

**AN EVALUATION OF SPECIES-AREA CURVES
CREATED FROM A SYSTEMATICALLY SAMPLED DATA SET
USING AN AUTOMATED ACCUMULATION TECHNIQUE**

by

Becci Dale

A RESEARCH PAPER

submitted to

THE GEOSCIENCES DEPARTMENT

in partial fulfillment of the
requirements for the
degree of

MASTER OF GEOSCIENCES

GEOGRAPHY PROGRAM

January 2002

DIRECTED BY
Dr. A. Jon Kimerling

ACKNOWLEDGEMENTS

Many thanks are due to all of the people that made my graduate education a fun and enlightening one. In particular, I would like to thank Dr. K. for being a wonderful mentor, editor, cartographer, and sailor. Thanks go to Julia Jones her patience and good advice, as well as being a lot of fun in the field (who knew professors scrambled over boulders with students?). I would also like to express gratitude toward Aaron Wolf for keeping food on my plate and giving me opportunities that I never could have imagined and Dr. Matzke for persistently convincing me to complete this paper. Jeannie Sifneos gets many thanks for her time, statistical expertise, and wedding advice. Mark Meyers kept me laughing and caffeinated for the duration of this research, as well as offering excellent advice on this and other projects. Many, many thanks go to my parents for their continued support throughout graduate school and life. It is impossible to thank Bill enough for all his help, support, and love that got me through graduate school. I'd also like to thank Mr. Ruyle, my 5th grade teacher for making learning fun, Mr. Heinzl, my 8th grade science teacher for teaching me that science is cool, Dr. Patricia Burton for giving me confidence in my scholastic abilities, Dr. Diane Janick-Buckner for being a great mentor and a wonderful biologist, and Dr. Aileen Buckley for introducing me to cartography. Lastly, I'd like to dedicate this paper to Denis White for his thoughtful persistence, patience with my "pontification", and tons of encouragement. Without his enthusiasm and dedication to this project this paper would not exist.

TABLE OF CONTENTS

INTRODUCTION.....	1
METHODS.....	7
Data Sets.....	7
Automated Accumulations and Regressions.....	9
Hypotheses and Methods for Data Set and Methodological Tests.....	10
<i>SPECIES - differences in species range sizes</i>	
<i>METHOD - nested and random accumulations</i>	
<i>AREA - the size of the accumulation area for regression</i>	
RESULTS AND DISCUSSION.....	19
<i>SPECIES - differences in species range sizes</i>	
<i>METHOD - nested and random accumulations</i>	
<i>AREA - the size of the accumulation area for regression</i>	
SUMMARY AND CONCLUSIONS.....	25
REFERENCES.....	27
APPENDICES.....	28
APPENDIX A: Species-area curve statistical results.....	28
APPENDIX B: Species-area curves.....	29
APPENDIX C: Histograms for east and west side taxonomic groups.....	34
APPENDIX D: Nested accumulation technique program.....	35
APPENDIX E: Random accumulation technique program.....	37
APPENDIX F: Regression curve fitting program.....	38
APPENDIX G: Histogram creation program.....	39

FIGURES

Figure 1	The first species-area curve.....	1
Figure 2	A species-area curve for islands in the Caribbean.....	2
Figure 3	Species richness and abundance.....	3
Figure 4	An island and a mainland species-area curve.....	5
Figure 5	The sampling grid.....	7
Figure 6	The study area.....	8
Figure 7	The east-west splitting of the data sets.....	9
Figure 8	Random accumulation method.....	15
Figure 9	Nested accumulation method.....	17

TABLES

Table 1	Occurrence status of the original TNC data sets.....	8
Table 2	The twenty data set/accumulation types.....	10
Table 3	RSRS index values for each data set.....	13
Table 4	The results for the SPECIES test.....	20
Table 5	The results for the METHOD test.....	21
Table 6	The results from the AREA test.....	24

AN EVALUATION OF SPECIES-AREA CURVES CREATED FROM A SYSTEMATICALLY SAMPLED DATA SET USING AN AUTOMATED ACCUMULATION TECHNIQUE

ABSTRACT: *This paper examines the influence of species, method, and area on automated species-area curve created from a systematically sampled data set. Because this is a new approach to species-area curves, three tests were run to evaluate the performance of the data sets and accumulation technique. The tests involved 1) range size, 2) accumulation method, and 3) amount of area used in the regression, and were evaluated using the slopes (z-values) of the resultant species-area curves.*

INTRODUCTION

The species-area curve is a statistical summary of the principle that as area increases, the number of species increases in a predictable manner. Although ecologists have understood the species-area relationship since the early eighteenth century (Figure 1), it wasn't until 1920 that Arrhenius adapted and applied an equation used for scaling morphology and metabolic processes to quantify the relationship (Brown and Lomolino, 1998). Considered by Schoener (1976) to be "one of community ecology's few laws", the species-area relationship is of great importance to both theoretical and applied

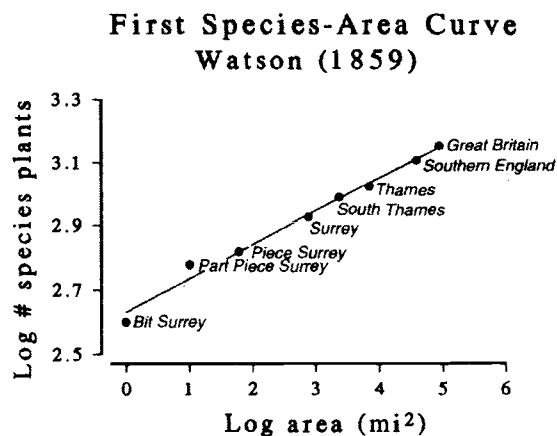


Figure 1. The oldest known example of an empirical ecological pattern. This recreation of the first species-area curve represents a continuum of area with the smallest area being a very tiny area in the county of Surry, England and the largest area being the entire island of Great Britain. (Reproduced from Rosenzweig, 1995).

ecology and biogeography. Within theoretical biogeography, the species-area relationship gives clues to the underlying processes that control species diversity. From an applied perspective, the relationship revealed by the curves can be used to estimate the amount of area needed to maintain a certain level or predict the approximate loss of biodiversity in an area (May and Stumpf, 2000).

Through landmark studies on the equilibrium theory of island biogeography published in their book *Island Biogeography* (1967), MacArthur and Wilson popularized the notion that as island areas increase, so do the number of species (Figure 2). In 1962, Frank Preston noted that the species-area relationship that had been applied to islands is only one case of the more general principle of the species-area relationship.

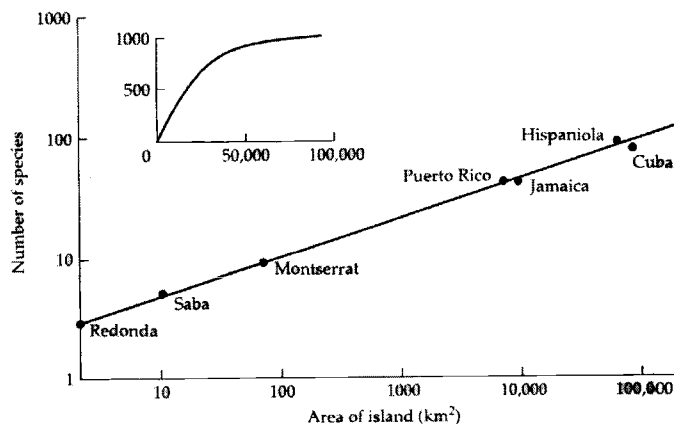


Figure 2. A species-area curve for islands in the Caribbean (Reproduced from Rosenzweig, 1995).

Preston realized that in a given area, species richness is higher for rare species than for abundant species. The relationship of richness and abundance nearly fits a normal curve when the x-axis is log transformed, meaning that the peak of the log-normal relationship is actually quite skewed toward rare animals when the data is transformed back to normal space (Figure 3). With this relationship plotted, it is evident that the most

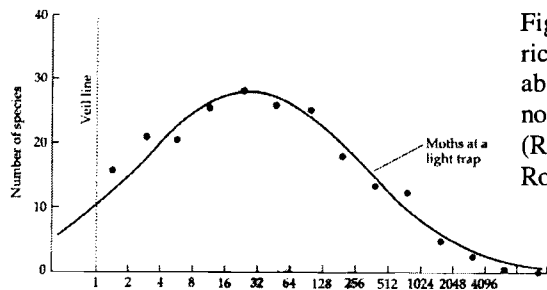


Figure 3. Species richness and abundance in log-normal space. (Reproduced from Rosenzweig, 1995)

rare species are not measured because the abundance, or number of individuals, of these species is so low they are

missed in the sampling. However, as more area is added, it is likely that the abundance of the rare species will increase to the level that they are counted, even though these same rare species were contained within the smaller area. It is the combination of the continual addition of rare species along with the accumulation of new habitats containing specialized species restricted to those habitats that account for the species-area relationship (Brown and Lomolino, 1998).

