Appendix G: Rarefaction Analysis

Measurement of Biodiversity: A: **Rarefaction Analysis**

Rarefaction analysis is a technique to assess species richness as a function of sampling effort. Rarefaction allows the evaluation of species richness for a given number of individual sampling events, based on the construction of so-called rarefaction curves. This curve is a plot of the number of species found as a function of the sampling intensity. These curves generally grow rapidly at first, as the most common species are found, but they plateau as only the rarest species remain to be sampled.

This technique of rarefaction was developed in 1968 by Howard Sanders in a biodiversity assay of marine benthic ecosystems. Following Sanders work, this technique has, through peer review, become more refined and is now generally accepted as a valuable tool to understand species diversity. Central to the modern concept of rarefaction is it estimates biodiversity when information is imperfect, for as more data is collected, and the number of species found is plotted against sampling effort, a curve is produced which reaches an asymptote. Further, statistical tools can be used to derive the maximum number of species present, and conversely, estimate how much sampling effort is necessary to quantify diversity.

Although, a caveat is that rarefaction works best when no taxon is extremely rare or very common and when the beta diversity is very high. It assumes that the number of occurrences of a species reflects the sampling intensity, but if one taxon is especially common or rare, the number of occurrences will be related to the extremity of the number of individuals of that species, not to the intensity of sampling. However, this technique does not recognize species abundance, only species richness. A true measure of diversity accounts for both the number of species present and the relative abundance of each.

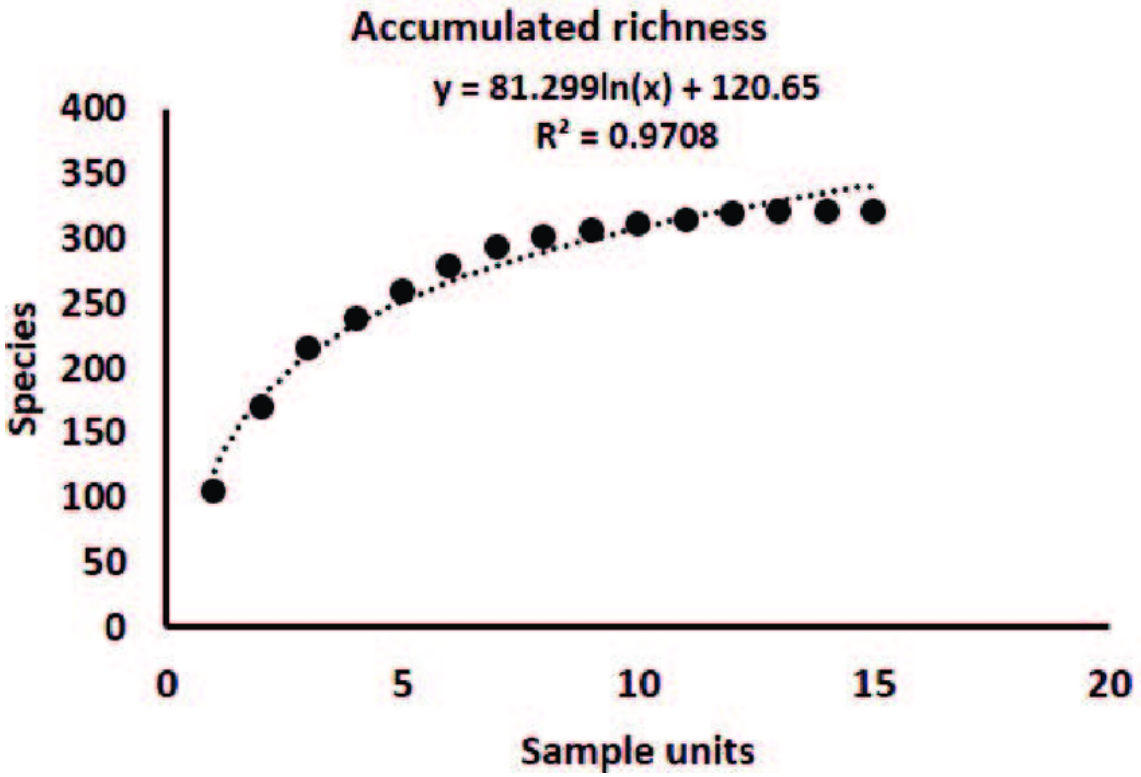

Figure 1. Typical species richness curve showing rarefaction. (From Zamora and Morales (2019) https://www.researchgate.net/publication/330117427

Although the rarefaction analysis was developed when ecologists studied relatively large plants and animals, with the advent of microbial ecology in the early 2000's, such analysis became an imperative, thus the technology advanced yet further. Australia has a total of 18,448 species of vascular plants (https://www.dcceew.gov.au/science-research/abrs/publications/other/numbers-living-species/discussion-plants) whereas a single gram of soil contains between 30,000 – 50,000 taxa of bacteria and fungi, and in the order of 10 billion individuals. The standard way to identify microbes is by doing DNA sequencing of the 16S ribosome. The question became how long to sequence? The answer, as with the biodiversity studies of bigger species, lies with rarefaction analysis. Above a sequencing depth of 1500 base pairs, you get very little additional information for the additional effort and expense

