a tendency towards segregation from trees belonging to the other species. The most evident tendency towards segregation can be observed for *Croton* and *Gardenia* which show S-index values of 0.13 resp. 0.19. The index value for the pure Teak stand is 1 per

definition, but not shown in figure 7 in order to pronounce the differences between the spec groups in the natural forest. Comparing figure 6 and figure 7 it may be seen, that the high aggregated species groups *Croton* and *Gardenia*, as revealed by the aggregation index R seem be segregated from other species to the highest degree.

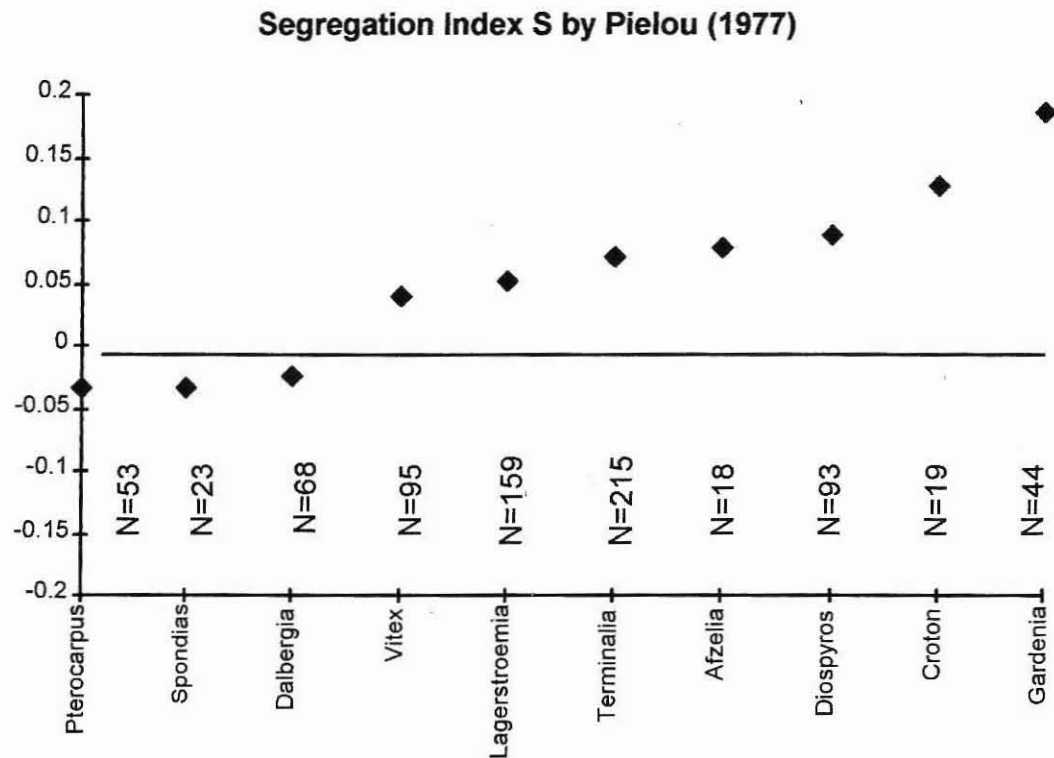


Figure 7: Calculated index values for the segregation index S. Negative values express tendency towards association of the subject species groups with trees belonging to other species groups. Positive values show a tendency towards segregation of the regarded species group from trees belonging to the other groups.

4.3. Species Profile Index A by PRETZSCH (1996)

As can be seen in figure 8 the species profile index A behaves quite differently for the pure teak stand and the natural forest stand. While the teak stand data lead to an index value of only 0.85, we get an A-index of 2.74 for the natural forest. This was expected, considering the multi-layered multispecies-built natural forest in comparison to the mono-layered pure teak stand.