The species-area curve summarizes the species-area relationship through a regression of the number of species of a given taxonomic group and area. The regression plot is created by charting increasing area on the x-axis and number of species on the y-axis. This taxonomic group can be anything from a taxonomic kingdom, such as plants, to a very rich taxonomic family such as Coleoptera (beetles). The areas span a range from several islands to a continent, or even all of the world's landmasses.

Because the data are collected and axes are expressed using a logarithmic interval scale ($a/b = b/c = c/d \dots$), a power function best describes this relationship (Stevens, 1957). Although there are several regression equations that can be used to fit the species-area data, the most common equation is referred to as the Arrhenius equation or power [function] model (eq. 1). It has the form:

$$S = cA^z \text{ (eq. 1)}$$

where S is the number of species, c is the y-intercept, A is the area, and z is the slope (Plotkin *et. al.*, 2000). It should not be surprising that, the case of fauna, the equation often is expressed in log-log space to fit the regression to a straight line. The regression equation then has the form

$$\log S = z \log A + \log c \text{ (eq. 2)}$$

The species-area relationship is most commonly expressed through the z -value (slope) and c -value (y-intercept) of the regression. The z -value is of primary concern in the equation, as it quantifies the relationship of number of species to area. The c -value is also of interest as it measures the “starting” level of diversity for each curve. The y-intercept is specific to each taxonomic group and reflects the average number of species within an area the size of the smallest plot of the species-area curve (May and Stumpf, 2000).

There are two types of species-area curves: island and mainland (Figure 4). Island curves are accumulated by plotting different islands on the x-axis at their appropriate areas. The total number of species for each island is plotted on the y-axis against that island. Islands can be “real islands in the ocean, or virtual islands such as hilltops (where the surrounding lowland presents a barrier to many species), lakes, or wooded tracts surrounded by open land” (May and Stumpf, 2000). The data are fitted to

a curve through the use of the regression equation, and the z- and c- values are obtained for analyses.

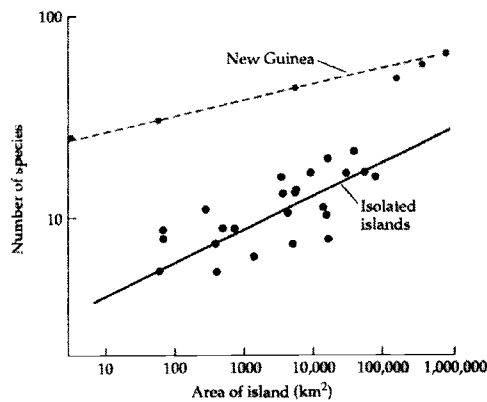


Figure 4. An island species-area curve and a mainland species-area curve.

Mainland curves are broken out into nested mainland curves and random mainland curves. Both types are accumulated differently than island curves. In island curves, neither the species nor the area is cumulative. This means that the total number of species for each island is plotted without regard to whether species on separate islands are the same species. In mainland curves, species and area are always cumulative. An amount of area is chosen as the smallest plot and the species for that area are recorded and plotted. As the area increases along the x-axis, it always contains the previously accumulated area. Because the area involved in mainland curves is cumulative, the species must also be cumulative. Island curves tend to have higher z-values than mainland curves because the smaller islands have a lower overall diversity that weights down the first part of the regression line, causing a steeper slope. The lower diversity occurs because it is easier for them to lose species and more difficult for them to recover locally extinct species than their well-connected mainland counterparts.

Nested mainland species-area curves begin with one main, or kernel, area and all of the subsequently accumulated area is adjacent to this kernel area. Random mainland

species-area curves also start out with a kernel area, but then add area randomly from non-contiguous areas. Rosenzweig (1995) warns that if area is accumulated using a random method, the z-value will be higher than it is with nested accumulations, a fact that could lead to inaccuracies in comparing species-area curves.

The bulk of the literature concerning species-area curves focuses on creating curves from specific data that has been collected by the researcher, comparing curves from many data sets over many different taxonomic distinctions and geographic areas, or on creating new and evaluating old regression equations. Although countless species-area curves have been constructed, few, if any, have been created using a systematically sampled data set and automated accumulation technique. As ecological sampling becomes more computerized through the use of satellite technology and other advances, and statistical software more usable, species area curves will become easier to create and perhaps more common.

This study combines data sets of regulated taxonomic distinctions and geographic areas with an automated accumulation technique. As these types of tools and research are used more commonly, it will become necessary to understand the biases we place in the data sets and the methods of accumulation. Although there are many factors influencing the outcome of species-area curves, three very basic factors can be used to evaluate this type of data set and automated accumulation method: species, method, and area. In this paper, we consider some of the effects of species range size, accumulation method, and the amount of area used in the regression on these species-area curves.

METHODS

Data Sets

The data sets to be used for this project are particularly well suited for understanding the effects of species, accumulation, and area on species-area curves. The study area is broken up into a regular sampling grid and the species have been recorded in a similar manner for all hexagonal areas. The basic structure of the data sets is a tessellation, or grid, of hexagons approximately 640 square kilometers in

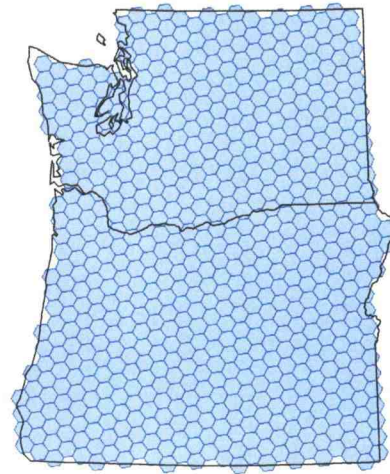


Figure 5. Hexagonal Sampling Grid

area, with the average distance from the center of one hexagon to the center of another being approximately 27 kilometers (Figure 5) (White *et.al.*, 1992). These data sets come from the first comprehensive nationwide database that includes standard range information from published literature and expert sources plus specific location data on plants and animals thought by The Nature Conservancy (TNC) to be of conservation concern (Master, 1996). The Nature Conservancy and its cooperating network of state Natural Heritage Programs developed the database and data sets; taxonomic experts in each state extensively reviewed the assignments of species to hexagons. Table 3 contains a listing of occurrence information recorded for each species sampled initially. These data were simplified into the data sets used in this study by considering all species ranked 80% or higher (probable to confident) to be present and all others to be absent.

Occurrence Status of Original TNC Data Sets	<ul style="list-style-type: none"> - confident or certain (>95% chance of occurrence) - predicted or probable (80% - 95% chance of occurrence) - possible (10% - 80% chance of occurrence) - not present or not significant (< 10% chance of occurrence)
--	--

Table 1. Occurrence status of the original TNC data sets.

The data sets are courtesy of the US EPA Corvallis Western Ecological Unit.

Washington and Oregon states have been broken up into a sampling grid of 660 adjacent hexagons, and all hexagons have at least 50% of their area within Washington and/or Oregon (Figure 6). The projection under the sampling grid is a Lambert azimuthal equal-area, ensuring that all hexagons retain their correct area. Vertebrate species for hexagons



Figure 6. Study Area

covering the states of Washington and Oregon are contained within these data sets. They are broken out into five taxonomic groups: amphibians, birds, fish, mammals, and reptiles, the original taxonomic groups that were delineated by The Nature Conservancy (White *et. al.*, 1999). Although studies often group reptiles and amphibians into one group termed herptiles, they remained as separate groups

throughout this study due to the significant differences in their species richness patterns.

