

**Assessing sampling plot shape and size on measures of tropical forest diversity:  
A simulation study**

Sophie You

**ABSTRACT**

Forest sampling is a crucial step in studies of species richness and diversity. Tropical forests require specialized sampling techniques because they have a greater numbers and density of species compared to deciduous forests. The inventory of stems from the 50 ha Barro Colorado Island plot in Panama provides a unique opportunity to learn about alpha diversity sampling in tropical forests and to compare different sampling approaches. The objectives of this study were to (i) determine the sampling scheme (size, shape, and number of plots) that best estimates the alpha diversity of the plot and (ii) minimize sampling effort required by finding a smaller sampling area that still estimates alpha diversity precisely. In order to create combinations of different sampling schemes, I varied the subplot shape, size and number of subplots. By running simulations of each combination of the sampling scheme characteristics, I calculated the estimated alpha diversity of the plot using the trees obtained under each of the schemes. The true values for Shannon-Weiner and Simpson indices, calculated using the entire dataset, were 3.6849 and 0.9999, respectively. By comparing estimated values to these true values, I found that the optimal individual sampling subplot characteristics were twenty rectangular subplots of 50 m<sup>2</sup>. However, the optimal sampling scheme is ten rectangular subplots of 50 m<sup>2</sup>, which reduced sampling area by half and increased the sampling efficiency while maintaining sufficient accuracy.

**KEYWORDS**

sampling design, alpha diversity, estimation, forest inventory

## INTRODUCTION

Forest sampling is a crucial component of forest species richness and diversity studies. Sampling is the first step in a study and decisions made in the sampling process will influence the rest of the study (Kenkel, Juhász-Nagy, & Podani, 1998). In evergreen forests of India and Malaysia, estimations of species richness and species diversity indices are all impacted by the sampling design (Gimaret-Carpentier, Pélissier, Pascal, & Houllier, 1998). Species richness represents the number of species in an area and diversity is a measure of both the number of species in an area as well as their relative abundance. Since both measures depend on the number of species present in an area, it is logical that an accurate sampling design is necessary. In forest sampling, as in many other sampling designs, researchers try to maximize both the efficiency and accuracy of these estimates. Accuracy usually increases with larger sample sizes (Condit, Hubbell, Lafrankie, Sukumar, Manokaran, et al., 1998) but efficient measurements attempt to achieve the best estimates with the fewest number of trees sampled (Laurance, Ferreira, Rankin-De Merona, & Hutchings, 1998). Choosing an appropriate sampling design for forests requires a balance between efficiency and accuracy. Most sampling schemes try to maximize both efficiency and accuracy, but sampling techniques also vary greatly depending on the research context of different study systems (Vanclay, 1998).

Since tropical forests exhibit species richness and species density than deciduous forests, they require specialized sampling techniques (McRoberts, Tomppo, & Næsset, 2010). For example, tropical forest tree species richness estimates increase with an increasing number of samples because of the high species diversity and number of rare species in these forest systems (Gimaret-Carpentier et al., 1998). Due to the unique nature of tropical forests, sampling design should be chosen based on both the site of the plot and the goals of the research (Kenkel et al., 1998). Though there have been studies examining the sampling schemes in tropical forests, including a study of diversity for pre-felling forest inventory (Potts, Kassim, Supardi, Tan, & Bossert, 2005), there has yet to be a study that determines the optimal sampling schemes to measure tropical forest diversity.

The 50 ha Barro Colorado Island (BCI) plot in Panama provides a unique opportunity to learn about alpha diversity sampling in tropical forests as well as to compare different sampling approaches. A complete inventory of stems in this plot was recorded in 1981 (CTFS, 2010).

Using the dataset, I created sampling schemes and compared the diversity indices calculated from these schemes to the true diversity value of the plot. I chose to investigate the alpha diversity of the plot because this metric helps to identify areas with high biodiversity (Potts et al., 2005) and inform conservation decisions by applying the findings about sampling procedures to other tropical forests.