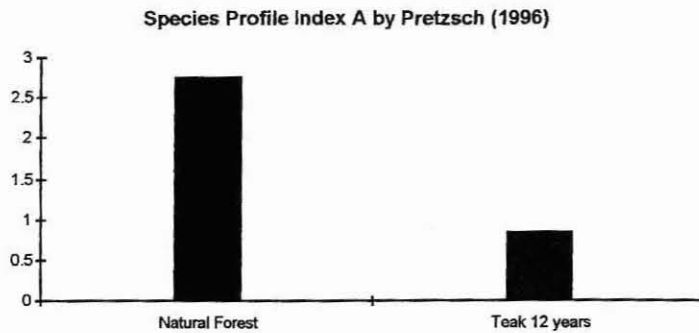


Figure 8: Obtained index values for the species profile index A. Higher index values indicate higher biodiversity and more pronounced vertical distribution of the trees.

4.4. Pair- and Mark Correlation Functions

Although we introduced the concepts of pair- and mark correlation functions separately from each other before, it seems here quite reasonable to interpret their results obtained for our data together. Calculating the mark correlation functions we used the dbh as mark and $f(m_1, m_2) = m_1 \times m_2$ as test function. Figure 9 shows the empirical functions for the 12 year old teak stand data.

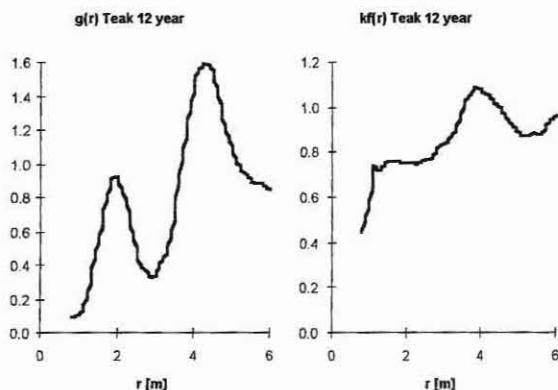


Figure 9: Empirically found pair- (left chart) and mark correlation function (right chart) for the 12 year old teak stand.

The pair correlation-function gives a clear sight on the horizontal tree distribution of the 12 year old teak stand. The point process realized seems to be a more or less "hard-core" one, as we do not observe considerable numbers of tree-pairs at distances below 1 m. Further, the pair correlation function shows a first peak for distances about 2 m and a second, even higher peak for distances about 4 m. This reflects very well the planting pattern applied when the plot was established 12 years ago, maintaining tree distances of 2 m within the rows and 4 m between the rows. In the meantime, thinning and competition led to a disappearance of many trees at low distances from their next neighbours (see figure 3), so we observe the 4 m peak being much higher than the 2 m peak, as most of the trees have a first or second neighbour at a distance about 4 m, but not all trees have a neighbour at 2 m. In general, the pair correlation function indicates -

in contrast to the aggregation index R calculated before- a quite regular pattern. This apparent contradiction will be discussed below.

The mark correlation function has only one single peak at radii about 4 m which indicates an average dbh-combination greater than expected for dbh's distributed independently. At low radii we observe the opposite, as the diameter-products of tree pairs decrease as their distance decrease. This means, that pairs of thick, dominant trees have not established themselves at distances much lower than 4 m. The mark correlation function's peak for distances about 4 m is not only maintained by competition effects, but also by the planting pattern, evidently. Tree pairs at lower distances seem to consist mostly of smaller, non-dominant trees in the stand.

Regarding the natural forest data, it would be beyond the limits of this paper to show the pair and mark correlation functions for all observed species groups. So, we limit ourselves to interpret these functions only for all species together and some selected species groups showing interesting or typical patterns.

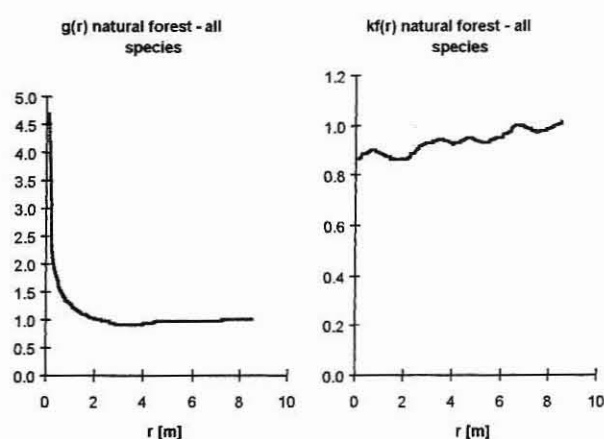


Figure 10: Empirically found pair- (left chart) and mark correlation function (right chart) for all species of the natural forest stand.