High variance in the species richness data due to extreme ecological differences in eastern and western Washington and Oregon made it necessary to divide each of the five taxonomic group data sets into two separate data sets (Figure 7). A geographic information system (GIS) was used to delineate Washington and Oregon into an east and

west side using the crest of the Cascade Mountains as the dividing line. The study area border of the Interior Columbia Basin Ecological Monitoring Project was used as a baseline for the delineation, and hydrologic unit boundaries from the US Geological Survey HUC system were used elsewhere (<http://www.ICBEMP.gov>). This splitting created ten data sets from the original five--an east and west side data set for each taxonomic group.

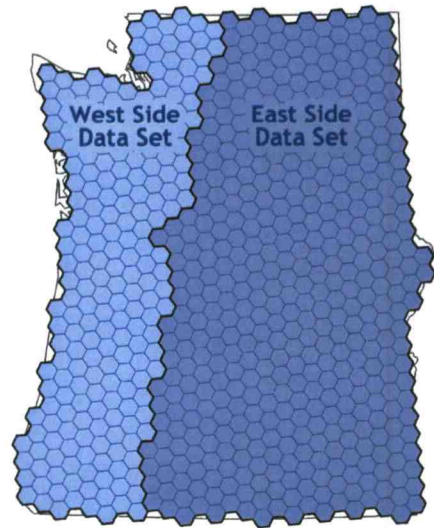


Figure 7. The delineation line used to split the data sets into the east side and west side.

Automated Accumulations and Regressions

The species-area curves were created using a program written by Denis White and Becci Dale in the statistical software R. The program automatically accumulates hexagons until all hexagons in the data set have been acquired. The ten aforementioned data sets were accumulated using two separate methods, leading to the twenty separate data set/accumulation types listed in Table 2. Although the information given in the sections below furnishes more details about the specifics of the accumulations, the data set/accumulation types, such as “bird west random”, were run independently of the hypotheses and had the output of twenty separate *z*- and *c*-values that were later used to evaluate the hypotheses. Each data set/accumulation type was run 200 times, with the data points for the number of species associated with each area stored in a table. The points were then averaged for each area and a curve was fit using the Arrhenius regression equation of this average number of species and the area. The Arrhenius

equation was chosen because it has a history of being the equation most often used, which makes the statistical outputs of the species-area curves comparable to other studies. As noted below, the regressions were run for two different sized areas, culminating in forty c- and z-values of the averaged accumulations that were returned for analyses.

Taxa	Geographic Region	Accumulation Method
Birds	East Side	Nested
Birds	East Side	Random
Birds	West Side	Nested
Birds	West Side	Random
Mammals	East Side	Nested
Mammals	East Side	Random
Mammals	West Side	Nested
Mammals	West Side	Random
Amphibians	East Side	Nested
Amphibians	East Side	Random
Amphibians	West Side	Nested
Amphibians	West Side	Random
Reptiles	East Side	Nested
Reptiles	East Side	Random
Reptiles	West Side	Nested
Reptiles	West Side	Random
Fish	East Side	Nested
Fish	East Side	Random
Fish	West Side	Nested
Fish	West Side	Random

Table 2. The twenty data set/accumulation types.

Hypotheses and Methods for Data Set and Methodological Tests

To better understand how well these data sets and automated accumulation methods performed, several hypotheses were used as tests. All of the hypotheses are mathematically valid and, therefore, can evaluate how well the data sets summarize the

species-area relationship and if the automated accumulation methods properly create the species-area curves.

SPECIES - differences in species range sizes

Species Hypothesis:

The relative z -values of species-area curves for different data sets should be predictable by the sizes of the species ranges within the taxonomic groups. Given regressions of log of number of species against log of area on many different taxonomic types, the taxonomic type with the greatest number of small species ranges and fewest number of large species ranges should have the highest z -value; the taxonomic type with the fewest number of small species ranges and greatest number of large species ranges should have the lowest z -value.

When many of the species ranges are large, it is very likely that the species ranges will overlap and many species will be recorded with one accumulation. The most extreme example of this would be if all species ranges were the size of the entire study area. Then, all species would be recorded with the first accumulation of data and there would be no more species accumulated throughout the remainder of the accumulations. This would result in a regression line with a slope of zero. Likewise, if each of the species ranges were exactly the size of an accumulation area and no species ranges overlapped, then every time an accumulation took place a new species would be recorded. This would result in a regression line with a slope of one. Thus, the smaller the species ranges, the higher the slope.

Species Test Methods:

To test the effect that species range size has on the curves, separate accumulations for each of the taxonomic groups were created and analyzed. The first column of Table 2 shows these taxonomic groups. Histograms were created for each of the ten data sets to determine the relative sizes of the species ranges (Appendix C). Because the east side data sets have 441 hexagons and the west side data sets have 219 hexagons, it was necessary to normalize the area in the histograms. There is also variance in the number of total species available for accumulation depending on the taxa. For example, there are 275 species of birds in Washington and Oregon, but only 33 species of amphibians. This made it necessary to normalize the number of species in the histograms as well. The axes of the histograms were divided into deciles such that for the east side, the 441 hexagons were divided into ten deciles, each representing approximately 44 hexagons. The species were divided in a similar manner such that if there were total 26 species, each decile would represent about two and a half species. To better summarize the data from the histograms, the Relative Species Range Size (RSRS) index was created by adding the proportions of species for the first three area deciles together (representing small species ranges) and dividing that number by the addition of the proportions of species for the last three deciles (representing large species ranges).

All of the taxa were then ranked based using their histograms and RSRS index value into a continuum from the taxa with the greatest number of small species ranges/lowest number of large species ranges (highest RSRS index value) to lowest number of small species ranges/greatest number of large species ranges (lowest RSRS index value). The z-values were then compared to see if the data held with the hypothesis

that those with greatest number of small species ranges/lowest number of large species ranges would have the highest z -values.

The creation of the RSRS index mentioned above resulted in the following ranking which predicts the order or z -values for the ten data sets (Table 3). The data set containing the most small species ranges and fewest large species ranges is ranked first:

Data Set	RSRS Index Rank	RSRS Index Value	
fish west	1	16.33	<i>Predicted steeper slope</i>
fish east	2	15.0	
amphibian east	3	10.0	
reptile west	4	2.5	
mammal east	5	2.13	
reptile east	6	1.88	
amphibian west	7	1.5	
mammal west	8	1.41	<i>Predicted flatter slope</i>
bird west	9	1.32	
bird east	10	1.05	

Table 3. Ranked Relative Species Range Size index values for each data set.

METHOD - nested and random accumulations

Method Hypothesis:

The z -value should be higher for randomly accumulated curves than for nested accumulated curves (Rosenzweig 1995). Following Tobler's first law of geography that all things are related, but closer things are more related than farther things, it is more likely that adjacent hexagons will contain similar species than that non-adjacent hexagons will contain similar species. By the nature of the algorithm, adjacent hexagons will be necessarily accumulated when using a nested accumulation method. On the other hand, there is a greater chance of accumulating non-adjacent hexagons when using a random

accumulation method. Therefore, during a nested accumulation, it is likely that the species contained within an added hexagon will contain a set of species similar to those of the previously accumulated hexagons. During a random accumulation, it is less likely that the species contained within an added hexagon will contain a set of species similar to those of the previously accumulated hexagons. The result of this should be a steeper slope, on average, for random accumulations because more new species would be encountered and accumulated than for the nested accumulations given the same amount of area.

Method Test Method:

To test the effect of accumulation method on these species-area curves, all separate data sets were run using nested and random accumulation methods. The accumulation type for the data set/accumulation type can be seen in the second row of Table 2. Because the area for both the random and nested accumulation methods are cumulative, the species are also cumulative. If an area (hexagon) is encountered that contains no species or only species that have already been recorded, then the area is accumulated, but the number of species remains the same.

The random accumulation method uses a random number generator to give an order of accumulation to all of the hexagons in the data set. The area is accumulated and the species are recorded in a table used later in the regressions.

