DRIVERS OF MODULARITY IN PLANT-POLLINATOR NETWORKS OF MONTANE MEADOWS

LYDIA MILLER
GOSHEN COLLEGE | OREGON STATE UNIVERSITY
ECO-INFORMATICS SUMMER INSTITUTE 2017

ABSTRACT

Mutualistic plant-pollinator networks are complex and ecologically valuable assemblages, and deeper understandings of their structure could be critical to conservation and forest management. Modularity analysis, the study of subgrouping in a network, could bring insight to network structure at a finer level, and help elucidate the flow of resources, resiliency, and species roles in a community.

I analyzed a longitudinal dataset of plant-pollinator interactions collected at the H. J. Andrews Experimental Forest in the Western Cascades of Oregon by calculating modularity statistics for several time periods, and comparing the results of weighted and binary approaches to modularity. I aimed to explore the various formulations of modularity for these networks, with species defined at the family level. I sought also to determine how phylogenetic relatedness of plants might play into module makeup.

Modularity was calculated for four networks, defined by year, using Beckett's DIRTLPawb+ algorithm [1], with plants and pollinators defined at the family level. Modularity scores and module counts were determined for each network, using both binary and weighted approaches, and results were compared to modularity results for a distribution of null networks matched to the observed networks for size, connectance and link density. Intra-module phylogenetic relatedness of plants was calculated for each network using Huestis' phylogenetic distance metric [8]. These results were also compared to null distributions using Z-scores.

Modularity scores were found to be low in all observed networks relative to null models, with a more extreme result for weighted networks. Weighted and binary formulations gave the same number of modules for every network. Intra-module relatedness was not significantly different from null models, except for the binary network in 2011-14, which was significantly lower than expected (p-value = .041).

My findings of low modularity contrast other findings in the literature, where significantly high modularity is the norm. However, other studies have not considered networks at the family level. My result could be due to super generalist nodes common to the Andrews Forest: families with high degree and many links between modules, whose presence decreases modularity scores. Modules do not appear to be defined by closely phylogenetically related plant families,
suggesting that phylogeny might not be a driving factor for module makeup, at least at the family level. Continuation of this analysis at finer phylogenetic resolutions could be valuable in determining if modular structure does exist for these networks and if phylogenetic relatedness is associated with module membership when the network is viewed at the level of genus or species.

INTRODUCTION

Plant-pollinator networks are complex structures consisting of a great diversity of species interacting in non-random ways. The plant-pollinator networks studied in this paper occur in the montane meadows of the H.J. Andrews Experimental Forest (hereafter Andrews forest), located in the Western Cascades of Oregon. These communities are under a variety of pressures from climate change, invasive species, and forest encroachment, making the study of plant-pollinator networks critical at this time for conservation and forest management [12].

Plant pollinator networks are bipartite, meaning that two classes of nodes exist, such that nodes in one class are only connected to nodes of the other class. In this context, the classes are plants and pollinators, but other bipartite networks have been studied in the ecological literature, including host-parasitoid networks [5]. Bipartite networks are either binary or weighted, depending on the presence of weights on the links. A binary network gives information only on whether links between plants and pollinators are present or absent, while a weighted approach takes into account the strength or weakness of a link based on the number of interactions between a pair of species [1,5]. Weighted networks hold more information than binary networks, and as such can be collapsed to their binary forms by removing link weights.

Modularity is a network characteristic quantifying the amount of subgrouping or “clumping” present within a larger network. Module-finding algorithms for binary networks seek to find groupings of nodes where the average number of within-module links is maximized relative to between-module links; that is, nodes are more connected in their groups, and less connected outside of their groups [1,2,3,5,10,11]. Weighted modularity compares sums of link weights rather than link counts, giving a very different picture of the modularity landscape [1,5].

A partition is defined here as a grouping of nodes into some number of modules. Modularity is characteristic of a network partition, and since there are a great number of possible partitions for a given network, there are likewise many possible modularity values. In general, when I refer to the modularity of a network (rather than a partition), I am referring to the maximum modularity obtained by any partition of the network, or an estimate of this.