The pair correlation function for all species of the natural forest stand (figure 10, left chart) indicates a typical clustered tree distribution pattern as detected also by the aggregation index I . Pairs of trees at distances between nearly 0 m and about 1 m are far more abundant than would be expected for a random pattern. Above distances about 1 - 2 m the curve approaches very soon a constant value of 1, indicating no more correlation between pair-distance and abundance of pairs. Applied to the same data, the mark correlation function shows a very balanced behavior in general. Although its values for distances lower than 7 m are somewhat lower than 1, the effect cannot be considered very sharp compared to the planted teak stand (figure 9, right chart). This means, that tree dimensions in our natural forest stand are distributed independently to a certain degree, including considerable numbers of pairs of dominating trees even at very low distances. Very similar pair- and mark correlation functions were obtained for the most abundant species groups in this stand, *Terminalia* and *Lagerstroemia* as expected.

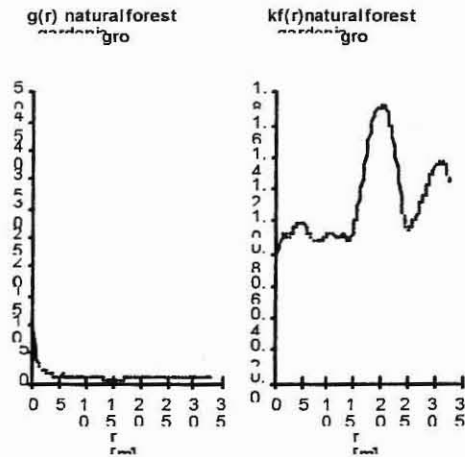


Figure 11: Empirically found pair- (left chart) and mark correlation function (right chart) for the *Gardenia* species group of the natural forest stand.

Regarding the *Gardenia* species group (figure 11), the pair correlation function indicates an even more evident clustering than observed for all species (figure 10). As we remember, the aggregation index R revealed the same tendency. The mark correlation function, however, forms an evident peak around tree pair distances of 20 m. So, we can conclude that strong *Gardenia* trees do occur mostly at distances about 20 m in the natural forest stand.

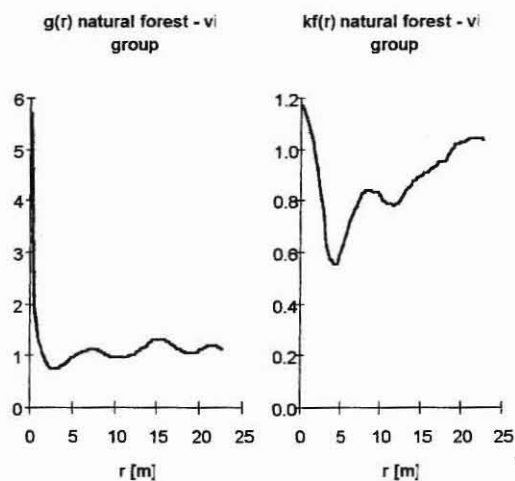


Figure 12: Empirically found pair- (left chart) and mark correlation function (right chart) for the *Vitex* species group of the natural forest stand.

