Network approach on characterizing floral diversity in the agroforestry zone of Mount Makiling Forest Reserve, Philippines

M P Martinez¹, A K S Amarga², G A Gestiada¹, M J F Lopez³, A L Nazareno¹ and R M L Roxas-Villanueva¹

¹Institute of Mathematical Sciences and Physics, College of Arts of Sciences, University of the Philippines Los Baños
²Laguna State Polytechnic University Siniloan Campus, Philippines
³Institute of Statistics, College of Arts of Sciences, University of the Philippines Los Baños

rlroxasvillanueva@up.edu.ph

Abstract. Complex networks have been used to characterize real world systems. The network structure may signify important relationships which may not be evident in other methods of analysis. In this study, we characterize the floral diversity in three study sites in the agroforestry zone of Mount Makiling Forest Reserve using network analysis. Plant species found in each study site are considered as nodes (N). Edges (E) are established to connect species with the same alternate role and habit. The dataset includes N = 157 and E = 4279 for Bagong Silang site, N = 145 and E = 3740 for the Karay site, and N= 122 and E = 2429 for the Magnetic Hill site. Network parameters such as degree, path length, clustering coefficient, modularity and number of connected components were calculated. Obtained values were compared to published diversity index. Results show that lower clustering coefficient and higher average path length signify higher diversity. A higher number of disconnected components also indicates diversity.

1. Introduction

Network analysis have been employed in investigating the links and relationship of components of complex systems. There are studies employing network analysis to investigate an ecosystem. Complex network analysis was used to study rainfall [1] and seismic data sets [2]. Furthermore, complex network analysis was employed for mining the topological and functional characteristics of rock masses [3].

In this study, we used a network approach to characterize the diversity in the agroforestry zone of Mount Makiling Forest Reserve, Philippines. Specifically, we characterized the floral diversity in three study sites (Bagong Silang, Karay, and Magnetic Hill) in the agroforestry zone of Mount Makiling Forest Reserve by calculating network parameters such as degree, path length, clustering coefficient, and modularity. We also analyzed and compared the network parameters of the three sites and we compared these values to their corresponding diversity indices.

Biodiversity is used as an indicator of ecosystem health. Optimal species diversity let the available resources in an ecosystem to be used efficiently and diversity lets the ecosystem to adapt to environmental fluctuations [4]. Regulation of functions in an ecosystem is very much dependent on the level of its plant biodiversity [5]. Therefore, it is very important to monitor diversity.
In community ecology, floristic characterization and vegetative profiling were usually done using floral survey techniques. Floral communities can be assessed on the basis of intrinsic properties of floral communities such as aggregation and nestedness [6, 7]. Diversity, abundance, and community structure of vegetation is traditionally measured using statistical tools such as diversity indices (e.g. Shannon- Wiener; Simpson), evenness, and relative abundance [8, 9].

Recent trends in computational botany include network analysis used in the field of phytocchemistry [10], plant pathology [11], plant metabolomics [12], plant phylogenetics [13], and as a tool for measuring vegetation similarity [14]. Herein, we present a network approach on the characterization of vegetation community in agroforest zone of Mount Makiling Forest Reserve.

Agroforest zone influences biodiversity in a specific area [15]. This zone is characterized by presence of agricultural system within a secondary lowland forest. Since this area has more plant species compared to other agricultural systems, it provides diverse habitat to various beneficial faunal groups such as pollinators and seed dispersers [5, 16]. Even though this zone is less diverse in flora as compared to other forest zones such as secondary forests and montane forests, agroforest zones can serve as foraging sites and breeding grounds for a variety of frugivorous and insectivorous taxa. Furthermore, this zone serves as wildlife corridors which in turn aid in the increase of biodiversity via edge effect [17, 18]. Thus, in turn, promoting biodiversity conservation [19-21].

Agroforestry zones have both ecological and economic interactions between different components and a lot of benefits can be obtained from these, but these can be attained from specific species combinations, that is why consistent monitoring of the species in this type of ecosystem is necessary especially when increased biodiversity is to be maintained.