Much of the literature on modularity of plant-pollinator networks takes a binary approach, disregarding link weights [1,2,10,11,5]. Beckett (2016), proposes a weighted module-finding algorithm for bipartite networks superior to any in the literature so far, and as of yet, no analysis contrasting modularity for weighted and binary networks has been concluded using this algorithm.

Null models as tools for standardizing modularity results are a recent area of focus in the context of modularity and plant pollinator networks. Standardization is valuable since it allows comparison among networks with different sizes, connectance, and link densities, three network characteristics that are associated with modularity. This approach has been used only once before, by Carstenson et al in a 2016 study of the modularity plant-pollinator networks in Brazilian rupestrian grasslands, and their interaction turnover across space [2].
Modules are often thought to be co-evolutionary units, where the plant and pollinator species in a module have evolved together over time, and have developed compatible traits that allow for a mutualistic relationship [2]. If this is true, it is possible that plants within modules would be more likely to be closely related phylogenetically. No study has yet explored relationships between phylogenetic relatedness of plants and modularity.

I analyzed plant-pollinator data collected at the Andrews forest in order to examine the following questions:

1. **How modular are the networks observed in the Andrews forest?** In particular, how do weighted and binary approaches compare, and how do these results compare to null distributions?
2. **Is phylogenetic relatedness of plants a possible driver of module formation?**

I expected that the networks would be more modular than a null distribution, and that binary networks would yield different modularity and module count results from weighted networks. Assuming that that phylogenetic relatedness played a part in module formation, I expected to find high intra-modular relatedness relative to a null model as evidence of this.

**METHODS**

**STUDY SITE**

The meadows sampled in this study are fragmented and located across several ridgelines in the Andrews forest. Archeological evidence suggests that prehistoric people created and maintained these open areas for hunting or foraging purposes, and the meadows have persisted due to both intentional human burns and wildfires [9]. The recent efforts to suppress wildfires in the Cascades have caused the meadows to decrease in area since the mid-19th century [12].

Twelve of these meadows were surveyed, with significant variation among them in meadow size, location, aspect, and soil depth and moisture. All meadows are located at around 1500 feet of elevation.

**FIELD METHODS**

Surveying was conducted in five watches at each meadow during every summer from 2011 to 2016. Each meadow contains ten 3m x 3m survey plots, 15m apart, in two parallel transects. Surveys consisted of two parts:

---

**Anthesis Portion**

An exhaustive survey of blooming flowers within the plot was conducted. Observers counted the number of stalks of a given species present, and calculated an average number of flowers by stalk by selecting up to 10 stalks. The total number of flowers for a species was estimated by multiplying these values.
A “flower” was considered as a unit of floral resource from the perspective of a pollinator, and as such sometimes differed from the strict botanical definition. For instance, umbels or umbelletes (tight clusters of tiny flowers) were counted as single flowers.

15 Minute Interaction Watch

For 15 minutes, each observed interaction between a pollinator and a plant was recorded, where an interaction was defined as contact between a potential pollinator and the surface of a flower. For each interaction, both the plant and pollinator species were recorded, as well as the number of plants visited. When the observer could not identify a pollinator, the pollinator was captured, pinned, and identified by an expert (Dr. Andy Moldenke, professor emeritus at Oregon State University).

Capturing of pollinators was done as infrequently as possible so as not to influence the meadow ecology. Given that our survey plots represented a small fraction of the meadows, and that our watches lasted only 15 minutes, 5 times a year, we don’t believe that our pollinator sampling affected pollinator populations in a significant way.

For a more detailed account of surveying methods, see the study description for the LTER at http://andltter.forestry.oregonstate.edu/data/abstractdetail.aspx?dbcode=SA026

STATISTICAL METHODS

Data Used

Although both plants and pollinators were identified to the species level, in this study I observe networks defined at the family level for both plants and pollinators, since phylogenetic relatedness information was available only on the family level. My network data sets include the following years: 2011-14 combined, 2015, 2016, and an aggregated network made up of interactions from all 6 years (referred to as ALL). These groupings were chosen in order to see changes present from one year to the next, and also to provide sufficient interactions to contrast larger networks (those with more species and interactions) with smaller ones.

Table 1. Descriptive statistics of networks studied in this paper.
The number of possible links is the number of plant families times the number of pollinator families. The number of interactions is the sum of all link weights, and would be equal to the number of realized links in a binary network.

