Distribution of infected oil palms with Ganoderma basal stems root disease

Assis Kamu 1, *, Chong Khim Phin 1, Idris Abu Seman 2, Ho Chong Mun 1

1Faculty of Science and Natural Resource, Universiti Malaysia Sabah, UMS Road, 88400 Kota Kinabalu, Sabah, MALAYSIA
2Ganoderma and Diseases Research for Oil Palm Unit, Malaysian Palm Oil Board, 6, Persiaran Institusi Bandar Baru Bangi, 43000, Kajang, Selangor, MALAYSIA

Abstract: The objective of this study is to spatially identify the pattern of Ganoderma disease under natural field epidemic using three different spatial point pattern analyses, which are nearest neighbor analysis and refined nearest neighbor analysis for first order spatial analysis, and Ripley’s K function for second order spatial analysis. Three commercial oil palm sites with three replicate areas per site were selected in this study with different age of palm tree. The nearest neighbor analysis showed that the spatial distribution of the infected palms in all the replicate areas at SKE0224 was clustered. Refined nearest neighbor analysis and Ripley’s K function also showed that the distribution of the infected palms in all the areas studied was clustered. But statistical test through Monte Carlo simulation showed that the cluster distribution of the infected palms in most of the areas studied was not statistically significant. Nevertheless, the study has proved that the spatial distribution of the infected palms under natural epidemic is not random, disperse or uniform but it is more to cluster pattern. This suggests the spread of the disease could be from tree to tree possibly through root contact as proved by many past studies.

Key words: Oil palm; Ganoderma; Spatial distribution; First order; Second order

1. Introduction

Oil palm or scientifically known as Elaeis guineensis Jacq is a golden crop of Malaysia with total area planted with oil palm is more than five millions hectare (Anonymous, 2013). The industry is well-established but yet facing with some problems arise lately. Disease is one of the major problems faced by oil palm industry. As far as disease is concern, the Basal Stem Rot (BSR) caused by Ganoderma boninense. Ganoderma Basal Stem Rot (BSR) remains the most important disease, present in more than 50% of the oil palm fields in Malaysia (Idris, Mior, Maizatul, & Kushairi, 2011). Ganoderma disease is the most widely studied and knowledge available of oil palm disease in Malaysia (Idris, 2012). There are many control measures or methods have been used or developed in order to minimize the economic loss due to the disease, such as removing or destroying the infected palms, applying some treatments to the infected palms (also called as a curative control), or protecting the healthy or young palms from being infected (also called as a preventive control). Currently, there is still no effective cure for Ganoderma infections in an existing stand (Lim et al., 2012). But there is still limited study on the spatial as well as temporal pattern or distribution of the disease especially under natural field epidemic condition in oil palm plantation (Azahar et al., 2011).

Identification of the type of disease pattern and spread in a field is critical in epidemiological investigations. It can help the authorities in selecting a strategy to combat the outbreak. A random pattern of infected plants suggest that, at the time of observation, the pathogen is not spreading from plant to plant. Conversely, aggregations (clusters) of infected plants suggest that the pathogen is spreading from plant to plant within a field (Campbell & Madden, 1990; Suriya Rao et al., 2007). Epidemics of plant disease vary both in time and space. It is widely accepted by plant pathologists and epidemiologists that the spatial component of plant disease epidemics is as important as the temporal component. Spatial pattern refers to the arrangement of disease entities relative to each other.

It may be classified into three categories, namely uniform, random, and clustered or clumped (Campbell & Madden, 1990). In a uniform pattern there is a regular arrangement of infected and healthy plants. ‘Random’ means that all distinguishable arrangements of infected and healthy plants are equally possible, or, in other words, that in every point in the crop there is the same probability of a plant being infected. That means there is no correlation exists between locations of the infected plants. In a clustered pattern, every plant in the field does not have an equal probability of being infected so that a diseased plant increases the probability of nearby plants being infected. Subgroups of infected plants tend to be significantly closer to each other than to other subgroups of infected plants. Information on the spatial pattern of pathogens and plant diseases can
be useful in various ways, such as to identify sampling plan and to identify the suitable analytical models.

Knowledge of whether infection occurs at random is important to an understanding of the epidemiology of infection by soil borne pathogens (Gilligan, 1983). The need to determine if clusters or random exist in a given data set has prompted the development of several techniques that come from a field of study known as point pattern analysis. There are many point pattern analysis or spatial analytical techniques that can be used to identify the spatial pattern of plant disease, and geospatial analytical technique is one of them which is based on geographical epidemiology (the geographical incidence of disease) (Gatrell et al., 1996). The simplest theoretical model for a spatial point pattern is that of complete spatial randomness (CSR), in which the events are distributed independently according to a uniform probability distribution over the region.