Compared to *Gardenia* the *Vitex* species group obviously shows a totally different behaviour (figure 12). Although the pair correlation function indicates a considerable clustering which could also be shown by the aggregation index R , the mark correlation function forms a totally different shape. *Vitex* tree-pairs, where both trees have considerable diameters, occur at the smallest distances observed. This may be due to ecological peculiarities, but also to trees having a bifurcation lower than the dbh measurement height, from which two dbh's and two stem center points are reported. At distances around 5 m we find a sharp minimum indicating only pairs with each tree having small diameters. For distances greater than 5 m the function raises and finally

approaches values around 1. Back to the pair correlation-function we can see another interesting aspect of the horizontal tree distribution of *Vitex*: at distances about 8 m and 16 m there are two slight peaks whose extent is partly concealed by the scale of the $g(r)$ -axis. These 'wiggles' indicate, that, in addition to the clustering at low distances there is a certain regularity at high distances, perhaps in the pattern, the clusters are distributed according to. Such effects evidently cannot be detected using single number indices.

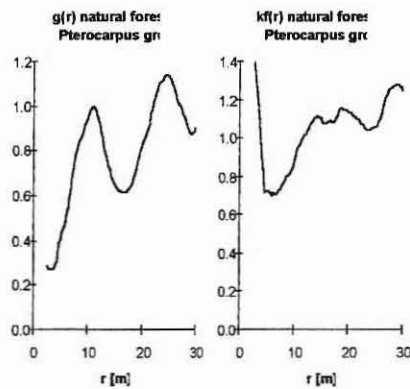


Figure 13: Empirically found pair- (left chart) and mark correlation function (right chart) for the *Pterocarpus* species group of the natural forest stand.

As last species stratum we analyse the *Pterocarpus* group (figure 13). Similar to the young Teak stand but not as pronounced, we observe two sharp local maxima of the pair correlation function: the first one at about 12 m, the second and even higher one at about 25 m. This indicates a considerable regular tree distribution pattern. As the maxima are less sharp as they are for the young Teak stand, we can conclude that the degree of regularity is lower for the *Pterocarpus* group. In contradiction, the aggregation index R calculated for *Pterocarpus* group does not indicate regularity. A rather similar contrast between this single-number index and the pair correlation function was observed already for the Teak stand. The mark correlation function forms a similar shape as it does for *Vitex* group (figure 12). At the lowest tree-distances observed we tendentially find the pairs of the strongest trees. A second, relatively pronounced local maximum can be observed at distances of approximately 30 m, so, the strongest *Pterocarpus* trees seem to stand either very near or very far from each other.

5. Discussion

When we consider the single-number indicators calculated for horizontal stand structure, R and R' , the results obtained seem to be quite reasonable. The different pronounced, but for every species group in the natural forest observed tendency towards an aggregated distribution of the trees could be due to distinct micro-site conditions on the plot. On such micro-sites certain species clearly show a quite different vitality, which leads to a clustered occurrence of the species at the preferred sites. Also fruit dispersion mechanisms must be taken into account. Tropical forest ecologists surely could discuss the results found here according to the recent knowledge about

the ecology of the tree species-groups we dealt with, but as the point of this study is a methodical one, we will not proceed in this direction, although it would be highly interesting in another context. Regarding the segregation index *S*, most species are found to be slightly segregated from the collective of all other species. This is going well together with the aggregation tendencies found by calculating the index *R*, especially as the species groups *Croton* and *Gardenia*, which are detected by *R* as the first and third most aggregated groups, are at the same time the groups for which the most pronounced tendency towards segregation is found. The fact, that these tendencies revealed by the indices *R* and *S* are not easy to recognize visually in the tree distribution pattern itself (figure 2) underlines the necessity to deal with numerical methods if we want to receive a distinguished view on spatial stand structure.

Unexpectedly, we found an aggregation index less than one for the 12 year old teak stand, although it was setup as a plantation and the plant rows still can be seen in figure 3. Therefore we had expected an aggregation index greater than 1, indicating a tendency towards a regular tree distribution. The aggregation index of 0.90 is in fact due to the planting pattern applied, where the distances between the rows are far larger than the tree distances within the rows (figure 3). So, the next-neighbour-distances will be found only within the rows, which are regarded as a type of clusters from this sight. This obviously renders the slight tendency towards clustering found for the young Teak stand, when applying the aggregation index *R*.