The objectives of this study were to (i) determine the sampling scheme (size, shape, and number of plots) that best estimates the alpha diversity of the BCI plot and (ii) reduce sampling effort required by minimizing sample size necessary to estimate alpha diversity precisely. I hypothesized that (i) sampling schemes with a large number of large rectangular plots would give most precise estimates of diversity. Potts et al. (2001) found that line transects were most effective for sampling the largest number of species, so it follows that a plot covering a length of forest might be the best sampling plot type. Secondly, I hypothesized that (ii) a sampling scheme with a few medium sized plots of around 60 m<sup>2</sup> would be able to measure alpha diversity efficiently, because it will cover enough area to get a sizable sample of trees but not so large of an area to require extensive sampling. Since most estimates of species richness are able to evaluate differences between study sites and can used for comparisons between these sites (Palmer, 1990), most of the sampling schemes at BCI should yield results close to the true alpha diversity value.

## METHODS

### Study site

My study site is the 50 ha BCI plot in Panama (latitude 9.1543 N, longitude -79.8461 W). This 1500 ha island in the Panama Canal is part of the Barro Colorado Nature Monument, a protected reserve established in 1923 (CTFS, 2010). The island is considered a lowland tropical moist forest, containing a relatively large amount of well-known flora compared to plots of similar sizes around the world (CTFS, 2010). The entire plot contains nearly 300 species and over 200,000 stems (CTFS, 2010).

### *Data collection*

The online dataset from BCI (Hubbell, Condit, & Foster, 2005; Condit, 1998; Hubbell, Foster, O'Brien, Harms, Condit, et al., 1999) allowed for a calculation of the actual alpha diversity of the entire plot. In 1982, Smithsonian Tropical Research Institute and Princeton University researchers organized workers to measure all stems greater than 1 cm diameter breast height (dbh), tag the stems, and map the trees using coordinates relative to the southeast corner of the plot. Each of the species was also identified with a unique species code. The researchers conducted several subsequent resurveys; in my study, I used the 1995 resurvey data. Resurveying meant that any changes in the species dbh were recorded, new stems were recorded, and any deceased trees were noted and removed. Because this dataset contains updated records of all species in the plot, I was able to calculate the true value of alpha diversity using counts of the unique species in the dataset and the x and y location coordinates of the stems.

### *Creating sampling schemes*

In order to create combinations of different sampling schemes, I varied the subplot shape, size and number of subplots. The subplots were one of three shapes: squares, rectangles, or circles. I varied the size (i.e. area) of each individual subplots from 10 m<sup>2</sup> to 100 m<sup>2</sup>, with sizes at every 10 m<sup>2</sup> interval. I varied the number of subplots, with 1, 5, 10, 15, or 20 subplots. I created sampling schemes by making combinations of subplot shape, size and number of subplots until there was one scheme for each possible combination. Under each sampling scheme, I used all of the stems in each sampled subplot in alpha diversity calculations. All the combinations add up to 150 different sampling schemes, each of which I replicated 100 times in the statistical package R (R Development Core Team, version 2.9.2, 2009).

### *Estimating Alpha Diversity*

Using the trees obtained from the sampling schemes, I calculated the estimated alpha diversity of the plot using Shannon-Weiner and Simpson indices. I used two diversity indices to find the sampling scheme that was consistently the best estimator of the alpha diversity across

both indices. The Shannon-Weiner index (Eqn. 1) uses  $p_i$ , the relative abundance of each species,  $i$ , to calculate the biodiversity, summing over the total number of species ( $S$ ). Simpson's diversity index (Eqn. 2), takes the proportion of each species ( $n_i$ ) over the total number of species ( $N$ ) into account. These indices are both commonly used in calculating biodiversity and also provide more than one standard to test sampling schemes against. An ideal sampling scheme would be able to estimate both Shannon-Weiner and Simpson's diversity indices precisely and accurately.