Geostatistical technique is one of the methods used in oil palm to quantitatively characterize the spatial pattern of BSR disease (Azahar et al., 2011; Idris et al., 2009; Idris et al., 2010; Somayeh et al., 2012). This technique is conducted by plotting the data into geographical information system (or GIS) map before conducting spatial and geostatistical analyses. Many studies have confirmed that the spatial pattern of BSR disease in oil palm plantation was random even after controlling some of the confounding factors such as planting density (Azahar et al., 2011). This finding confirms that the spread of BSR disease is not from tree to tree.

This paper is an extension of the previous published paper with different spatial point pattern analysis and with additional one oil palm study plot (Assis et al., 2015). The objective of this study is to spatially identify the pattern of Ganoderma disease under natural field epidemic by using three spatial point pattern analyses, which are nearest neighbour analysis and refined nearest neighbour analysis for first order spatial analysis, and Ripley’s K function for second order spatial analysis.

2. Materials and methods

2.1. Study sites

The production phase of oil palm can further be divided into three, namely step ascent phase (3 to 10 years after planting) where the yield is in increasing phase, plateau phase (10 to 15 years after planting) where the yield is in flat or maximum phase, and declining phase (older than 15 years after planting) where the yield is in declining phase (Schmidt, 2007). In order to cover all of these three phases, three commercial oil palm sites were selected in this study. The details of each of the study sites are as shown in Table 1. All of the study sites are under management of the same oil palm company and are located in the same district which is Tawau district of Sabah, Malaysia. Three replicate areas per study site with the size of 50,625 m² per area were selected for the purpose of spatial point pattern analysis of the infected palms with Ganoderma disease.

<table>
<thead>
<tr>
<th>Table 1: Descriptions of the study sites</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
</tr>
<tr>
<td>Production phase</td>
</tr>
<tr>
<td>ID</td>
</tr>
<tr>
<td>Location (latitude)</td>
</tr>
<tr>
<td>Longitudes</td>
</tr>
<tr>
<td>Year of planting (Age)</td>
</tr>
<tr>
<td>Variety of oil palm</td>
</tr>
<tr>
<td>Planting density</td>
</tr>
<tr>
<td>Topography</td>
</tr>
<tr>
<td>Soil type</td>
</tr>
<tr>
<td>Previous crop</td>
</tr>
<tr>
<td>Planting generation</td>
</tr>
</tbody>
</table>

2.2. Disease detection method

The assessment of Ganoderma disease in this study was based on the single point disease assessment. By following the standard operating procedure (SOP) of Ganoderma census published by Malaysian Palm Oil Board (Anonymous, 2014), the census has been conducted based on the external visible symptoms of the disease, which are presence of basidiocarps or fruiting body of Ganoderma at the bottom or in the lower part of the trunk, presence of unopened new fronds (spears), wilting of green fronds hanging downward like a skirt, yellowing fronds due to nutrient deficiency, and small canopy due to production of smaller fronds (Chung, 2011). The classification of the disease severity is as shown in Table 2. There were no treatments being applied in the study sites to control the disease.
Table 2: Ganoderma disease severity index (DSI) by MPOB

<table>
<thead>
<tr>
<th>DSI</th>
<th>Label</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>Uninfected</td>
<td>Palm showing foliar symptoms and stem rotting at the base. Using early detection methods (e.g. GSM or PCR-DNA) showing no (negative) Ganoderma</td>
</tr>
<tr>
<td>R2</td>
<td>Mild</td>
<td>Presence of white mycelium or fruiting body (e.g. small white button form) when using early detection methods (e.g. GSM or PCR-DNA) showing positive Ganoderma. No foliar symptoms and slightly or no stem rotting (&lt;10%) at the base</td>
</tr>
<tr>
<td>R3</td>
<td>Moderate</td>
<td>Presence of white mycelium or fruiting body (e.g. small white button or bracket shape form). Palm showing foliar symptoms (&lt;50%) and slightly stem rotting (&lt;30%) at the base. Confirmed presence of Ganoderma fungus using early detection methods (e.g. GSM or PCR-DNA)</td>
</tr>
<tr>
<td>R4</td>
<td>Severe</td>
<td>Presence of white mycelium or fruiting body (e.g. small white button or bracket shape form). Palm showing foliar symptoms (&gt;50%) and stem rotting (&gt;30%) at the base. Confirmed presence of Ganoderma fungus using early detection methods (e.g. GSM or PCR-DNA)</td>
</tr>
<tr>
<td>R5</td>
<td>Very severe</td>
<td>Presence of white mycelium or fruiting body (e.g. small white button or bracket shape form). Palm dead/collapsed showing severe foliar symptoms and stem rotting at the base. Confirmed presence of Ganoderma fungus using early detection methods (e.g. GSM or PCR-DNA)</td>
</tr>
</tbody>
</table>

Source: (Anonymous, 2014)

2.3. Spatial point pattern analysis

There are three spatial point pattern analyses used in this study, namely Nearest Neighbor Analysis, Refined Nearest Neighbor Analysis (Boots & Getis, 1988), and Ripley’s K-Function (Dixon, 2002).