<table>
<thead>
<tr>
<th>year</th>
<th>number of plant families</th>
<th>number of pollinator families</th>
<th>number of possible links</th>
<th>percent of realized links</th>
<th>number of interactions</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011-14</td>
<td>29</td>
<td>73</td>
<td>2,117</td>
<td>23.5%</td>
<td>97,315</td>
</tr>
<tr>
<td>2015</td>
<td>23</td>
<td>47</td>
<td>1,081</td>
<td>16.7%</td>
<td>9,055</td>
</tr>
<tr>
<td>2016</td>
<td>24</td>
<td>49</td>
<td>1,176</td>
<td>17.1%</td>
<td>11,472</td>
</tr>
<tr>
<td>ALL</td>
<td>32</td>
<td>78</td>
<td>2,496</td>
<td>22.0%</td>
<td>117,842</td>
</tr>
</tbody>
</table>
Modularity calculations

I used Beckett's DIRTLPawb+ algorithm for all modularity calculations. I chose this algorithm since it appears to be the best and most consistent method for estimating modularity in weighted networks; that is, it returns equal or higher modularity values compared to other algorithms in the literature [1]. Furthermore, the algorithm has the ability to return a binary network modularity given a binary input. DIRTLPawb+ is a stochastic optimization algorithm, attempting to find a partition that optimizes modularity, and thus may give a different value every time it runs. To get an accurate estimate of network modularity, I took the maximum of 20 independent trials for each network.

Modularity varies depending on the link density (number of links per species, or sum of weights of links per species, for weighted networks), size (total number of nodes), and connectance (proportion of realized links) of a network [1,11]. These characteristics vary among the years studied, and must be controlled for if modularity values are to be compared across networks. To accomplish this, I compared the modularity score of each observed network to a distribution of modularity scores calculated for 20 null model network partitions, where link density, size, and connectance were kept constant (see fig 1). R code for null model generation can be viewed in additional materials. Comparison to a null distribution was used to find the Z-score of the observed modularity value, where the Z-score was interpreted as a normalized modularity score that could be compared across networks. I developed this method independently but later found that it had been used by others, including Carstenson et al. in 2016 [2].

![Example frequency histogram of the distribution of null network modularities.](image.png)

Figure 1. **Example frequency histogram of the distribution of null network modularities.** This example is from all years (2011-16) aggregated. The mean and standard deviation of this distribution can be used as a baseline from which to compare observed modularity values.

Phylogenetic Relatedness Data

Phylogenetic plant distance data was taken from Huestis’ (2017), and is calculated pairwise between two plant families. Huestis uses a shorthand metric to calculate pairwise distances
between terminal taxa on a phylogenetic tree. This metric takes into account the distance travelled in the tree in moving from one taxon to the other. Tree path distance is a sum of the link weights travelled in the shortest possible path, where link weights vary on an ordinal scale from 1 to 9 based on the resolution of the link. Courser-resolution links, i.e. those further from the terminal taxa, are more heavily weighted, while links closer to the terminal taxa have less weight. The phylogenetic distances for all pairs ranged from 1 to 58, so phylogenetic relatedness was defined simply as 59 – (distance). An illustration of Huestis’ method is provided in Additional Materials.

The network index intra-module phylogenetic similarity or \( S_p \), was defined as the average phylogenetic similarity of pairs of plants sharing a module. That is, for each module, phylogenetic similarities were determined for each pair of plants in that module, and then added to a list. All values of all lists were then averaged to find a network mean of \( S_p \). For comparison, a distribution of null models was created for each network. For a given observed network, each null network was exactly the same, but the partition was randomized, since intra-module phylogenetic similarity is a function of a network partition. The null partitions had the same number of modules as the observed network, but the grouping of nodes into modules was randomized (methods and codes for null partition generation are found in Additional Materials). For each observed network, the \( S_p \) value was calculated and compared to the distribution of \( S_p \) values for 50 null partitions, and compared as Z-scores.

\[ P \text{ values of less than 0.05 were considered to be significant in all parts of this study.} \]

**RESULTS**

**HOW MODULAR ARE THE NETWORKS OBSERVED IN THE ANDREWS FOREST?**

Modularity values and module counts were calculated with 20 runs of DIRTLPawb+ for both binary and weighted networks. Modularity fell between .09 and .23 for weighted networks, and between .20 and .32 for binary networks. Module counts were identical for every year between weighted and binary networks, with 4 or 5 modules each.