**Eqn. 1. Shannon-Wiener Index.**

$$H' = -\sum_{i=1}^S (p_i \ln p_i)$$

**Eqn. 2. Simpson's Index.**

$$D = 1 - \frac{\sum_{i=1}^S n_i(n_i-1)}{N(N-1)}$$

## Data analysis

I calculated the true alpha diversity values using all the species in the 50 ha plot, and compared this value to the estimated mean and median values from each of the schemes. I plotted the estimated indices for each repetition of the scheme using boxplots. I found the medians and interquartile ranges for each of the schemes, as well as the minimum and maximum values. After looking at these statistics, I chose the best sampling schemes to approximate alpha diversity based on the accuracy of the true alpha diversity estimate. I then used qualitative comparisons to choose the sampling scheme that also minimized sampling effort.

## RESULTS

### Data set characteristics

The BCI plot contained a large number of species as well as high species diversity. A total of 306 unique species were present, adding up to 263,934 individual stems. The species

with the most individual stems in the plot was *Hybanthus prunifolius* with 46,175 stems, comprising 17.5% of the entire plot (Table 1).

**Table 1. Top 10 most populous species in the plot.**

<b>Species Name</b>	<b>Count</b>	<b>Percentage of Total Stems</b>
<i>Hybanthus prunifolius</i>	46175	17.5%
<i>Faramea occidentalis</i>	24733	9.4%
<i>Trichilia tuberculata</i>	13552	5.1%
<i>Desmopsis panamensis</i>	12959	4.9%
<i>Oenocarpus mapora</i>	8323	3.2%
<i>Alseis blackiana</i>	7930	3.0%
<i>Mouriri myrtilloides</i>	7145	2.7%
<i>Psychotria horizontalis</i>	6765	2.6%
<i>Piper cordulatum</i>	5365	2.0%
<i>Swartzia simplex</i>	5044	1.9%

The median value for the number of stems of a particular species was 113, with an interquartile range of 563.25.

### **Estimated alpha diversity**

Values for Shannon-Wiener's Diversity Index varied from a minimum of 1.0778 for the sampling scheme with one circular subplot of 10 m<sup>2</sup>, to a maximum of 3.7506 for the sampling scheme with twenty rectangular subplots of 50 m<sup>2</sup> (Figure 1a). The standard deviation for these estimates was 0.5643.

For Simpson's Diversity Index, values varied from a minimum of 0.3981 for the sampling scheme with one circular subplot of 10 m<sup>2</sup>, to a maximum of 0.9944 for the sampling scheme with twenty rectangular subplots of 100 m<sup>2</sup> (Figure 1b). The standard deviation for these estimates was 0.1250.

---

**A****B**

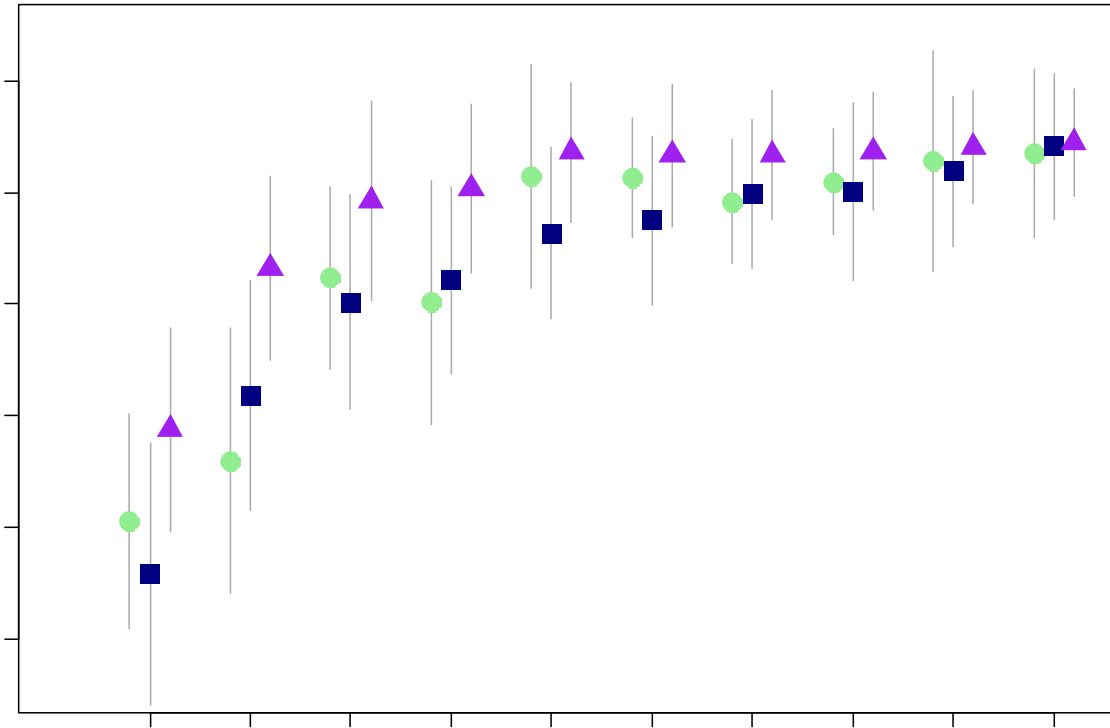
**Figure 1. Boxplots of estimates from all sampling schemes for (a) Shannon-Wiener Index and (b) Simpson's Diversity Index.** The bolded line represents the median, the box shows the 1<sup>st</sup> and 3<sup>rd</sup> quartiles, and the whiskers show data within 1.5 times the interquartile range. Any points outside of the box and whiskers are outliers.