2. Methodology

2.1. Data collection

The data were obtained from the published work of Gruezo [22] on floral diversity profile of Mount Makiling Forest Reserve. Three sites in the agroforestry zone were considered: Bagong Silang, Karay, and Magnetic Hill. A total of 305 plant species (weed, forest, or naturalized) from the three sites were considered for the study. Alternate role and habit for each species were also identified. Alternate role refers to function of the plant species while habit refers to the overall appearance or form of the plant species. The alternate roles considered were: 3 cover crops, 75 ecological, 1 edible fruit, 9 fiber sources, 14 forages, 3 fruit crops, 6 fruit tree, 1 herbal medicine, 118 landscapes, 8 medicinal, 44 ornamental, 5 poisonous, 2 spices, 4 vegetables, 4 weeds, 1 dye, 1 soil erosion control, and 1 parasite. The habits considered were: 4 epiphytes, 25 ferns, 85 herbs, 19 medium-sized trees, 55 large trees, 6 lianas, 1 parasite, 41 shrubs, 30 small trees, 4 tree-like, and 36 vines.

2.2. Network analysis

2.2.1 Construction of the network. A network or graph, G, can be represented as G = (N, E), where N is the number of nodes or entities in a network and E is the number of edges or the links or interactions between the nodes. In this study, a network representing the floral diversity of each site was constructed. The nodes represent the plant species and the edges represent the links between species with the same alternate role and habit.

2.2.2. Calculation of network parameters. The calculated network parameters are number of nodes (N) and edges (E), degree, clustering coefficient, path length, modularity, number of disconnected components, and number of modularity classes. For each node, the presented values are individual degree and betweenness centrality. These values were analyzed to determine if these can represent the diversity index of each site.
• **Degree (D):** For a node, it refers to the number of edges connected to it. For a network, it refers to the average of the degrees of all its nodes.

• **Clustering coefficient (C):** The clustering coefficient of a network is the tendency of the nodes to cluster together which is also the average of the clustering coefficients of all its nodes. It can be calculated using: 
  \[ C = \frac{1}{N} \sum_{i \in V} \frac{2l_i}{k_i(k_i-1)} \]  
  where \( V \) is the set of all nodes in a network, \( k_i \) is the degree of node \( i \), and \( l_i \) is the number of edges between the neighbors of node \( i \).

• **Path length (L):** This refers to the shortest path between two nodes. It can be calculated using: 
  \[ L = \frac{2}{N(N+1)} \sum_{i \neq j} d_{ij} \]  
  where \( N \) is the total number of nodes in a network and \( d_{ij} \) is the smallest number of links between nodes \( i \) and \( j \).

• **Modularity (M):** This value quantifies the quality of division or community formation in a network [23]; it determines how many links are inside a community compared to the number of links between different communities. It can be calculated using: 
  \[ M = \frac{1}{2m} \sum_{i,j} A_{ij} - \frac{k_i k_j}{2m} \delta(c_i, c_j) \]  
  where \( A_{ij} \) is the number of edges between \( i \) and \( j \), \( c_i \) is the community where node \( i \) is assigned, \( \delta(c_i, c_j) \) is 1 if \( c_i = c_j \) and 0 otherwise, and \( m = \frac{1}{2} \sum_{i,j} A_{ij} \).

2.2.3. **Comparison of network parameters to diversity indices.** The trends of the degree and average path length with changing diversity indices were investigated. The average clustering coefficients of the sites were compared to each other. Test of normality was conducted for all the sites’ clustering coefficient. If the assumption for normality was violated, Analysis of Variance cannot be used, and, in this case, the Kruskal-Wallis Test can be used to determine if there is a significant difference between the clustering coefficients for the three sites. If the calculated value of the Kruskal-Wallis test is less than the critical Chi-square value, then the null hypothesis (\( H_0: \) There is no significant difference between the clustering coefficient of the sites.) is rejected.

2.2.4. **Characterization of individual nodes.** The species with the highest betweenness centrality and degree in each network were determined. The species with the highest betweenness centrality occurred in the greatest number of shortest paths between pair of nodes in the network. The species with the highest degree has the greatest number of edges connected to it. The properties of these identified species were determined and its possible implication to the diversity in each site was investigated.