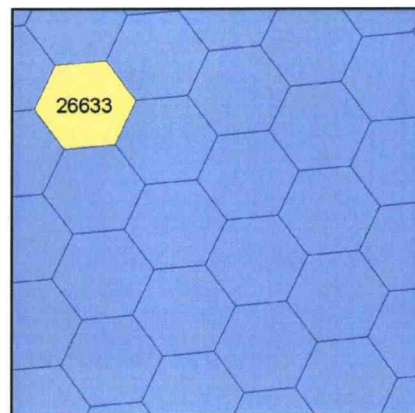
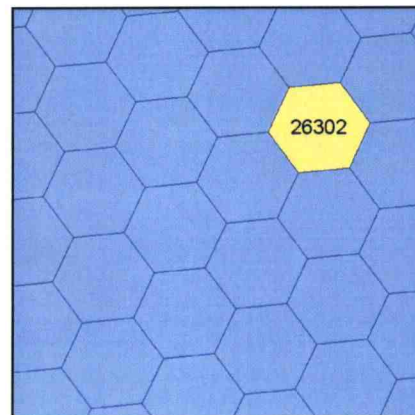
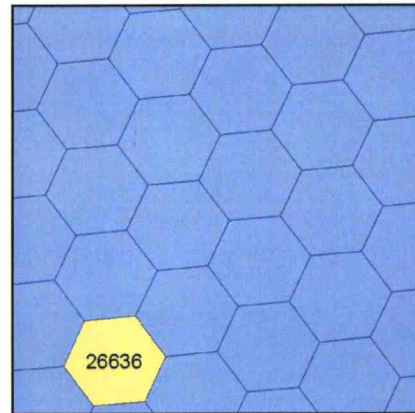
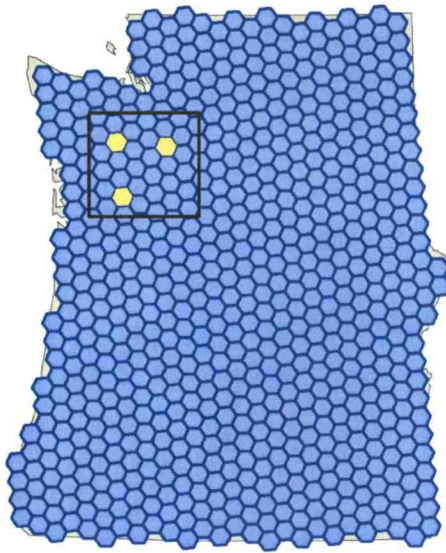


Figure 8. Random accumulation method.
Hexagons are accumulated in the order produced
through a random number generator

Random list: 26636, 26302, 26333

The nested accumulation method also relies on a hexagon being chosen at random, but only the first, or kernel, hexagon is selected this way. Area adjacent to the first hexagon is added such that the first kernel always remains approximately in the middle of the accumulation pattern. For hexagons that make up the border of the data set, there are less than six hexagons available, and it is possible to have as few as one available.

After the kernel hexagon (k1) has been selected, the program uses a hexagon adjacency look-up table, listing all hexagons and their adjacent neighbors, to determine which hexagons should be added. The hexagon adjacency table has two sections: a column listing the each hexagon as a kernel and six other columns listing each hexagon adjacent to the kernel. The accumulation method appears to “walk” around the kernel hexagon (k1), adding one hexagon at a time until all six adjacent hexagons have been accrued. The program then chooses the hexagon listed in the first column of the adjacency look-up table to become the new kernel (k2). The program then jumps to the line listing the new kernel hexagon (k2) in the kernel column and all hexagons listed in the table next to it are accumulated. Then, the program returns to the line listing the first kernel (k1) and chooses the second hexagon in the list as the new kernel (k3). This continues until all hexagons are accumulated. Throughout the accumulation the number of hexagons is recorded in a table that can be used later for creation of the species-area curves. The resulting z -values are then used to determine if the hypothesis that the random accumulations would have a greater slope is true.

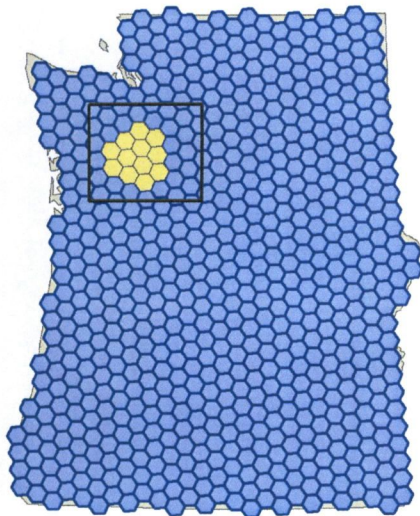


Figure 9. Nested accumulation method.

Step 1: Beginning with the area marked on the map to the left, hexagons are accumulated beginning with the first kernel, k1, assigned as hexagon number 26525.

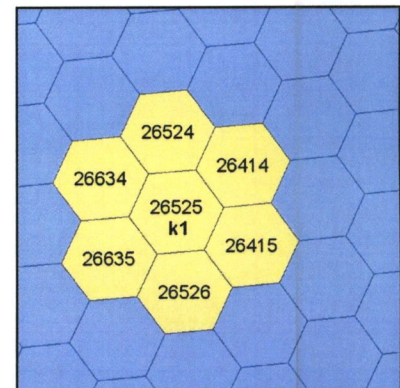
Step 2: After all six hexagons surrounding k1 have been accumulated, the program chooses the hexagon listed in the first adjacency column of k1 as the second kernel, or k2. Here, it is 26414. The hexagons that have already been accumulated (26302, 26303, and 26413) are ignored, and all others are added.

Step 3: The program returns to the line listing k1 in the kernel column. The hexagon listed in the second column is chosen as the next kernel, k3. All hexagons listed on the line with k3, 26415, as the kernel are then accumulated.

This process repeats until all hexagons are accumulated.

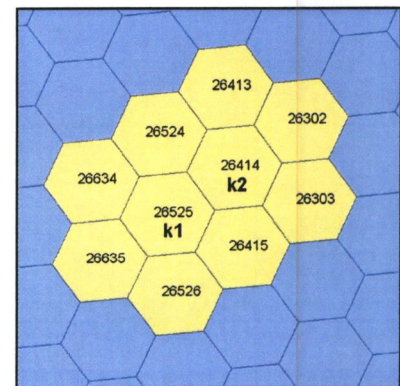
Step 1: k1

Kernal	Adjacent hex 1	Adjacent hex 2	Adjacent hex 3	Adjacent hex 4	Adjacent hex 5	Adjacent hex 6
26414	26302	26303	26415	26525	26524	26413
26415	26303	26304	26416	26526	26525	26414
26524	26413	26414	26525	26634	26633	26523
26525	26414	26415	26526	26635	26634	26524
26526	26415	26416	26527	26636	26635	26525
26634	26524	26525	26635	26743	26742	26633
26635	26525	26526	26636	26744	26743	26634



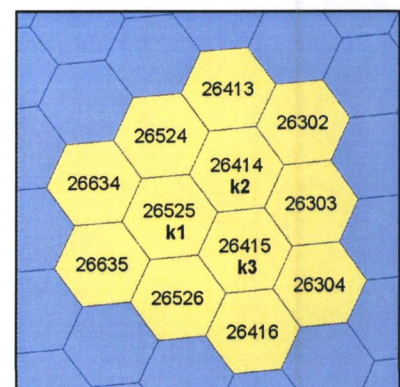
Step 2: k2

Kernal	Adjacent hex 1	Adjacent hex 2	Adjacent hex 3	Adjacent hex 4	Adjacent hex 5	Adjacent hex 6
26414	26302	26303	26415	26525	26524	26413
26415	26303	26304	26416	26526	26525	26414
26524	26413	26414	26525	26634	26633	26523
26525	26414	26415	26526	26635	26634	26524
26526	26415	26416	26527	26636	26635	26525
26634	26524	26525	26635	26743	26742	26633
26635	26525	26526	26636	26744	26743	26634



Step 3: k3

Kernal	Adjacent hex 1	Adjacent hex 2	Adjacent hex 3	Adjacent hex 4	Adjacent hex 5	Adjacent hex 6
26414	26302	26303	26415	26525	26524	26413
26415	26303	26304	26416	26526	26525	26414
26524	26413	26414	26525	26634	26633	26523
26525	26414	26415	26526	26635	26634	26524
26526	26415	26416	26527	26636	26635	26525
26634	26524	26525	26635	26743	26742	26633
26635	26525	26526	26636	26744	26743	26634



AREA - the size of the accumulation area for regression

Area Hypothesis:

The z-value generally should be lower for 200 accumulations than 50 accumulations. After the hexagon accumulations reach a certain point, generally near 200 hexagons, very few species have not been accumulated in most of the cases. This lack of adding new species causes many of the accumulated area measurements near the end of the accumulation to be repetitive, perhaps only gaining one new species for several accumulations in area. This repetition causes the data to level off at a certain point. When the regression line is fit to these data, the equation attempts to fit the entire line. The clustering of the repetitions unnecessarily causes a more gentle slope than would exist if the regression just took place over the hexagons where many additional species are being added. Therefore, if the accumulation is truncated, much of this repetition is removed. Because truncating at 200 hexagons retains more of the repetition, it should have a lower slope than truncating the accumulation at 50 hexagons, in which most of the repetition has been removed.