### **Comparing true alpha diversity to estimated alpha diversity**

Calculated using the entire plot dataset, the true value for Shannon-Wiener's Diversity Index was 3.6849, and the true value for Simpson's Index was 0.9999.

#### *Subplot Size*

Estimates approach the true diversity value as subplot size increases (Figure 2). Schemes with subplots 50 m<sup>2</sup> or greater had estimates closest to the true values of the diversity indices. Standard deviations of estimates decreased as subplot size increased.



**Figure 2. Graph of Shannon Wiener Index estimates of sampling schemes with ten plots.** Circles, squares, and triangles represent the mean values of plots. The grey lines represent the standard deviation of each of the estimates, and the dashed line shows the true Shannon Wiener value of the plot.

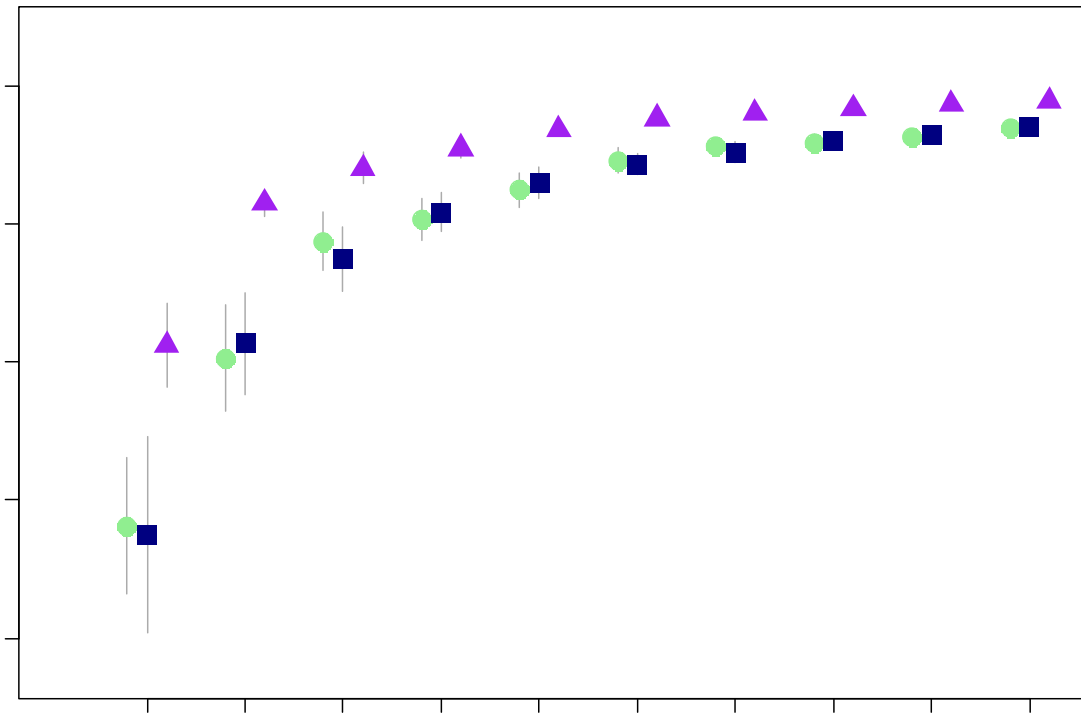
### *Subplot shape*

The rectangular subplots had more accurate estimates of the indices than equal sampling schemes with different shapes (Figure 2). Each of the three shapes had similar standard deviations.



*Number of subplots*

Sampling schemes with twenty subplots had the estimates closest to the true values of the indices (Figure 2, Figure 3). These schemes also had smaller standard deviations than schemes with fewer subplots.



**Figure 3. Graph of Simpson's Diversity Index estimates of sampling schemes with twenty plots.** Circles, squares, and triangles represent the mean values of plots. The grey lines represent the standard deviation of each of the estimates, and the dashed line shows the true Simpson's Diversity value of the plot.

## DISCUSSION

### **True alpha diversity**

The true values of alpha diversity indices confirm the high species richness of the BCI plot. The Shannon-Wiener diversity value of 3.6849 is considered very high in both species richness and evenness. Values of this index above 3.5 are considered indicators of high species richness and diversity, and values below 1.5 are considered indicators of low species richness and evenness (Magurran, 1988). Values outside of this range are possible, but not common. The Simpson's Index value of 0.9999 further validates the diversity of the BCI plot: a value of 0 represents no diversity, while a value of 1 represents infinite diversity (Magurran, 1988). Though the true values of alpha diversity are high, the ranges of the estimates found from the sampling schemes varied greatly. These high true diversity values as well as significant variability confirm the need for a unique scheme for estimating the high species diversity of the tropical forest.