In the context of the DSE training course where this case study was embedded, it was not possible to perform significance tests for the tendencies found by the indices *R* and *S*. As the tree numbers for the species groups spread widely (figure 1) we have to suspect, that at least the tendencies found for groups covered only by few trees cannot be proved statistically. So, for further studies it is recommended to apply statistical tests on the indices *R* and *S* as can be found at PRETZSCH (1993) and QUEDNAU and FRÖHLICH (1994).

The species profile index *A* behaves just as expected. Compared to the mono-layered pure Teak stand the natural forest shows a more than three times higher value for *A*. As the example stand, shown in the left hand side of figure 4, representing a considerable rich structured forest in temperate regions, shows an index value for *A* of 1.72, the *A*-index value of 2.74 for the natural tropical forest gives an impression of the high biodiversity and the species richness and diversity typical for tropical forests (BOONTAWEE et. al. 1995). Interpreting this value we have to consider additionally, that we had aggregated the species of the natural forest to ten groups, on which structure analysis was applied. If we would have taken into account the real number of about 70 species we would obviously have obtained a even higher value for *A*, so the index calculated here indeed underestimates the real diversity of the natural forest stand.

Being so-called 'second-order characteristics' (PENTTINEN et al., 1992) pair- and mark correlation functions give a much more detailed insight into spatial stand structure and the ecological peculiarities related with, than single-number indices can do. In contrast to the latter ones we are able to see not only the general structural tendencies but also their scale, which may be important for example, when the spatial ecology of tree species shall be investigated or when a forest should be classified as a habitat for certain animals. Additionally, these functions reveal even different types of distribution patterns occuring at different scales, as we could observe regarding the pair correlation function for the *Vitex* species group in the natural forest (figure 12). As we remember, applying the pair correlation function and the aggregation index *R* to the pure Teak stand data lead to contradictory results. The former clearly suggested a regular, the latter a

slightly clustered tree distribution pattern. We explained above in this discussion, that the aggregation index R is forced by its concept, the next-neighbour-method, to interpret such patterns as clusters, where the tree distances within the rows are far smaller than the distances between the rows. The pair correlation function is not trapped by such special patterns and indicates regularity, according very well to our intuitional idea of regular point patterns. For the *Pterocarpus* species group in the natural forest, we seem to encounter a quite similar effect where the aggregation index R clearly shows a certain weakness due to its simplicity. So, when performing structural analyses, it is a good recommendation to apply more than one method and to interpret the results synergetically.

Finally it is important to state, that the five methods proposed in this study are only some example parameters of the high dimensional vector 'stand structure'. There are much more methods and possibilities than could be shown in this paper. The interested reader will find many hints in the papers cited below. As one of many examples it seems to be a most interesting topic to apply pair- and mark correlation functions for different-species-pairs to get deeper insight into the complex ecology of tropical forest ecosystems or agroforest systems. The results could be of very practical silvicultural and ecological use. The aim of this study is reached, if the advantages of using numerical methods for characterizing forest stand structure could be shown and if it leads to further investigations.