Area Test Method:

To test the effect of accumulated area on the species-area curves, the amount of area accumulated by the automated programs was truncated for the regressions. Although the program ran until all hexagons were accumulated, for the regression only first 50 and 200 hexagons were used for the 50 hexagon and 200 hexagon regressions, respectively. When the R program for the regression was run, it would either read in the first 50 or 200 of the areas and species accumulations and the regression would take place

using these data. The resulting z-values were then used to determine if the z-values were lower for 200 accumulations than 50 accumulations.

RESULTS AND DISCUSSION

Overall, the results of the tests agreed with the predictions made in the hypotheses. In a few cases, discrepancies made it necessary to examine the plotted species-area curves more carefully to gain insight into the nature of the results. The raw data for the z-value, c-value, and R-squared results are listed in Appendix A. Plots of the species-area curves resulting from this test data are in Appendix B.

SPECIES - differences in species range sizes

The species range sizes hypothesis states that the more small species ranges and less large species ranges a taxonomic group contains, the higher the z-value should be.

When the regressions are listed in the same order as the ranking from the RSRS index (Table 3), they tend to follow the predicted pattern as noted in Table 4.

RSRS Index Rank	Data Set			z-value	
				50 Hexagons	200 Hexagons
1	fish	west	nested	0.34	0.32
2	fish	east	nested	0.4	0.38
3	amphibian	east	nested	0.23	0.33
4	reptile	west	nested	0.13	0.21
5	mammal	east	nested	0.14	0.16
6	reptile	east	nested	0.16	0.18
7	amphibian	west	nested	0.15	0.19
8	mammal	west	nested	0.12	0.17
9	bird	west	nested	0.13	0.16
10	bird	east	nested	0.12	0.12
1	fish	west	random	0.33	0.23
2	fish	east	random	0.44	0.31
3	amphibian	east	random	0.35	0.27
4	reptile	west	random	0.17	0.09
5	mammal	east	random	0.19	0.11
6	reptile	east	random	0.2	0.12
7	amphibian	west	random	0.19	0.12
8	mammal	west	random	0.17	0.11
9	bird	west	random	0.15	0.12
10	bird	east	random	0.13	0.08

Table 4. The results for the SPECIES test

Although the descending order of the z -values does not exactly match the ranking of the data sets, it does follow a trend running from approximately 0.30 to 0.10. This result gives evidence that both the data set and automated accumulation method are valid for creating species-area curves. However, it is important to properly delineate the data sets to avoid splitting species ranges and creating “stragglers”.

METHOD - nested and random accumulations

The z -value was predicted higher for the random accumulation method than the nested accumulation method (Table 5). For the regressions with 50 hexagons, the slope

is always higher in random than nested, with the exception of “fish west,” for which the values of each were very similar (.34 and .33). However, the hypothesis falls short for the regressions with 200 hexagons. In these, the z-value is always higher in nested than random accumulations.

			z-value	
Nested Method			50 Hexagons	200 Hexagons
amphibian	west	nested	0.15	0.19
amphibian	east	nested	0.23	0.33
bird	west	nested	0.13	0.16
bird	east	nested	0.12	0.12
fish	west	nested	0.34	0.32
fish	east	nested	0.4	0.38
mammal	west	nested	0.12	0.17
mammal	east	nested	0.14	0.16
reptile	west	nested	0.13	0.21
reptile	east	nested	0.16	0.18
Random Method			50 Hexagons	200 Hexagons
amphibian	west	random	0.19	0.12
amphibian	east	random	0.35	0.27
bird	west	random	0.15	0.12
bird	east	random	0.13	0.08
fish	west	random	0.33	0.23
fish	east	random	0.44	0.31
mammal	west	random	0.17	0.11
mammal	east	random	0.19	0.11
reptile	west	random	0.17	0.09
reptile	east	random	0.2	0.12

Table 5. The results for the METHOD test

The random accumulations of 50 hexagons act as predicted. The random accumulations record more species in a smaller amount of area than do the nested accumulations, causing a higher z-value for random accumulations as predicted.

For both 50 and 200 hexagon accumulations, there is a critical number of species that is reached at some point in the accumulation, beyond which it takes many hexagons to accumulate just a few species. This leads to a greater amount of repetition of the species at the right end of the accumulation. The regression line is then tipped away from

the left side of the plot and toward the right end due to the many observations clustered together. Often for 50 hexagons, the regression equation fits the nested accumulations better than the random. This is because the random usually reaches the critical number of species before the nested does, so for the same amount of area accumulated, there is more clustering in the right end observations for the randomly accumulated data.

To understand why the z -values are higher in nested than random 200 hexagonal accumulations, an analysis of the plotted regression curves was necessary. Based on the species-area curve plots, it appears that the regression lines with 50 hexagons fit better than the regression lines for 200 hexagons. This is what accounts for the results being in opposition with the hypothesis. For the random accumulations with 200 hexagons, the slope seems to underestimate the steepness of the data at the left end of the plot. Here, there are many data points that lie below the regression line, but the many points clustered the right end of the line cause the data at the left end to have little effect on the regression. Although the left end points do have some weight, they are not properly accounted for by the regression. The opposite is true for the nested 200 hexagon regressions. It seems that the slope has been overestimated in this case, as many of the data points at the left end of the regression plot lie above the regression line. Again, these observations are not properly fit to the regression line. The outcome of the hyper-critical number of species effect on the 200 hexagon accumulations is that the random accumulations have a lower z -value than they otherwise would and the nested accumulations have a higher z -value than they otherwise would if the effect did not exist. Therefore, if the effect did not exist, it is quite possible that the random accumulations would have a higher z -value than the nested accumulations.

There is an additional, interesting difference between the random and nested accumulations- many of the nested accumulations have small curves in them toward the right hand side of the regression plots. This is most likely due to different assemblages of species being encountered in groups with the nested accumulations, whereas the random accumulations are smoothed of these assemblages. The assemblages could very possibly denote where the different ecoregions are.

These results reinforce the idea that the automated accumulation method is valid for creating species-area curves, although the preferred method of accumulation is nested over random.

AREA - the size of the accumulation area for regression

The z-value should be lower for 200 accumulations than 50 accumulations. For this test, the slope is always lower for 200 accumulations than 50 accumulations when the method is random (Table 6). The slope is lower for 50 accumulations than 200 accumulations in almost all cases when the method is nested. The only exceptions are “fish west nest” with a 50 hexagon value of 0.34 and a 200 hexagon value of 0.32, “fish east nest” with a 50 hexagon value of 0.40 and a 200 hexagon value of 0.38, and “bird east nest” where the values were equal at 0.12. This type of inversion of z-value levels only occurs when the fit of the regression line is as close to perfect as possible (R -squared = 1). These results fit with the theory that asserts the 200 hexagon regressions will have a poorer fit than the 50 hexagon regressions because of the hyper-critical number of species effect. The R -squared terms (see Appendix A) support the observation that the data fit the regression better for the 50 hexagon regressions than the 200 hexagon

regressions is supported by the R-squared terms. The *R*-squared terms for the 50 hexagon regressions are always higher or equal to that of the 200 hexagon regressions, but the 200 hexagon regressions are never higher than the 50 hexagon regressions. This means that area and clustering of data at the right end of the accumulation has a large effect on the outcome of the curves and the ability of the *z*-values to correctly predict what is truly happening with the data.

Data Set			z-value	
			50 Hexagons	200 Hexagons
amphibian	west	nested	0.15	0.19
amphibian	west	random	0.19	0.12
amphibian	east	nested	0.23	0.33
amphibian	east	random	0.35	0.27
bird	west	nested	0.13	0.16
bird	west	random	0.15	0.12
bird	east	nested	0.12	0.12
bird	east	random	0.13	0.08
fish	west	nested	0.34	0.32
fish	west	random	0.33	0.23
fish	east	nested	0.4	0.38
fish	east	random	0.44	0.31
mammal	west	nested	0.12	0.17
mammal	west	random	0.17	0.11
mammal	east	nested	0.14	0.16
mammal	east	random	0.19	0.11
reptile	west	nested	0.13	0.21
reptile	west	random	0.17	0.09
reptile	east	nested	0.16	0.18
reptile	east	random	0.2	0.12

Table 6. The results from the AREA test

These results give evidence that the amount of area used for the regressions is extremely important when creating these types of species-area curves. Because the results are in line with the hypothesis, it reaffirms that both the data set and automated accumulation method are valid for creating species-area curves.