### **Estimated alpha diversity**

The results found schemes that are best with respect to each separate sampling scheme component and also found some general trends in sampling schemes. First, my results showed that schemes covering larger areas better estimated diversity, which is consistent with findings in similarly sized tropical forest plots in Malaysia (Potts, 2005). However, it is unrealistic to sample many large plots in the field. The plot shape that gave the best estimates of alpha diversity was rectangular, which is consistent with findings in Malaysia where rectangular plots were able to sample the highest variety of species compared to other plot shapes (Potts, 2005). The rectangular plot shapes allow for a longer plot size to be sampled, which helps reach a greater number of species. While the trends support my hypotheses, another important consideration is the accuracy of the estimator – the optimal sampling scheme should not require sampling large area to get an accurate estimate.

The range and distribution of the estimated diversity indices suggest that at a certain threshold point, sampling additional plots does not increase the accuracy of the estimates. As sampled subplot area increases, both estimators of alpha diversity reach asymptotes. After these

threshold values, adding more subplots or increasing the size of the subplots no longer improves the estimate of alpha diversity. However, the large range of values indicate that there are sampling schemes that will not give good estimates of alpha diversity, and it is still important to sample enough of the plot area. Insight gained from these analyses of the separate components of the sampling scheme helped to determine the best sampling scheme.

### **Optimal sampling scheme**

Although the optimal individual sampling subplot characteristics were twenty rectangular subplots of 50 m<sup>2</sup>, the optimal sampling scheme is not one that combines these three aspects into one scheme. Since 50 m<sup>2</sup> was the threshold value for the sampling size, I chose the optimal scheme of ten rectangular subplots of 50 m<sup>2</sup> area. This reduced the sampling area by 50% and gives a Shannon-Weiner index estimate of 3.6718, which is within 0.35% of the true value, and a Simpson's Diversity Index estimate of 0.9657, which is within 3.45% of the true value. This sampling scheme would be more practical for field sampling in tropical forests, giving reasonably accurate diversity estimates with less sampling time required.