3. **Results and discussion**

3.1. **Network structure of the three sites**

The network of each agroforestry site was constructed as shown in Figure 1. The circles represent the nodes (plant species) and connecting lines were drawn for each pair of nodes with similar alternate role and habit. Each colour represents one modularity class or group of species that are well-connected.

3.2. **Network parameters and their ecological/diversity implication**

The calculated network parameters for each of the three sites are summarized in Table 1.

| Site            | N  | E  | D    | M   | C   | L   | # of disconnected components |
|-----------------|----|----|------|-----|-----|-----|-------------------------------|
| Bagong Silang   | 156| 4057| 52.01| 0.43| 0.82| 1.80 | 1                             |
| Sitio Karay     | 146| 3637| 49.42| 0.26| 0.76| 1.82 | 1                             |
| Magnetic Hill   | 122| 2397| 39.29| 0.39| 0.85| 1.83 | 2                             |
The published diversity indices of the three sites are as follows: Bagong Silang – 3.81, Sitio Karay – 3.94, and Magnetic Hill – 4.14 [31]. These values were compared to the network parameters and the possibility of using the network parameters to quantify diversity were determined.

**Figure 1.** Network structure of floral species in a) Bagong Silang Site and vicinity, Los Baños (N = 156, E = 4057), b) Sitio Karay Site, Pansol, Calamba (N = 146, E = 3637) and c) Magnetic Hill Site and vicinity, Timugan, Los Baños, Laguna (N = 122, E = 2397). The nodes represent the plant species and the edges were drawn for each pair of nodes with similar alternate role and habit.

3.2.1. **Nodes and edges.** Bagong Silang site has the greatest number of plant species followed by Sitio Karay and lastly, Magnetic Hill. Bagong Silang also showed the greatest number of links between species, with 4057 edges. These values do not necessarily explain the diversity in each site. These represent the number of plant species and the number of times there are two plant species with common alternate role and habit.

3.2.2. **Degree.** Bagong Silang has the greatest average degree while Magnetic Hill has the lowest. This means that on the average, each node of Bagong Silang site shares the same alternate role and habit with around 52 other species in this site. Large average degree means that there is more visible links or connectivity between species in the site. Figure 2 shows the trend of the average degree for each site against the published diversity index of that site.

It can be observed that for high values of diversity index, the calculated average degree was low. That is, Bagong Silang, which has the highest degree, has the lowest diversity index.
Figure 2. Relationship between average degree and diversity index. For high values of diversity index, the calculated average degree is low.

3.2.3. Clustering coefficient. Magnetic Hill has the largest clustering coefficient. A large clustering coefficient means that a network has high concentration or in other words, there is a higher tendency for the nodes to form a cluster with one another.

Figure 3. Distribution of the clustering coefficient of each site (Legend: bs – Bagong Silang, kar – Sitio Karay, and tim – Magnetic Hill). The line inside each box plot represents the median of the clustering coefficient of each site. A large clustering coefficient means that there is a higher tendency for the nodes to form a cluster with one another.

Furthermore, the distributions of the clustering coefficient for the sites were investigated (Figure 3). The box plot for Magnetic Hill site showed several outliers which may affect the results. The line inside each box plot represents the median of the clustering coefficient of each site. It was observed from the figure that the medians were numerically different. Whether they are significantly or not significantly different from one another was determined.

Table 2. Test of normality of the three sites.

| Site       | Normality |
|------------|-----------|
| Bagong Silang | Not normal |
| Sitio Karay   | Not normal |
| Magnetic Hill | Not normal |
Test of normality showed that the clustering coefficients for each site were not normally distributed (Table 2). Since the assumption of normality was violated, Kruskal-Wallis Test was used. The null hypothesis states that there is no significant difference among the median clustering coefficients of the three sites. The results of the Kruskal-Wallis Test showed that the $p$-value is $<0.0001$, so the null hypothesis was rejected. That is, at least one of the sites has a different median clustering coefficient. The three sites were then compared, and the results are presented in Table 3.