SUMMARY AND CONCLUSIONS

Taken as a whole, both the data sets of regulated taxonomic distinctions and geographic areas and automated accumulation methods seem to work well to create species-area curves that have the same properties as traditionally used data sets and accumulation methods. The largest controlling factor in the results of the species-area curves was the regression equation used. The inability of the regression equation to properly fit the data is at the center of many of the unexpected results. Such as with the case of the nested accumulations and random accumulations, the outcomes of regressions for the same taxonomic groups in the same geographic area can be very different if the regression equation does not fit the data well. Although it did not always fit the data properly, this does not suggest that the equation should be “tossed out with the bath water”. There are many newly created regression formulas that contain extra parameters and can return a fit with a higher *R*-squared values, but as May and Stumpf (2000) stated, “Give me five parameters and I will fit elephants.” Instead of trying new equations when one does not fit, it is important to be sure the data being used are appropriate and non-redundant, as was the case with the 50 and 200 hexagon regressions.

In support of the outcome that these data sets and the automated accumulation method are valid ways of creating species-area curves, it is interesting and important to compare the *z*-values produced by this study and those reported by others. The *z*-values for species-area curves are generally listed in the literature as falling between .20 and .30. The *z*-values for these data sets fell between 0.8 and 0.44. The mean for all twenty regressions is 0.20, with a standard deviation of 0.095, and a median of 0.17.

Although species-area curves were first created nearly 150 years ago, they are still the best way to quantify the relationship of species richness to area (Rosenzweig, 1995). It is currently estimated that we are losing many species daily, an amount equal to the loss of biodiversity during the great natural disasters of history. Thus, species-area curves are important in both a theoretical light to understanding species-richness, and in an applied manner by giving managers a tool for maintaining biodiversity.

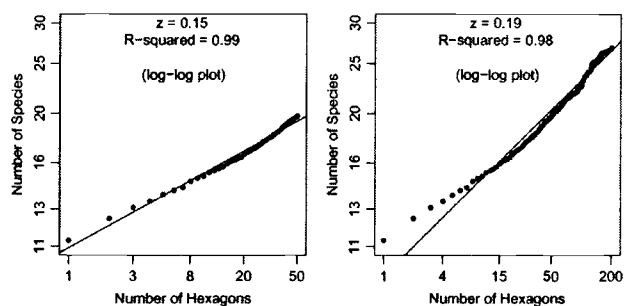
REFERENCES

- Brown, J.H., and Lomolino, M.V. 1998. *Biogeography*. Sunderland, Massachusetts: Sinauer Associates, Inc.
- Lyons, S.K., and Willig, M.R. 2002. Species Richness, Latitude, and Scale-Sensitivity. *Ecology* 83(1):47-58.
- Master, L. 1996. Predicting Distributions for Vertebrate Species: Some Observations. In *Gap Analysis: A Landscape Approach to Biodiversity Planning*. ed. J. M. Scott, T. H. Tear, and F. W. Davis, pp. 171-176. Bethesda, MD: American Society for Photogrammetry and Remote Sensing.
- May, R.M., and Stumpf, M.P.H. 2000. Species-Area Relations in Tropical Forests. *Science* 290:2084-2086.
- Plotkin, J.B. *et. al.* Predicting Species Diversity in Tropical Forests. *PNAS* 97(20):10850-18054
- Preston, F.W. 1962a. The Canonical Distribution of Commonness and Rarity: part I. *Ecology* 43:185-215.
- , 1962b. The Canonical Distribution of Commonness and Rarity: part II. *Ecology* 43: 410-432.
- Rosenzweig, M.L. 1995. *Species Diversity in Space and Time*. Cambridge, UK: Cambridge University Press.
- Stevens S. S. 1957. On the psychophysical law. *Psychological Review* 64:153-181.
- US Forest Service/Bureau of Land Management Interior Columbia Basin Ecological Management Project. <http://www.ICBEMP.gov> [accessed 3:00 p.m. June 15, 2001].
- White D., Kimerling, A.J., and Overton, W.S. 1992. Cartographic and Geometric Components of a Global Sampling Design for Environmental Monitoring. *Cartography and Geographic Information Systems* 19(1):5-22.
- White, D., Preston, E.M., Freemark, K.E., and Kiester, R.A. 1999. A Hierarchical Framework for Conserving Biodiversity. In *Landscape Ecological Analysis: Issues and Applications*. ed. J. M. Klopatek and R. H. Gardner, pp. 127-153. New York: Springer-Verlag.
- Williamson, M., Gason, K.J., and Lomdale, W.M. 2001. The Species-Area Relationship Does Not Have an Asymptote! *Journal of Biogeography* 28:827-830.

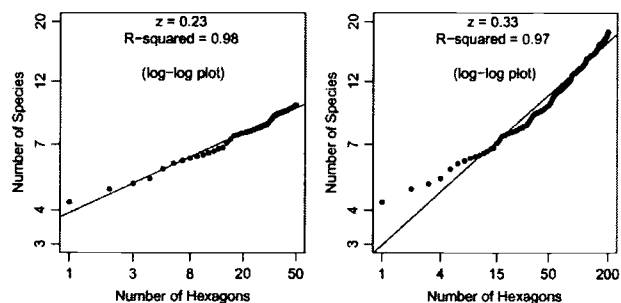
	z-value		c-value		R-squared	
	50 Hexagons	200 Hexagons	50 Hexagons	200 Hexagons	50 Hexagons	200 Hexagons
amphibian west nested	0.15	0.19	10.92	9.61	0.99	0.98
amphibian east nested	0.23	0.33	3.91	2.97	0.98	0.97
bird west nested	0.13	0.16	102.05	94.48	0.99	0.99
bird east nested	0.12	0.12	120.33	121.21	1.00	1.00
fish west nested	0.34	0.32	9.66	10.17	1.00	1.00
fish east nested	0.40	0.38	4.86	5.12	1.00	1.00
mammal west nested	0.12	0.17	48.67	41.84	1.00	0.95
mammal east nested	0.14	0.16	45.49	43.25	0.99	0.99
reptile west nested	0.13	0.21	8.46	6.83	1.00	0.93
reptile east nested	0.16	0.18	9.36	8.83	1.00	0.99
amphibian west random	0.19	0.12	13.61	16.60	0.96	0.86
amphibian east random	0.35	0.27	5.04	6.31	0.98	0.96
bird west random	0.15	0.12	114.44	126.31	0.98	0.96
bird east random	0.13	0.08	142.54	161.57	0.91	0.87
fish west random	0.33	0.23	13.65	17.60	0.94	0.93
fish east random	0.44	0.31	8.32	11.83	0.95	0.92
mammal west random	0.17	0.11	56.67	67.09	0.97	0.88
mammal east random	0.19	0.11	57.27	70.66	0.93	0.84
reptile west random	0.17	0.09	11.36	14.19	0.87	0.73
reptile east random	0.20	0.12	11.81	14.74	0.94	0.84