### **Implications**

I have found accurate estimates of alpha diversity and specific sampling techniques that are effective in the BCI plot, which can be used to verify previous and future works in this plot. By examining other sampling schemes and comparing them to the optimal scheme found, this study can provide a benchmark for comparison to other studies carried out in the plot. If a sampling scheme differs dramatically from the optimal one, this may serve as a "red flag" that the estimators may be incorrect.

Furthermore, the findings of this study have implications for tropical forest sampling and the verification of sampled areas across the world. The principles used in the optimal sampling scheme for the BCI plot can be applied to tropical forest plots in other areas with similar spatial variability of species. The findings help with the planning of future studies by providing better strategies for sampling in highly diverse tropical forests. They also help to minimize the time and effort needed to sample subplots.

Not only can accurate sampling schemes assist in finding diversity values for studies, but conservationists can also use results from diversity estimates to help guide management of tropical forests around the globe. Tropical forests should be preserved because their high species richness and diversity – there may be endemic or rare species present in these forests that cannot be found elsewhere in the world (Noor, Kean, Vun, & Mohamed-Hussein, 2011). Comprehending the diversity in an area will give a better idea of the importance of protecting these species. However, areas with high tropical forest diversity are usually difficult to protect and conserve (Condit et al., 2001). Logging in tropical forest may cause damage that cannot be undone even after 150 year periods of protection (Brown & Gurevitch, 2004), so conservation of rare species is crucial. Since using a given diversity estimate, regardless of accuracy, is an effective measure for comparing the diversity between different plots (Palmer, 1990), these estimates provide a deeper understanding of the sampling plots' relative alpha diversity.

### **Limitations and Future Studies**

Although we found an optimal sampling scheme for the BCI plot, there were some limitations regarding the dataset and the sampling schemes in this study that future studies could address. First, because only stems greater than 10 cm in dbh were sampled, smaller stems were left out of the study altogether. The calculated true diversity values do not take these stems into account, so the true values that the estimates were compared to could be inaccurate, specifically lacking the young saplings in the plot. A study of the BCI plot suggests that data from large plots should be double checked, because manual sampling and identification is easily a source of error in studies (Chave et al., 2004). A more accurate future study would compare estimated values to exact alpha diversity values calculated using every stem in the plot, though this is unrealistic for such a large plot. In addition, the randomly selected plots were not checked to see if they overlapped with each other. Since the sampling area never reached more than 0.05% of the plot, it is reasonable to assume that the plots do not overlap. However, a future study should definitely take this into account. Furthermore, because of computer processing limitations, the sampling schemes were only replicated 100 times. For the sampling schemes with high variance in estimator values, this could skew the data. Ideally, a future study could replicate each scheme thousands of times to reduce the variance. In addition, other studies could vary the shapes, sizes,

and number of subplots used to create a more comprehensive study that analyzes more sampling scheme variables.

In addition to limitations with the data and computation, there were also limitations with the applicability of the scheme to other tropical forests. One factor that I was not able to account for was the distance between subplots because the lack of research regarding the topic. Subplots that are closer together are more likely to have similar species than subplots that are further from each other, and this distance can affect the estimated values. Both the structure of the plot population and the distribution of species have a large impact on the diversity indices (Magnussen & Pelissier, 2005), so further research on this topic should take into account this structure and distribution to generate a more realistic estimator. Lastly, the accuracy of the optimal sampling scheme has yet to be applied to other plots and other indices. Future studies should apply this scheme to other tropical forests and verify its accuracy, as well as estimate other diversity indices. A long term objective is to apply this scheme across a variety of plots and a variety of indices and still accurately and precisely measure diversity of a tropical forest plot.