<table>
<thead>
<tr>
<th>Year</th>
<th>Weighted Modularity</th>
<th>Weighted # of Modules</th>
<th>Binary Modularity</th>
<th>Binary # of Modules</th>
</tr>
</thead>
<tbody>
<tr>
<td>2016</td>
<td>0.2240</td>
<td>5</td>
<td>0.3106854</td>
<td>5</td>
</tr>
<tr>
<td>2015</td>
<td>0.0939</td>
<td>5</td>
<td>0.3151543</td>
<td>5</td>
</tr>
<tr>
<td>2011-14</td>
<td>0.1952</td>
<td>4</td>
<td>0.2086982</td>
<td>4</td>
</tr>
<tr>
<td>ALL</td>
<td>0.1766</td>
<td>4</td>
<td>0.2181035</td>
<td>4</td>
</tr>
</tbody>
</table>

There is significant variation in modularity from year to year, especially in weighted results, with a notably low modularity value of 0.0939 in 2015. Because these are raw data, not compared to null models, values cannot be directly compared from one year to the next, or from weighted to binary.
**HOW DOES OBSERVED MODULARITY DIFFER FROM A DISTRIBUTION OF NULL MODELS?**

<table>
<thead>
<tr>
<th>Year</th>
<th>Weighted observed modularity</th>
<th># Modules</th>
<th>Weighted null average modularity</th>
<th># Modules</th>
<th>Z score</th>
<th>Weighted observed modularity</th>
<th># Modules</th>
<th>Weighted null average modularity</th>
<th># Modules</th>
<th>Z score</th>
</tr>
</thead>
<tbody>
<tr>
<td>2016</td>
<td>0.22</td>
<td>5</td>
<td>0.37</td>
<td>6.20</td>
<td>-15.19</td>
<td>-1.44</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td>0.09</td>
<td>5</td>
<td>0.38</td>
<td>5.95</td>
<td>-46.22</td>
<td>-1.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011-14</td>
<td>0.20</td>
<td>4</td>
<td>0.26</td>
<td>5.55</td>
<td>-15.79</td>
<td>-1.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALL</td>
<td>0.18</td>
<td>4</td>
<td>0.26</td>
<td>5.85</td>
<td>-12.63</td>
<td>-1.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>Binary observed modularity</th>
<th># Modules</th>
<th>Binary null average modularity</th>
<th># Modules</th>
<th>Z score</th>
<th>Binary observed modularity</th>
<th># Modules</th>
<th>Binary null average modularity</th>
<th># Modules</th>
<th>Z score</th>
</tr>
</thead>
<tbody>
<tr>
<td>2016</td>
<td>0.31</td>
<td>5</td>
<td>0.36</td>
<td>5.75</td>
<td>-6.11</td>
<td>-0.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td>0.32</td>
<td>5</td>
<td>0.38</td>
<td>5.80</td>
<td>-5.47</td>
<td>-1.15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011-14</td>
<td>0.21</td>
<td>4</td>
<td>0.25</td>
<td>5.55</td>
<td>-13.83</td>
<td>-2.26</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALL</td>
<td>0.22</td>
<td>4</td>
<td>0.25</td>
<td>5.85</td>
<td>-7.17</td>
<td>-2.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 3. **Standardized modularity and module count scores for binary and weighted networks.** Observed modularity and module count values were compared to null models (see Fig. 1), and the resulting Z-scores are interpreted as standardized values. All values are negative, meaning that for all networks, both modularity and module counts are less than what would be expected from random mixing.

Standardized modularity and module count scores are shown in Figure 3. All values are low, compared to null networks, modularity scores extremely so. Binary networks are more modular than weighted networks, and tend to have more modules.