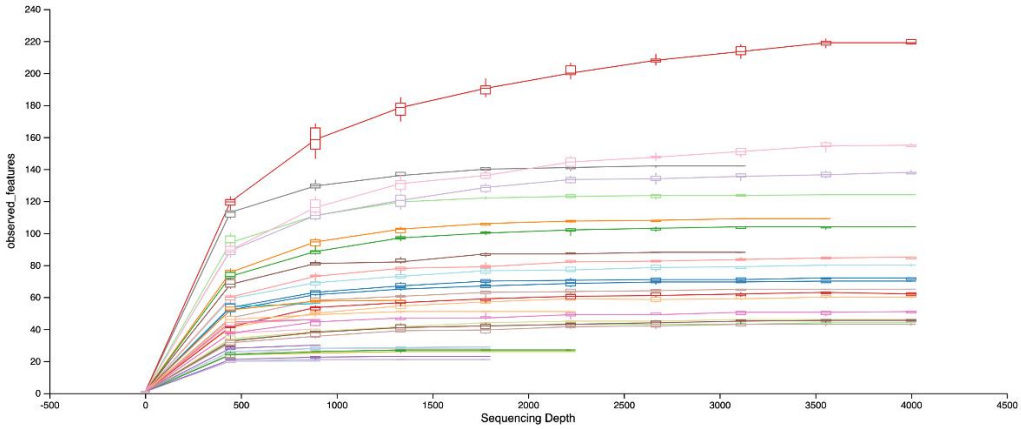

Figure 2: Rarefaction curves expressed as a function of DNA sequencing depth, from https://bioinformatics.ccr.cancer.gov/docs/qiime2/Lesson5/

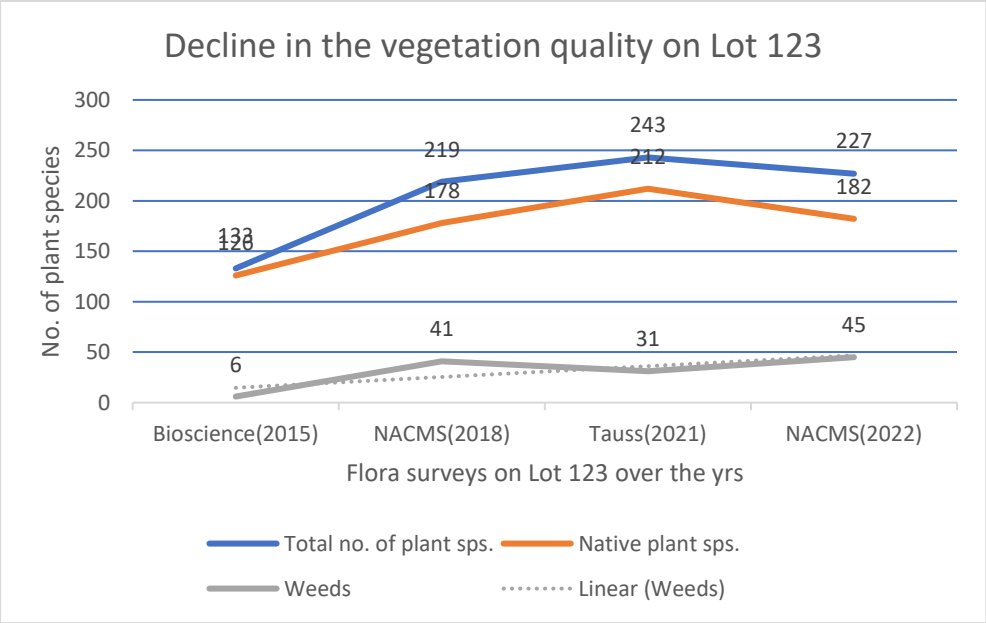

Figure 3: Rarefaction curves showing the total number of plant species and weeds present on the Lot 123 has reached an asymptote as confirmed by four separate flora surveys over the period of 2015-2022. The conclusion is that further analysis is not warranted as it would be subject to the law of diminishing returns for any additional effort.

Additionally, the figure 3 shows that the total number of native plant species is currently 212 (132 in 2015 according to Bioscience, 2015), whereas the total number of weeds present are 45(NACMS, 2022) as compared to 6 in 2015(Bioscience, 2015). This indicates a decline in the overall quality of vegetation on the land with an approximately 500% increase in the number of weeds over the years. Also, four flora studies were referred to and it was concluded that enough knowledge was obtained to determine that there was degradation of vegetation on the Lot 123. Further, a considerable effort in terms of weeding and investment in weed management treatments would be required from the owner for its ongoing maintenance.