### **ACKNOWLEDGEMENTS**

I thank Tina Mendez, Kurt Spreyer, Seth Shonkoff, and in particular, Lara Roman, for their guidance, support, and advice throughout this entire process. Thank you to Dr. Matthew D. Potts for helping me conceptualizing this project and locate the BCI plot dataset, and to Matthew Luskin for giving me initial project ideas. I thank my ES 196 work group – Jim Gao, Daphne Szutu, and Michael Young – for their help in peer editing.

The BCI forest dynamics research project was made possible by National Science Foundation grants to Stephen P. Hubbell: DEB-0640386, DEB-0425651, DEB-0346488, DEB-0129874, DEB-00753102, DEB-9909347, DEB-9615226, DEB-9615226, DEB-9405933, DEB-9221033, DEB-9100058, DEB-8906869, DEB-8605042, DEB-8206992, DEB-7922197, support from the Center for Tropical Forest Science, the Smithsonian Tropical Research Institute, the John D. and Catherine T. MacArthur Foundation, the Mellon Foundation, the Celera Foundation, and numerous private individuals, and through the hard work of over 100 people from 10 countries over the past two decades. The plot project is part the Center for Tropical Forest Science, a global network of large-scale demographic tree plots.

## REFERENCES

- Brown, K. and Gurevitch, J. (2004). Long-term impacts of logging on forest diversity in Madagascar. *Proceedings of the National Academy of Sciences*, 101(16), 6045-6049.
- Chave, J., Condit, R., Aguilar, S., Hernandez, A., Lao, S., & Perez, R. (2004). Error propagation and scaling for tropical forest biomass estimates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 359(1443), 409-420.
- Condit, R. (1998). *Tropical Forest Census Plots*. Springer-Verlag and R. G. Landes Company, Berlin, Germany and Georgetown, Texas.
- Condit, R., Hubbell, S., Lafrankie, J., Sukumar, R., Manokaran, N., Foster R., & Ashton, P. (1996). Species-area and species-individual relationships for tropical trees: A comparison of three 50-ha plots. *Journal of Ecology*, 84(4), 549-562.
- Condit, R., Robinson, D., Ibanez D., Aguilar, S., Sanjur, A., Martinez, R., Stallard, R. F., Garcia, T., Angehr, G. R., Petit, L., Wright, S. J., Robinson, T. R., and Heckadon-Moreno, S. 2001. The status of the Panama Canal watershed and its biodiversity at the beginning of the 21st century. *BioScience* 51: 135-144.
- Center for Tropical Forest Science (CTFS). (2010). Smithsonian Tropical Research Institute. *Barro Colorado Island*. Retrieved November 13, 2010, from <http://www.ctfs.si.edu/site/Barro+Colorado+Island>.
- Gimaret-Carpentier, C., Pélissier, R., Pascal, J. & Houllier, F. (1998). Sampling strategies for the assessment of tree species diversity. *Journal of Vegetation Science*, 9, 161-172.
- Hubbell, S. P., Foster, R. B., O'Brien, S. T., Harms, K. E., Condit, R., Wechsler, B., Wright, S. J., and Loo de Lao, S. (1999). Light gap disturbances, recruitment limitation, and tree diversity in a neotropical forest. *Science*, 283, 554-557.
- Hubbell, S. P., Condit R., and Foster, R. B. (2005). Barro Colorado Forest Census Plot Data. Retrieved November 11, 2010 from <http://ctfs.amarb.harvard.edu/webatlas/datasets/bci>
- Kenkel, N., Juhász-Nagy, P., & Podani, J. (1989). On sampling procedures in population and community ecology. *Vegetatio*, 83, 195-207.
- Laurance, W., Ferreira, L., Rankin-De Merona, J., & Hutchings, R. (1998). Influence of plot shape on estimates of tree diversity and community composition in central Amazonia. *Biotropica*, 30(4), 662-665.
- Magnussen, S. and Pelissier, R. (2005). Sample-based estimation of tree species richness in a wet tropical forest compartment. *Proceedings of the Seventh Annual Forest Inventory and Analysis Symposium*, 269-272.