**IS PHYLOGENETIC RELATEDNESS OF PLANTS A POSSIBLE DRIVER OF MODULE FORMATION?**

<table>
<thead>
<tr>
<th>Year</th>
<th>observed</th>
<th>null average</th>
<th>z score</th>
</tr>
</thead>
<tbody>
<tr>
<td>2016</td>
<td>33.21622</td>
<td>34.74377</td>
<td>-0.8053037</td>
</tr>
<tr>
<td>2015</td>
<td>34.34</td>
<td>33.72441</td>
<td>0.3221022</td>
</tr>
<tr>
<td>2011-14</td>
<td>35.35644</td>
<td>35.31099</td>
<td>0.03486598</td>
</tr>
<tr>
<td>ALL</td>
<td>34.11024</td>
<td>33.02974</td>
<td>0.9304292</td>
</tr>
</tbody>
</table>

Table 2. **Intra-module phylogenetic relatedness (S_p) scores for observed networks and corresponding null partitions.** Weighted values are shown in (a), and binary values in (b)
Figure 4 shows $S_p$ values, a network metric for phylogenetic relatedness of plants in modules, for all four years, and for binary and weighted partitions. All binary standardized $S_p$ values were less than 0, indicating that there is a tendency in these networks for plants to be in modules with other plants that are less phylogenetically related. The weighted results vary from year to year, with the aggregated network of all six years of data having the highest standardized score.

These results vary slightly from 0, but overall, they indicate modules do not have more closely related species than any other subset. The binary graph in 2015 is a possible exception. A Z-score of -1.73 indicates that this value is significantly lower than expected ($p = .041$), and that modular structure may not be independent of phylogenetic relatedness of plants. This relationship is opposite of the relationship I expected to see, however. The low values for binary networks indicate that closely related plants are less likely to share a module than would be expected.

### DISCUSSION

#### MODULARITY

Weighted and binary network modularity scores were calculated for observed plant-pollinator networks in the Andrews forest with the dual intentions of learning more about the structure of these networks, and comparing inferences from the binary and weighted approaches.
Modularity was found to be extremely low compared to null model distributions for both weighted and binary networks, with a more extreme result (lower modularity) for weighted networks. This contrasts what has been previously found in the literature; plant-pollinator networks are generally found to be significantly modular, [1,2, 3,11,14], and binary networks have been found to be less modular than weighted networks [2]. It is possible that this difference could be explained by my choice to evaluate plants and pollinators at the family level. All other analyses cited considered nodes as plant or pollinator species. Grouping plants by family rather than species would likely decrease modularity values if modules in species-defined networks have high numbers of species with families in common sharing a module.

My results could be explained by a variety of ecological and statistical interpretations. Connections across modules decrease modularity by definition, thus generalist plants and pollinators that form links with a wide diversity of species could play a part in the low observed modularity scores. Table 3 shows the top generalist plants and pollinators, defined as the top three families with highest degree. These nodes have extremely high degrees, and many (perhaps most) of their links are with nodes outside their module. This effect helps to explain the low modularity scores for both weighted and binary networks.

These generalist nodes have high interaction counts, all falling above the 90th percentile. In fact, for both plants and pollinators, the family with the highest degree also has the highest number of interactions. This means that in the weighted approach, inter-module links for these nodes have the potential to be very heavily weighted, making their modularity-decreasing effects even greater.

Table 3. Top generalist pollinators and plants and their connectance. The top 3 plant and pollinator families for degree are shown below. There are 32 plant families total for all years, so Apidae with a degree of 28 is connected to all but 4 plant families. The top generalist plant family, Asteraceae is connected to 67 out of 78 possible pollinator families. The number of interactions for each species is shown along with the percentile for interactions. Families with high degree also have very high percentiles for number of interactions.

<table>
<thead>
<tr>
<th>family name</th>
<th>degree</th>
<th>number of interactions</th>
<th>interactions percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pollinators with highest degree</td>
<td>Apidae</td>
<td>28</td>
<td>72,527</td>
</tr>
<tr>
<td></td>
<td>Syrphidae</td>
<td>25</td>
<td>9,617</td>
</tr>
<tr>
<td></td>
<td>Halictidae</td>
<td>23</td>
<td>2,646</td>
</tr>
<tr>
<td>Plants with highest degree</td>
<td>Asteraceae</td>
<td>67</td>
<td>31,501</td>
</tr>
<tr>
<td></td>
<td>Apiaceae</td>
<td>51</td>
<td>9,954</td>
</tr>
<tr>
<td></td>
<td>Polemoniaceae</td>
<td>48</td>
<td>34,042</td>
</tr>
</tbody>
</table>

The negative effect of generalists on modularity could be tested by removing these generalists and recalculating modularity on the resulting network. A significantly higher modularity score for the new network would indicate modularity-undermining effects of these pollinators.