Table 3. Comparison of the three sites.

| Site                        | p-value     | Remarks   |
|------------------------------|-------------|-----------|
| Bagong Silang vs Sitio Karay | $<0.0001$   | significant |
| Bagong Silang vs Magnetic Hill | 0.8254      | not significant |
| Sitio Karay vs Magnetic Hill | $<0.0001$   | significant |

Table 3 shows that there is a significant difference between the clustering coefficient of Bagong Silang and Sitio Karay, and Sitio Karay and Magnetic Hill. However, with a large $p$-value, it was determined that there is no significant difference between Bagong Silang and Magnetic Hill. This may be attributed to the observation that Magnetic Hill site has two disconnected components whereas the two other sites only have one component. Also, the presence of outliers in Magnetic Hill site may have altered the result for this site.

The observations for the degree and clustering coefficient in relation to the diversity indices were supported by several previous studies. A node with low clustering coefficient has a greater tendency to act as a bridge node or to be a spreader in a network [32]. Therefore, high clustering coefficient, like what was observed in Bagong Silang, can have a less tendency to act as a spreader which may cause a lower diversity index.

3.2.4. Path length. As shown in Figure 5, there is a positive relationship between path length and diversity index, that is, as the diversity index increases, the path length also increases. Average path length is the average number of steps or links between each pair of plant species. Short path length means that there is a lower number of paths or links to be taken for the information to be spread out from one species to another. Also, smaller value of path length indicates a more compact network. That is why, Bagong Silang, which has the smallest diversity index among the three, has the smallest average path length.
Figure 5. Relationship between path length and diversity index. As the diversity index increases, the path length also increases.

3.2.5. Modularity. The modularity of the network can be used in community detection in the network. That is, the modularity obtained in this study shows the presence of communities of plant species in each site. A community of plant species may show a unique set of characteristics among these plant species and set of interactions.

It was found out that modularity was highest for Bagong Silang, then Magnetic Hill, and lastly, Sitio Karay. High modularity indicates higher number of communities or higher tendency for a network to form communities.

In relation to the diversity indices, there was no observed direct relationship on the trend of these two variables. Although as observed in Bagong Silang, high modularity translates to low diversity index.

3.3. Individual nodes and their ecological significance

Table 4. Summary of the species with the highest betweenness centrality and degree per site.

| Site          | Highest Degree | Highest Betweenness Centrality |
|---------------|----------------|--------------------------------|
| Bagong Silang | *Ficus ulmifolia* | *Ficus ulmifolia*              |
| Sitio Karay   | *Amomum elegans, Anaxagorea luzonensis* | *Illigera luzonensis* |
| Magnetic Hill | *Canna indica, Ficus ulmifolia* | *Canna indica, Ficus ulmifolia* |

The goal of betweenness centrality is to quantify the importance of a node or its role as a bridge between other groups of nodes. High betweenness centrality indicates that the node is more critical in the network which may tell that this node can control the functionality or connectivity of the network. Table 4 shows the species with highest degree and betweenness centrality from the three sites. *Ficus ulmifolia, Illigera luzonensis*, and *Canna indica* seem to be the most important species in Bagong Silang, Sitio Karay, and Magnetic Hill, respectively.

4. Summary and recommendation

In this study, we used network analysis to quantify the floral diversity of three sites in the agroforestry zone of Mount Makiling Forest Reserve, namely, Bagong Silang, Sitio Karay, and Magnetic Hill. Calculated network parameters for the three sites were compared to their corresponding published diversity indices. It was found out that low number of disconnected components characterizes low diversity index. For networks with the same number of disconnected components, high degree, high clustering coefficient, low path length, and high modularity corresponds to low diversity index. For further study, other network parameters can be investigated. Network analysis can be used to study the diversity of more ecological sites to somehow generalize the trend of the network parameters against the diversity indices.
References

[1] Agarwal A, Marwan N, Maheswaran R, Merz B and Kurths J 2018 Quantifying the roles of single stations within homogeneous regions using complex network analysis J. Hydrol. 563 802–10