- Magurran, A. E. 1988. *Ecological Diversity and its Measurement*. Princeton University Press, Princeton, NJ.
- McRoberts, R., Tomppo, E., & Næsset, E. (2010). Advances and emerging issues in national forest inventories. *Scandinavian Journal of Forest Research*, 25(4), 368-381.
- Noor, N., Kean, C., Vun, Y., Mohamed-Hussein, Z. (2011). In vitro conservation of Malaysian biodiversity – achievements, challenges and future directions. *In Vitro Cellular & Developmental Biology – Plant*, 47(1), 26-36.
- Palmer, M. (1990). The estimation of species richness by extrapolation. *Ecology*, 71(3), 1195-1198.
- Potts, M., Kassim, A., Supardi, M., Tan, S., & Bossert, W. (2005). Sampling tree diversity in Malaysian tropical forests: an evaluation of a pre-felling inventory. *Forest Ecology and Management*, 205(1-3), 385-395.
- Potts, M., Plotkin, J., Lee, H., Manokaran, N., Ashtons, P., & Bossert, W. (2001). Sampling biodiversity: effects of plot shape. *The Malaysian Forester*, 64(1), 29-34.
- R Development Core Team (2009). R: A Language and Environment for Statistical Computing (Version 2.9.2) [Software]. Vienna, Austria. Available from <http://www.R-project.org>
- Vanclay, J. (1998). Towards more rigorous Assessment of Biodiversity. *Assessment of biodiversity for improved forest planning*, 51, 211-232.

## APPENDIX

The following code finds 100 different estimates for the sampling scheme of five circular subplots of 10 m<sup>2</sup> and 20 m<sup>2</sup> in R (R Development Core Team, version 2.9.2, 2009).

```

belong = function(a,b,n){
  n >= sqrt((xcoord[1]-a)^2+(ycoord[1]-b)^2) |
  n >= sqrt((xcoord[2]-a)^2+(ycoord[2]-b)^2) |
  n >= sqrt((xcoord[3]-a)^2+(ycoord[3]-b)^2) |
  n >= sqrt((xcoord[4]-a)^2+(ycoord[4]-b)^2) |
  n >= sqrt((xcoord[5]-a)^2+(ycoord[5]-b)^2)
}

#####

cr10.05sh<-rep(NA,100)
cr10.05sp<-rep(NA,100)
radius <- sqrt(10/pi)

for (j in 1:10){
  TF<- rep(NA, 263934)
  xcoord<-runif(1,0,1000)
  ycoord<-runif(1,0,500)

  for (i in 1:length(gx)){
    TF[i]<-belong(gx[i],gy[i],radius)
  }

  NUM<- length(Latin[which(TF)]) #total number of species in plot
  num <-length(unique(Latin[which(TF)])) #number of unique species in plot
  psubi<-summary(Latin[which(TF)])/length(Latin[which(TF)])

  cr10.05sh[j] <- (-sum(psubi*log(psubi),na.rm=TRUE))
  cr10.05sp[j] <- 1-(sum(num*(num-1))/(NUM*(NUM-1)))
}

#####

cr20.05sh<-rep(NA,100)
cr20.05sp<-rep(NA,100)
radius<-sqrt(20/pi)

for (j in 1:10){
  TF<- rep(NA, 263934)

```



```
xcoord<-runif(1,0,1000)
ycoord<-runif(1,0,500)

for (i in 1:length(gx)){
  TF[i]<-belong(gx[i],gy[i],radius)
}

NUM<- length(Latin[which(TF)]) #total number of species in plot
num <-length(unique(Latin[which(TF)])) #number of unique species in plot
psubi<-summary(Latin[which(TF)]/length(Latin[which(TF)])

cr20.05sh[j] <- (-sum(psubi*log(psubi),na.rm=TRUE))
cr20.05sp[j] <- 1-(sum(num*(num-1))/(NUM*(NUM-1)))
}
```

R Development Core Team (2009). R: A Language and Environment for Statistical Computing  
(Version 2.9.2) [Software]. Vienna, Austria. Available from <http://www.R-project.org>