2.3.1. Nearest neighbor analysis

This analysis examines the distances between each point and the closest point to it. It then averages all these nearest neighbour distances. If the average distance is less than the average for a hypothetical random distribution or CSR (complete spatial randomness) pattern, the distribution of the features being analysed are considered clustered. If the average distance is greater than a hypothetical random distribution, the features are considered dispersed. CSR describes the pattern of points that would occur by chance in a completely undifferentiated environment (Fischer & Getis, 2010). CSR is generated by means of two assumptions: 1) that all places are equally likely to be the recipient of a case (event) and 2) all cases are located independently of one another. The sample mean of nearest-neighbour distance is calculated by using Equation 1.

\[ \overline{d} = \frac{1}{N} \sum_{i=1}^{N} d_i \]  

(1)

where \( d_i \) is the nearest-neighbour distance for point \( i \), and \( N \) is the number of points. The expected average distance for a random pattern is calculated by using Equation 2.

\[ E(\overline{d_x}) = \frac{1}{2\sqrt{\lambda}} \]  

(2)

where \( \lambda \) is density and calculated by using Equation 3.

\[ \lambda = \frac{N}{A} \]  

(3)

where \( N \) denotes the size of area studied. The variance of \( \overline{d} \) is calculated as Equation 4 and its standard deviation is calculated as shown in Equation 5.

\[ \text{var}(\overline{d}) = \frac{4 - \pi}{4\lambda \pi} \]  

(4)

\[ \sigma(\overline{d}) = \sqrt{\text{var}(\overline{d})} = \sqrt{\frac{4 - \pi}{m4\pi\lambda}} \]  

(5)

where \( \pi \) is the constant 3.141593. The null hypothesis of CSR is tested using the Z statistic (standard normal variate). A negative Z score indicates clustering and a positive score indicates dispersion or evenness. The Z statistic is calculated using the formula Equation 6.

\[ Z = \frac{\overline{d} - E(\overline{d})}{\sigma(\overline{d})} \]  

(6)

2.3.2 Refined nearest neighbor analysis

This analysis is used to test the null hypothesis of CSR. In nearest neighbour analysis, a few very large nearest neighbour distances associated with isolated points could obscure an otherwise clustered pattern. Therefore, refined nearest neighbour analysis overcomes this issue by examining the entire distribution of nearest neighbour distances (Fischer and Getis, 2010). If for each distance, the observed nearest neighbour distances higher, \( F(d_i < r) \), larger than the expected nearest neighbour distances for CSR, \( P(d_i < r) \), a clustered pattern is indicated, and when the observed nearest neighbour distances, \( F(d_i < r) \), lower than the expected nearest neighbour distances for CSR, \( P(d_i < r) \), a regular pattern is indicated. \( F(d_i < r) \) is obtained by taking the nearest neighbour distances, \( d_i \), and the nearest distances to study boundary, \( u \), for each point \( i \). The value of \( d_i \) then be ranked from the smallest to the largest. For every distance of interest, \( r \), the number of points, \( n_i \), for which \( d_i < u \), and the number of points, \( n_r \), for which \( u < r < d_i \), are counted. Then, the proportions of
2.3.3. Ripley’s K-function

One problem with nearest neighbour analysis is it examines only one scale of interaction at a time. Most commonly this technique detects clustering at short distances. Ripley’s K function can be computed over a range of distances and be used to identify the scales over which clustering occurs (Fischer & Getis, 2010). Ripley’s K-Function is a tool for analysing completely mapped spatial point process data. K-function is also called second-order analysis to indicate that the focus is on the variance, or second moment, of pairs of intervened distances. K-function analysis is also a test of the hypothesis of CSR. It considers all combinations of pairs of points. It compares the number of observed pairs with the expectation at all distances based on a random spatial distribution of points (Dixon, 2002). The density of points, the borders, and the size of the sample are taken into consideration. The expected value of \( L(d) \) under CSR is \( d^2 \). The confidence interval in this analysis is generated by examining the specified number of permutations of randomly generated patterns of \( N \) points over the whole study area. If for any distance, the observed \( L(d) \) falls above or below the expected \( L(d) \) the null hypothesis of CSR can be rejected at an appropriate level of significance. The level of significance is determined by the confidence envelope which is determined through Monte Carlo simulations under an appropriate null hypothesis. An observed \( L(d) \) below the envelope indicates that the points are dispersed at that distance, whereas an observed above the envelope indicates that clustering is present at that distance. The observed \( L(d) \) is calculated as shown in Equation 9.