[2] Pastén D, Torres F, Toledo B A, Muñoz V, Rogan J and Valdivia J A 2018 Non-universal critical exponents in earthquake complex networks Physica A 491 445–52

[3] Liu G, He J, Long W, Li W, Gao P, Lu J, Li R, Li L and Tang M 2018 Topological and dynamic complexity of rock masses based on GIS and complex networks Physica A 512 1240–8

[4] Bukvareva E 2017 The optimal biodiversity – a new dimension of landscape assessment Ecol. Indic. 94 6–11

[5] Altieri M A 1999 The ecological role of biodiversity in agroecosystems Agric. Ecosyst. Environ. 74 19–31

[6] Rezende E L, Jordano P, and Bascompte J 2007 Effect of phenotypic complementarity and phylogeny on the nested structure of mutualistic networks Oikos 116 1919–29

[7] Wassmuth B E, Stoll P, Tscharnkte T and Thies C 2009 Spatial aggregation facilitates coexistence and diversity of wild plant species in field margins Perspective in Plant Ecology, Evolution and Systematics 11 127–35

[8] Wilson J B, Gitay H, Steel J B and King W M 1998 Relative abundance distributions in plant communities: effects of species richness and of spatial scale J. Veg. Sci. 9 213–20

[9] Rad J E, Manthey M and Mataji A 2009 Comparison of plant species diversity with different plant communities in deciduous forests Int. J. Environ. Sci. Technol. 6 389–94

[10] Aristyanyi S, Widyarti S and Sumitro S B 2018 Network analysis of indigenous Indonesia medicinal plants for treating tuberculosis Pharmacognosy Journal 10 1159–64

[11] Garrett K A, Alcalá-Briseno R I, Andersen K F, Buddenhagen C E, Choudhury R A, Fulton J C, Nopsa J F H, Poudel R and Xing Y 2018 Network analysis: a systems framework to address grand challenges in plant pathology Annu. Rev. Phytopathol. 56 559–80

[12] Toubiana D, Fernie A R, Nokoloski Z and Fait A 2013 Network analysis: tackling complex data to study plant metabolism Trends Biotechnol. 31 29–36

[13] Chen S, Kim D P, Chase M W and Kim J H 2013 Networks in a large-scale phylogenetic analysis: reconstructing evolutionary history of asparagales (liliaceae) based on four plastid genes PLoS One 8 e59472

[14] Vivaldo G, Masi E, Pandolfi C, Mancuso S and Caldarelli G 2016 Networks of plants: how to measure similarity in vegetable species Sci. Rep. 6 27077

[15] Scales B R and Marsden S J 2008 Biodiversity in small-scale tropical agroforests: a review of species richness and abundance shifts and factors influencing them Environ. Conserv. 35 160.

[16] Varah A, Jones H, Smith J and Potts S G 2013 Enhanced biodiversity and pollination in UK agroforestry systems J. Sci. Food Agric. 93 2073–5

[17] Harvey C A and González-Villalobos J A 2007 Agroforest systems conserve species-rich but modified assemblages of tropical birds and bats Biodiversity Conservation 16 2257–92

[18] Fratini R, Riccioli F, Argenti G and Ponzetta M P 2016 The sustainability of wildlife in agroforestry land Agric. Agric. Sci. Procedia 8 148–57

[19] Perfecto I and Vandermeer J 1997 The agroecosystem: a need for the conservation biologist’s lens Conserv. Biol. 11 591–2

[20] Vandermeer J and Perfecto I 2007 The agricultural matrix and future paradigm for conservation Conserv. Biol. 21 274–7

[21] Asare R, Afari-Sefa V, Osei-Owusu Y and Pabi O 2014 Cocoa agroforestry for increasing forest connectivity in a fragmented landscape in Ghana Agrofor. Syst. 88 1143–56

[22] Gruzow W S M 1997 Floral diversity profile of Mt. Makiling Forest Reserve, Luzon, Philippines. In The Conditions of Biodiversity Maintenance in Asia: The Policy Linkages Between Environmental Conservation and Sustainable Development, East-West Center.

[23] Newman M E J 2004 Detecting community structure in networks Eur. Phys. J. B 38 321–30.