Similarly, it is important to note that the most common generalist pollinators are also some of the physically largest, and most easily identified pollinators. Pollinators that are missed by the observer are not recorded, and neither are pollinators that the observer sees but cannot identify. Sampling bias against small or rare species might result in over-sampling of common
generalists. Misidentification of an unknown species as a common generalist could have the same effect. Oversampling of generalists could result in artificially low modularity results.

These low modularity values indicate that modularity is not the most dominant, defining structure for these communities, at least at the whole network level. Exploration of modularity at the level of meadow or complex could yield different results. Observed plant-pollinator networks in the literature have aspects of both modular structure and nestedness – a network property where the plants visited by a pollinator with small degree is nested subset of the plants visited by a pollinator with larger degree, and vice versa with plant and pollinator roles reversed [6, 11, 12]. In highly connected networks, nestedness tends to be indirectly associated with modularity [6], and further research into the nestedness of the networks we observed could give a better understanding of the interplay of these two network tendencies.

It is important to keep in mind, however, that even if a network is estimated to have a low modularity score, the modularity-maximizing partition still contains important information, and there is value in looking for patterns in module membership for these networks. It is likely that species in a module tend to share certain characteristics, even if the modules are defined loosely, as indicated by a low modularity score, and uncovering these tendencies could elucidate driving factors of the modularity (or lack of modularity) of a network.

**INTRA-MODULE PHYLOGENETIC RELATEDNESS**

Intra-module phylogenetic relatedness tended to be lower in observed networks than in a null partition, refuting my hypothesis. I expected to see higher relatedness between species in modules, believing phylogeny to be a driving factor in the way modules formed. I expected, for instance, that if two plants were more phylogenetically related, they would be likely to be attractive to the same pollinators, creating triangles of links that might be the underlying structure of a module. As Huestis found, however, phylogenetic relatedness does not correlate strongly with shared pollinators, and as shown by my research, does not seem to have a positive relationship with module co-membership either.

The results from this section are intriguing, since while the $S_p$ values differed from my expectations, they also differed from null expectations for binary networks, indicating a possible relationship. This result suggests that something might be driving phylogenetically similar plants to attract different pollinators. It has been proposed that coexisting plants might evolve varied traits in order to decrease competition for pollinators, which may be an avenue by which phylogenetically similar plants come to have very distinct morphologies [13]. This non-alignment between phylogeny and traits could explain the low $S_p$ values for observed networks, assuming that traits are a strong driver of likelihood of interaction.

**VALIDITY CONSIDERATIONS**

The reference distributions used in this study consisted of only 20 null networks for modularity comparisons, and only 100 for intra-module relatedness comparisons, and these low numbers may be a threat to the validity of the results. A small distribution will give less accurate estimations of the mean and standard distribution of the true null distribution, and as such could give $z$ scores that are insufficiently accurate. A greater number of null networks could increase our confidence in the resulting $Z$-scores (normalized values).
Another source of possible inaccuracy is the data collection process, which only gives us a limited estimate of plant-pollinator interactions. Interaction watches occurred at different times of day, under different levels of sun and wind, and at slightly different points in the season from year to year (although all were in the months from June to August). The researchers were students with relatively little experience in this type of data collection, and there was high turnover of observers from year to year. The 3m x 3m plots are too large for one person to see every interaction that occurs during the busiest interaction periods, and pollinators that are fast or small could be easily missed. Misidentification of both plants and pollinators is also a concern, especially given the high number of species, and the acute similarities in appearance between many of them.

One type of sampling bias includes the under-sampling of rare species, which has been shown to lead to the characterization of these species as more specialist than they really are [4].

Using a weighted approach is one way to combat data inaccuracy. When considering link weights, we give more emphasis to very clear, well-documented relationships than to lightweight links that are possibly the result of a mistake or error. As noted above, this weighting also has disadvantages, and for some analyses, a binary approach is desired. In these cases, a third approach could involve truncating the lowest-weighted links from a weighted network, and then projecting this truncated network to its binary form.