\[
L(d) = \sqrt{\frac{\sum_{i=1}^{N} \sum_{j=1}^{N} k(i, j)}{\pi(N(N - 1))}}
\]

where \( A \) is the study area, \( N \) is the number of points, \( d \) is the distance, \( \sum_{i=1}^{N} \sum_{j=1}^{N} k(i, j) \) is the number of \( j \) points within distance \( d \) of all \( i \) points, and \( k(i, j) \) is the weight, which is estimated by Equation 10 if no edge corrections, by Equation 11 for border correction if a point \( i \) is closer to one boundary than it is to point \( j \), or by Equation 12 for weighting if a point \( i \) is closer to two right angle boundaries than it is to point \( j \).

\[
k(i, j) = \begin{cases} 
1 & \text{in case } d(i, j) \leq d \\
0 & \text{Otherwise} 
\end{cases}
\]

(10)

\[
k(i, j) = 1 - \frac{\cos^{-1} \frac{e}{d(i, j)}}{\pi}^{-1}
\]

(11)

where \( e \) is the distance to the nearest edge.

\[
k(i, j) = 1 - \frac{\cos^{-1} \frac{e_1}{d(i, j)} + \cos^{-1} \frac{e_2}{d(i, j)} + \frac{\pi}{2}}{2\pi}^{-1}
\]

(12)

where \( e_1 \) and \( e_2 \) are the distances to the nearest vertical and horizontal borders respectively.

3. Results and discussion

3.1. Nearest neighbor analysis

<table>
<thead>
<tr>
<th>Study site</th>
<th>Replicate area</th>
<th>Total number of points, ordered mean,</th>
<th>Expected mean, Variance, Z-value</th>
<th>Spatial pattern distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>MBE0702 A</td>
<td>Area 1</td>
<td>61 14.8222</td>
<td>14.6304 1.1153 0.1817 Dispersed</td>
<td></td>
</tr>
<tr>
<td>Area 2</td>
<td>63 14.3670</td>
<td>14.0798 0.9987 0.2874 Dispersed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Area 3</td>
<td>51 16.8738</td>
<td>16.0911 1.6274 0.6136 Dispersed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SKE0224 A</td>
<td>Area 1</td>
<td>37 16.7194</td>
<td>18.3003 2.9477 -0.9208 Clustered</td>
<td></td>
</tr>
<tr>
<td>Area 2</td>
<td>33 19.3761</td>
<td>20.3330 4.1028 -0.4724 Clustered</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Area 3</td>
<td>28 17.5278</td>
<td>21.2903 5.3458 -1.6273 Clustered</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDE8718 A</td>
<td>Area 1</td>
<td>76 14.0049</td>
<td>13.0263 0.7026 0.1371 Dispersed</td>
<td></td>
</tr>
<tr>
<td>Area 2</td>
<td>101 12.1588</td>
<td>10.9848 0.3714 1.9264 Dispersed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Area 3</td>
<td>120 12.5026</td>
<td>12.2560 0.2706 4.3189 Dispersed</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3 shows the results of nearest neighbour analysis for all of the study areas. Three out of nine areas show negative values of Z. These negative values indicate that the spatial distribution pattern
of the infected palms in these three areas (i.e. area 1, area 2, and area 3 at SKE0224) is clustered. A negative Z score indicates clustering and a positive score indicates dispersion or evenness. All the other study areas located at MBE0702 and MDE8718 showed positive value of Z which indicates the spatial distribution of the infected palms in those areas is dispersed. The results show that there is a mixed result. Cluster distribution of the infected palms in the study area of SKE0224 could justify that the primary route of infection of BSR disease appears to be through root contact with inoculum sources in the soil (Rees et al., 2009). But the results of MBE0702 and MDE8718 suggested that Ganoderma spread probably through basidiospore dispersal at greater distance, instead of root to root infection (Mohd Rashid et al., 2014).