6. Literature

- AMMER, U., DETSCH, R., SCHULZ, U. 1995: Konzepte der Landnutzung, Forstwissenschaftliches Centralblatt, Vol. 114, pp. 107-125
- BOONTAWEE, B., PLENGKLAI, C., KAO-SA-ARD, A. 1995: Monitoring and Measuring Forest Biodiversity in Thailand. Pp. 113-126 In: BOYLE, T.J.B. and BOONTAWEE, B. 1995: Measuring and Monitoring Biodiversity in Tropical and Temperate Forests: Proceedings of IUFRO Symposium. Held at Chiang Mai, Thailand. August 27th-September 2nd, 1994. Center for International Forest Research. Indonesia. 395 p.
- BLAB, J. 1986: Grundlagen des Biotopschutzes für Tiere. Ein Leitfaden zum praktischen Schutz der Lebensräume unserer Tiere, Kilda-Verlag, Bonn, Bad Godesberg, 257 p.
- CLARK, P.J. AND EVANS, F.C. 1954: Distance to nearest neighbour as a measure of spatial relationships in populations, Ecology, Vol. 35, No. 4, pp.445-453
- DEGENHARDT, A. 1993: Analyse der Entwicklung von Bestandesstrukturen mit Hilfe des Modells der zufälligen Punktprozesse in der Ebene. Beiträge für Forstwirtschaft und Landschaftsökologie, 27 (1993) 4, pp. 182-186.
- DONELLY, K. 1978: Simulation to determine the variance and edge-effect of total nearest neighbour distance, pp.91-95 in: HODDER, I.R. (ed.), Simulation methods in archeology, Cambridge University Press, London.
- ELLENBERG, H. VON EINEM, M., HUDECZEK, H., LADE, H.-J., SCHUMACHER, H.U. SCHWEINHUBER, M., WITTEKINDT, H. 1985: Über Vögel in Wäldern und die Vogelwelt des Sachsenwaldes, Hamb. Avifaun. Beitr., Vol. 20, pp.1-50
- HABER, W. 1982: Was erwarten Naturschutz und Landespflege von der Waldwirtschaft ? Schriftenreihe des Deutschen Rates für Landespflege, Vol. 40, pp.962-965
- KAHN, M. 1998: Modelling Teak Forest Growth and Scenario Analyses. In: MIES, E. (Ed.): Natural and Socioeconomic Analysis and Modelling of Forest and Agroforestry Systems in Southeast Asia. Publication by the German Foundation for International Development, Zschortau, Germany.
- PENTTINEN, A., STOYAN, D., HENTTONEN, H. 1992: Marked Point Processes in Forest Statistics. Forest Science, November 1992, pp. 806-824
- PIELOU, E.C. 1975: Ecological diversity, John Wiley & Sons
- PIELOU, E.C. 1977: Mathematical ecology, John Wiley & Sons, 385 p.

- PRETZSCH, H. 1992: Konzeption und Konstruktion von Wachstumsmodellen für Rein- und Mischbestände, Forstliche Forschungsberichte München, No. 115, 385 p.
- PRETZSCH, H. 1993: Analyse und Reproduktion räumlicher Bestandesstrukturen. Versuche mit dem Strukturgenerator STRUGEN, Schriften aus der Forstlichen Fakultät der Universität Göttingen und der Niedersächsischen Forstlichen Versuchsanstalt, Vol. 114, Sauerländer's Verlag, 87 p.
- PRETZSCH, H. 1996: Strukturvielfalt als Ergebnis waldbaulichen Handelns, Tagungsbericht der Sektion Ertragskunde des Deutschen Verbands Forstlicher Forschungsanstalten, Neresheim, 20.-22.Mai 1996, pp.134-154
- PRETZSCH, H. 1997: Analysis and modeling of spatial stand structures. Methodological considerations based on mixed beech-larch stands in Lower Saxony. Forest Ecology and Management 97 (1997) pp. 237-253
- QUEDNAU, H.-D. AND FRÖHLICH, M. 1994: Statistical analysis of the distribution pattern of natural regeneration in forests, Forest Ecology and Management 73 (1995), pp.45-57
- RIPLEY, B.D. 1977: Modelling spatial patterns, J. Roy. Stat. Soc., Series B, Vol. 39, No. 2, pp.172-212
- RIPLEY, B.D. 1981: Spatial statistics, John Wiley & Sons
- SHANNON, C.E. 1948: The mathematical theory of communication. In SHANNON, C.E., WEAVER, W. (editors): The mathematical theory of communication, Urbana, Univ. of Illinois Press, pp.3-91
- STOYAN, D. and STOYAN, H. 1992: Fraktale Formen und Punktfelder. Methoden der Geometrie-Statistik. Akademie Verlag GmbH, Berlin, 394 p.
- UPTON, G. J. G. AND FINGLETON, B. 1985: Spatial data analysis by example, Volume I: point pattern and quantitative data, John Wiley & Sons, 410 p.
- UPTON, G. J. G. AND FINGLETON, B. 1989: Spatial data analysis by example, Volume II: categorical and directional data, John Wiley & Sons, 416 p.