Measurement of Biodiversity: B: **Dendrogram analysis**

A dendrogram is a graphical representation of hierarchical similarly between two sets of data. It is very commonly used in flora analysis without explanation. Dendrograms are most commonly used to assign data from a quadrat and, thereby, assign it into a floristic community type. The statistical method most commonly used is Bray Curtis dissimilarity analysis, wherein a metric is used that is based on the species present, and the number of individuals of that species counted at a site. If the metric scores is 1, the sites are unrelated to each other. If the score is zero, the sites are essentially of identical composition. Bray Curtis dissimilarity requires that the same method of sampling and analysis is used at each compared site. An alternative method frequently used is called Jaccard analysis which, like Sorensen analysis compares only the presence or absences of species, thus, it partly avoids the requirement of identical sampling and analysis methods. As opposed to Bray Curtis, if two datasets share all members in common the score is 1, whereas, if there are no common members, the score is zero. However, it is mathematically valid to take the Jaccard score subtracted from 1 to obtain the Jaccard dissimilarity. The floristic community types (Vegetation Units) of the Swan Coastal Plain were defined through a major study undertaken and published in 1994, often referred to simply as Gibson et al. This study reported on 509 sites within the Coastal Plain, selecting the sites with the least disturbed vegetation. Each site was marked as a 10 m x 10 m quadrat, and in the 509 sites, with most sites (>95%) visited at least twice, a total of 1313 native species and 173 weed species were found (11.6% of the total species were weeds). Though it was noted that 525 taxa known to be in the Perth area at the time were not encountered in this study, and of these, 183 were native, and 342 (65%) were non-native weed species.

In order to determine vegetation units for the 509 sites, multivariate analysis was undertaken, but only after “singletons” (i.e., taxa only occurring in 1 of the 509 plots) were removed. Therefore, 272 taxa were thus removed from 166 of the plots. At the first pass, they identified 4 “supergroups’ corresponding to the major geomorphic (soil type) units. These were further broken into 30 floristic community types, some of which were further broken down to sub-types to produce a total of 43 subgroups.

It is common for vegetation surveys to use a dendrogram analysis to ascribe a vegetation unit which is closest to the 43 units as described in Gibson *et al.* (notwithstanding that the methods used in the SCP study may differ to those used in any particular vegetation survey), whereas a Bray Curtis analysis is frequently used, it is rare, if not unprecedented to obtain a value of 0, meaning the vegetation unit is identical to that described in the SCP study. This is mostly because this study had very broad findings due to most vegetation units having an insufficient number of quadrats sampled. The quadrat numbers sampled for each Floristic Community Group (FCG) ranged from 2 (FCG 14) to 67 (FCG 21). A rarefaction analysis for Lot 123 would thus predict that the assignment of vegetation units is more likely to be accurate in Group 21, than in group 2. However, the actual score in a Bray

Curtis analysis, can not be used to determine how accurate any particular analysis is, other than what FCG a quadrat is likely to be closest to.

Measurement of Biodiversity: C: **Using Gibson *et al.***

Notwithstanding that the findings of the SCP survey are very broad, as the study quantifies the mean species richness (i.e. number of different taxa present) and mean weed frequency, and the study deliberately selected sites in the best, most undisturbed condition. Therefore, the data can be reliably used in the "supergroups" with large number of quadrats (e.g. 142 in supergroup 3, 141 in supergroup 4) to determine the overall disturbance of quadrats today, by comparing the mean native species richness and weed frequency.

Furthermore, Gibson *et al* note that (p54) all sites in the 3rd major group are much more closely related to each other than in any other groups. It is noteworthy that all vegetation units at the Lot 123, Mortimer Rd are within the supergroup 3.

References:

Bioscience (2015). *Vegetation and Black Cockatoo Assessment.* Unpublished report for prepared for Mr I. Yujnovich. Submitted to the Index of Biodiversity Surveys for Assessments as IBSA-2023-0053.

Gibson, N., Keighery, B., Keighery, G., Burbidge, A., and Lyons, M., (1994). *A Floristic survey of the Swan Coastal Plain*, Unpublished Report for the Australian Heritage Commission prepared by then Department of Conservation and Land Management and the Conservation Council of Western Australia (Inc.).

Natural Area Management Consulting Services (2018). *Lot 123 Mortimer Road Flora and Vegetation Survey and Black Cockatoo Habitat Assessment.* Unpublished report for prepared for Mr I. Yujnovich.

Natural Area Management Consulting Services (2022). *Detailed Flora and Vegetation Survey Lot 123 Mortimer Road, Casuarina.* Unpublished report for prepared for Mr I. Yujnovich. Submitted to the Index of Biodiversity Surveys for Assessments as IBSA-2023-0049.

Tauss and Associates (2021). The flora, vegetation and wetlands of Lot 123 Mortimer Road, Casuarina in Western Australia — an independent assessment. Unpublished report for the City of Kwinana, Western Australia.