### 3.2. Refined nearest neighbour analysis

<table>
<thead>
<tr>
<th>Study site</th>
<th>Replicate area</th>
<th>The largest absolute distance, $d_r$,</th>
<th>At distance, $r$</th>
<th>Significance level</th>
<th>Spatial pattern distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>MBE0702</td>
<td>Area 1</td>
<td>0.164</td>
<td>9</td>
<td>21</td>
<td>0.340</td>
</tr>
<tr>
<td></td>
<td>Area 2</td>
<td>0.237</td>
<td>9</td>
<td>28</td>
<td>0.080</td>
</tr>
<tr>
<td></td>
<td>Area 3</td>
<td>0.257</td>
<td>9</td>
<td>21</td>
<td>0.100</td>
</tr>
<tr>
<td>SKE0224</td>
<td>Area 1</td>
<td>0.237</td>
<td>20.125</td>
<td>19</td>
<td>0.220</td>
</tr>
<tr>
<td></td>
<td>Area 2</td>
<td>0.249</td>
<td>20.125</td>
<td>16</td>
<td>0.250</td>
</tr>
<tr>
<td></td>
<td>Area 3</td>
<td>0.246</td>
<td>12.728</td>
<td>9</td>
<td>0.310</td>
</tr>
<tr>
<td>MDE8718</td>
<td>Area 1</td>
<td>0.116</td>
<td>12.728</td>
<td>34</td>
<td>0.690</td>
</tr>
<tr>
<td></td>
<td>Area 2</td>
<td>0.233</td>
<td>9</td>
<td>61</td>
<td>0.020</td>
</tr>
<tr>
<td></td>
<td>Area 3</td>
<td>0.166</td>
<td>9</td>
<td>62</td>
<td>0.150</td>
</tr>
</tbody>
</table>

*Significant at 0.05 level of significance

Table 4 shows the results of refined nearest neighbour analysis for all of the areas studied. All of the areas show that the observed nearest neighbour distances higher, $P(d, \leq r)$, larger than the expected nearest neighbour distances for CSR, $P(d, \leq r)$ at different distance ranging from 9 meter to 20.125 meter. This indicates that the spatial distribution of the infected palms in all the study areas is clustered. But the Monte Carlo test with a simulation run 99 times shows that only one out of nine areas shows significance result at 0.05 level of significance. The area is area 2 at MDE8718 with the largest absolute distance, $d_r$, is 0.233. This $d_r$ occurs at a distance of 9 meters. This means the spatial distribution of the infected palms in that area is significantly clustered at the distance of 9 meter. The results of this analysis which is refined nearest neighbour analysis suggested that Ganoderma disease spread probably through basidiospore dispersal at greater distance.

### 3.3. Ripley’s K-function

![Fig. 1: Ripley’s K-function analysis](image)
Fig. 1 shows the results of Ripley’s K function based on the maximum search distance, step size, and the number of permutation for significance envelope of 45 meter, 9 meter, and 9 respectively. All of the graphs (i.e. Figure 2a to Figure 2l) show the observed $\hat{L}(d)$ (dashed line) falls above the expected $L(d)$ (dotted line) which indicate that clustering is present at the distance up to 45 meter. But based on the significance level (i.e. maximum $\hat{L}(d)$ and minimum $L(d)$) set through Monte Carlo simulation, only the distribution of the infected palms in area 1 at SKE0224 shows a significance result as shown in Figure 2d. Again, this analysis which is Ripley’s K function analysis suggested that *Ganoderma* disease spread probably through root to root infection, instead of through basidiospore dispersal at greater distance.

4. Conclusion

The nearest neighbour analysis showed that the spatial distribution of the infected palms in all the replicate areas at SKE0224 was clustered. Refined nearest neighbour analysis and Ripley’s K function also showed that the distribution of the infected palms in all the areas studied was clustered. But statistical test through Monte Carlo simulation showed that the cluster distribution of the infected palms in most of the areas studied was not statistically significant. Nevertheless, the study has proved that the spatial distribution of the infected palms under natural epidemic is not random, disperse or uniform but it is more to cluster pattern. This suggests the spread of the disease could be from tree to tree possibly through root contact as proved by many past studies. The findings of the study are very crucial to the oil palm management in order to control the disease. It provides information on spatial point pattern distribution and mode of *Ganoderma* disease spread at three study sites which were useful in future studies to investigate factors associated with the diseases outbreak and site specific disease management. Cluster pattern will be more manageable as compared to random pattern.

Acknowledgement

We would like to acknowledge Malaysia Palm Oil Board (MPOB) for funding this project and also Sawit Kinabalu Sdn Bhd for allowing us to conduct the study in their selected oil palm estates in Tawau, Sabah. We would also like to thank the entire field assistant especially to Mr. Jumain Sinring, Mr. Mohd Irwan Salleh, and Mr. Sutrisno Sumarno for helping in data collection.

References