Aggregating all study years is another protection against these inaccuracy concerns. With larger data sets, we can be more confident in our results, trusting that over time, the trends that occur will better and better reflect the truth. Of course, this is only the case if mistakes are random, and many are not, including the consistent bias toward larger and more common pollinators. For this type of skew, a large dataset is no advantage.

The aggregated data set is advantageous, however, to correct for variation in climate from year to year. The wetness of the spring and summer, the date of the last snowpack, the humidity and sun intensity of the summer, and other climatic or weather variations from year to year have a noticeable effect on plant and pollinator activity. When data are combined over several years, the effects of this type of variation will be lessened.

CONCLUSIONS

I sought to explore the modularity of the plant-pollinator networks in the Andrews by comparing weighted and binary approaches to modularity, and by comparing modularity results to a distribution of null models. I found, contrary to my hypothesis, that modularity and module counts were lower in observed networks than in the null models for both binary and weighted networks. Furthermore, the modularity values for the weighted networks were significantly lower than for binary networks in all years. Interestingly, binary and weighted modularity approaches both gave the same number of modules for a network, which raises the question of whether the members of each module differ, or whether the two approaches give the same or very similar partitions.

I sought also to determine whether phylogenetic relatedness of plants might be a driving factor in the modular structure of these networks, and I did this by comparing intra-module
phylogenetic relatedness to a distribution of null partitions for each network. I expected intra-module relatedness to be higher than the null distribution, but I found that this relatedness metric was not significantly different from the null distribution for weighted networks, and was lower than the null average for all binary networks. This result suggests that intra-module phylogenetic relatedness of plants is not a defining module feature, and that there are likely other stronger driving factors for module structure. An exploration of intra-module phylogenetic relatedness of pollinators is a logical next step for future study.

In future studies, I recommend further study into module member roles, that is, the way that individual nodes function as part of their module and part of the network as a whole. Several studies have produced work in this area [11,14], and consideration of module member roles might be a natural continuation of the discussion in this paper of generalist and specialist roles in modularity.

Future work might also examine the nodes that make up the modules in more detail, with the intention of finding patterns and possible explanations for module membership. One relevant approach is the study of pollination syndromes: groupings of plants by shared traits, with the goal of predicting shared pollinators based on syndrome groupings [2, 3, 11]. Others have suggested overlapping phenology as a driver of modularity [7].

ACKNOWLEDGMENTS

Heartfelt thanks to my mentors, Dr. Julia Jones, Dr. Rebecca Hutchinson, and Stephanie Bianco. Julia, your ecological wisdom helped me see the world in a new way. Rebecca, your feedback and suggestions were invaluable to so many parts of my analysis. Steph, none of this could have happened without your support every step of the way – I’m so grateful for your mentorship and your friendship. To all my mentors: your guidance was huge in allowing me to achieve what I did this summer. It was an honor to work with such wise and kind people. Thank you.

Warm thanks to Dr. Andrew Moldenke, whose depth of knowledge of pollinators and plants is astounding. Your presence greatly improved the accuracy and depth of our data set, and we thoroughly enjoyed your humor and knowledge. Thanks also to Mark Schultze and H.J. Andrews administration for the opportunity to conduct research in the midst of not just beautiful ecology but wonderful people. Deep thanks to Oregon State University and the National Science foundation for making this opportunity possible. Finally, I’m so grateful to my teammates, Elaina Thomas, Jane Huestis, Andrew Guide, and Joshua Griffin, as well as the EISI Stream Team, Malia Gonzales, Peter Duin, Sashka Warner, Adam Zhang, and Brent Davis, for making this summer unforgettable. Thank you for your intelligent insight, hard work, humor, and friendship.

REFERENCES

1. Beckett SJ. 2016 Improved community detection in weighted bipartite networks. R. Soc. open sci. 3: 140536

3. Danielli-Silva, A et al., 2012 Do Pollination Syndromes Cause Modularity and Predict Interactions in a Pollination Network in Tropical High-altitude Grasslands? OIKOS 121. 35-43


7. Gonzáles et al, 2012 Species interactions in an Andean bird-flowering plant network: phenology is more important than abundance or morphology. PeerJ 4:e2789