APPENDIX B. Species-area curves

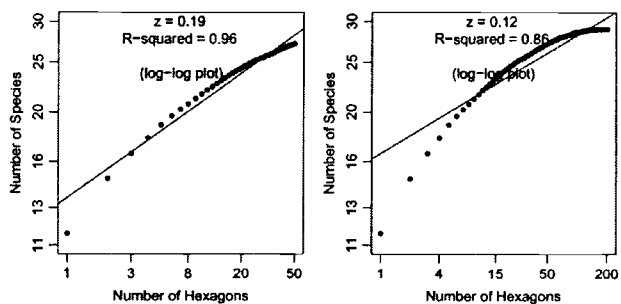
Amphibians, West, Nested



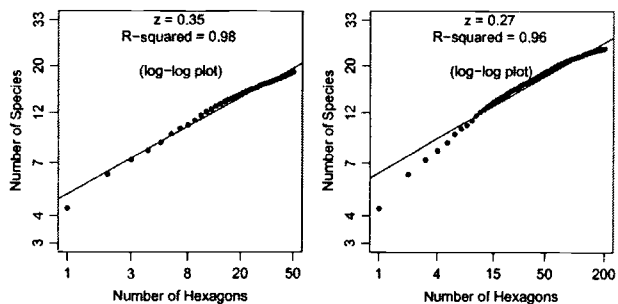
Amphibians, East, Nested



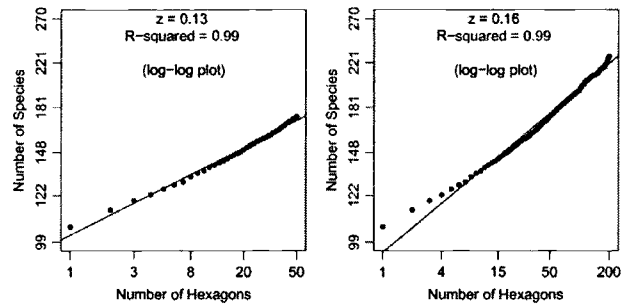
Amphibians, West, Random



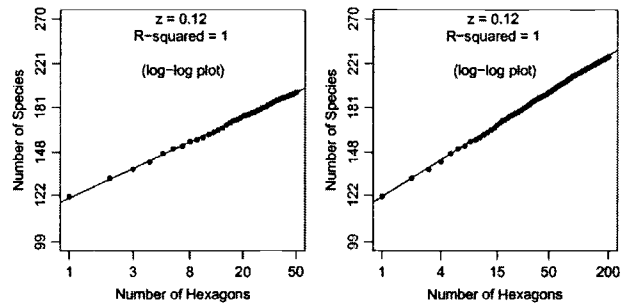
Amphibians, East, Random



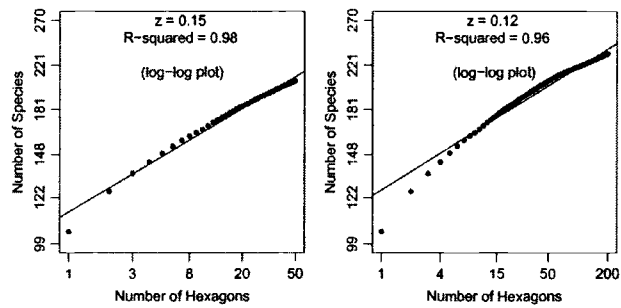
Birds, West, Nested



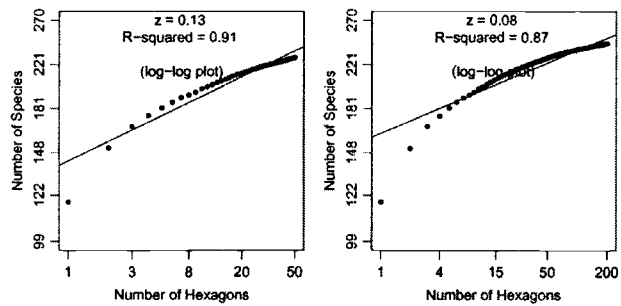
Birds, East, Nested



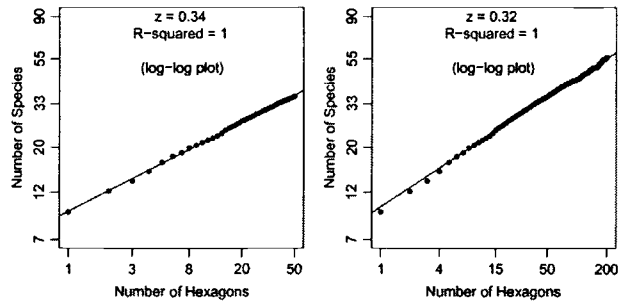
Birds, West, Random



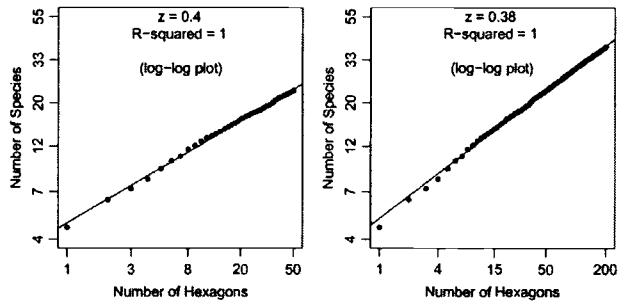
Birds, East, Random



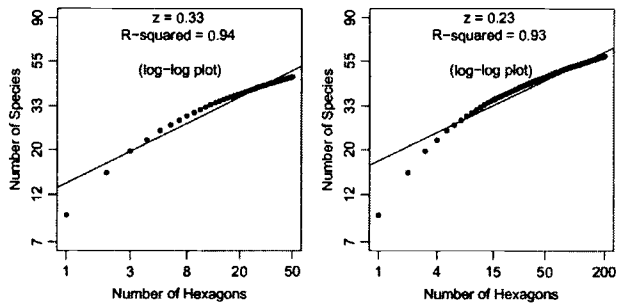
Fish, West, Nested



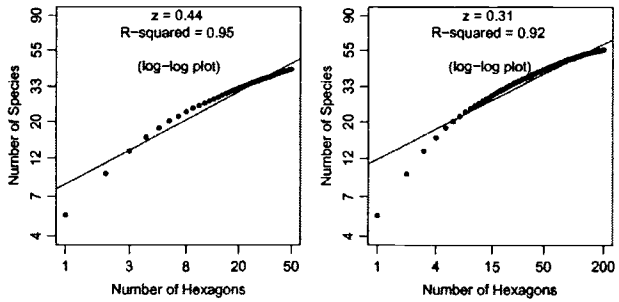
Fish, East, Nested



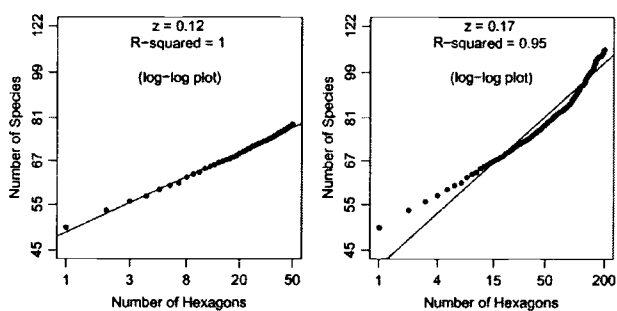
Fish, West, Random



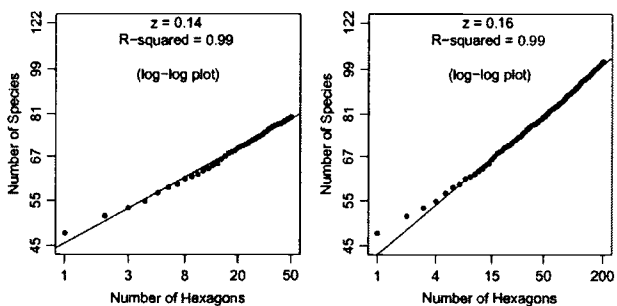
Fish, East, Random



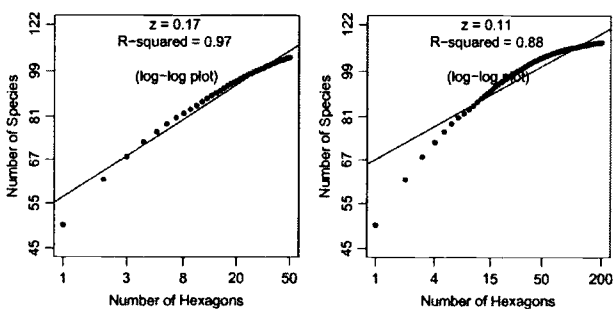
Mammals, West, Nested



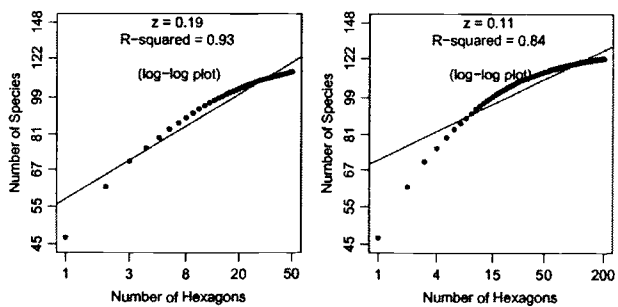
Mammals, East, Nested



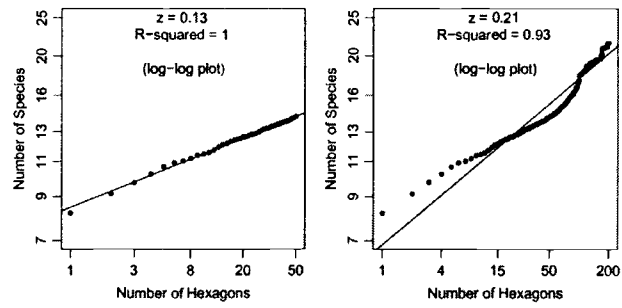
Mammals, West, Random



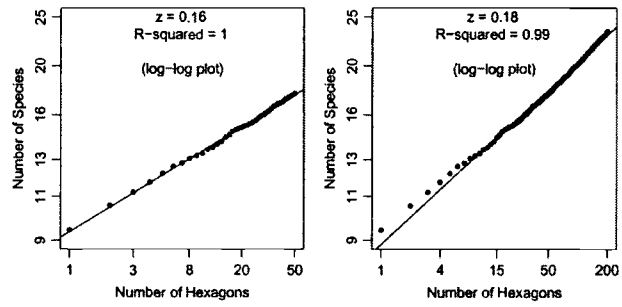
Mammals, East, Random



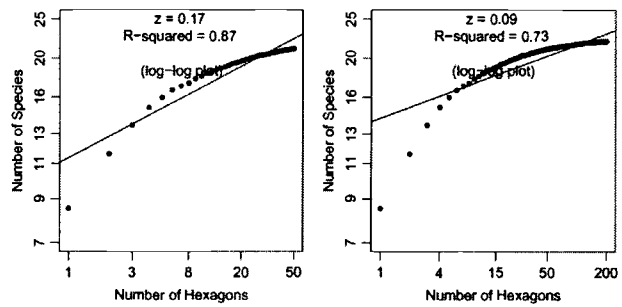
Reptiles, West, Nested



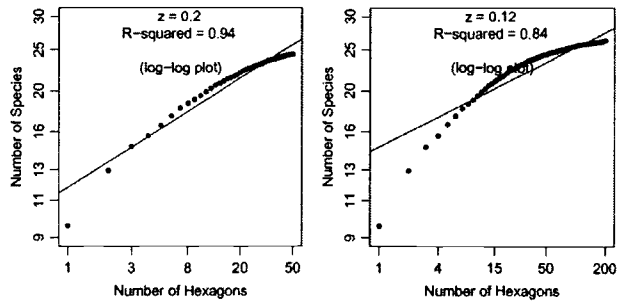
Reptiles, East, Nested



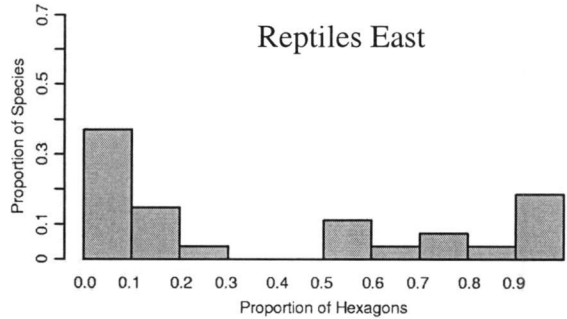
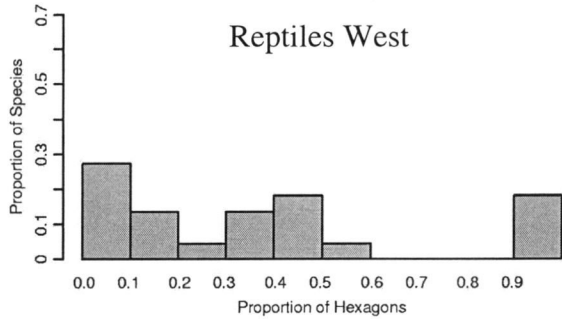
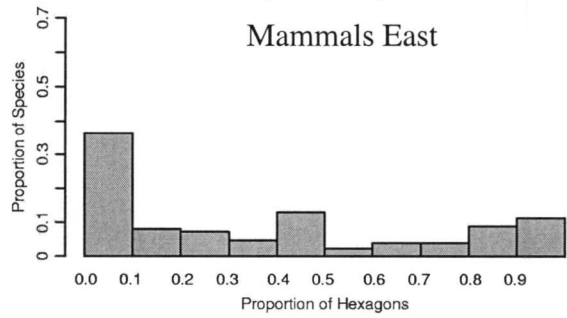
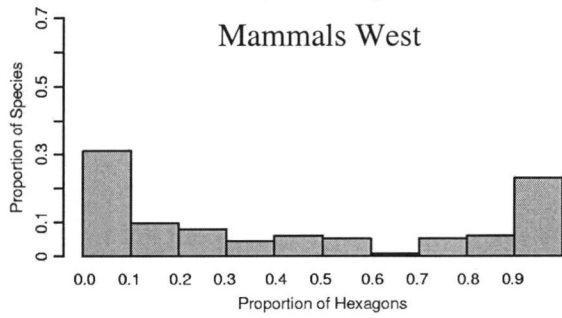
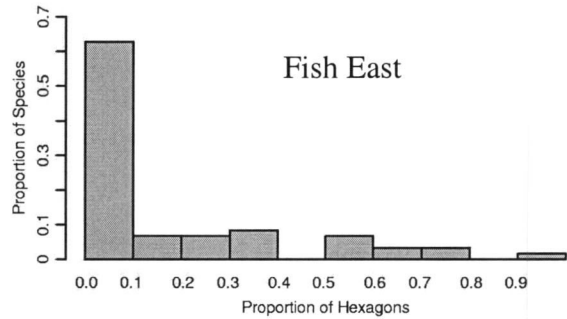
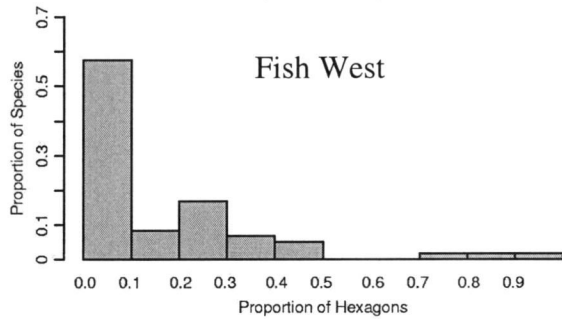
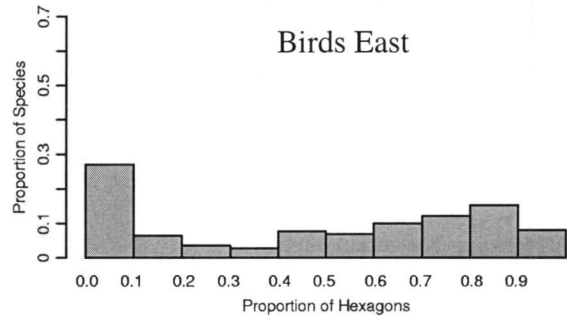
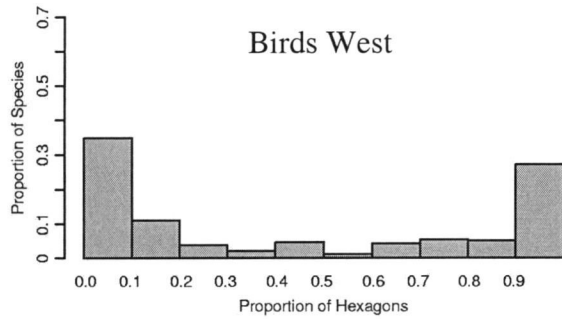
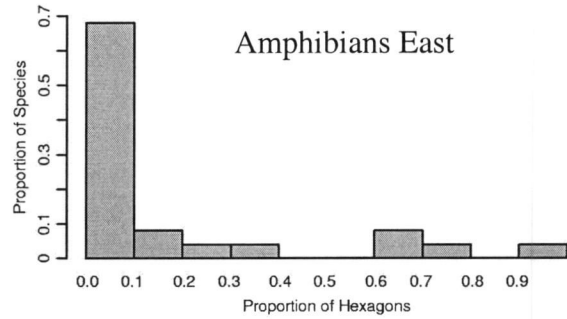
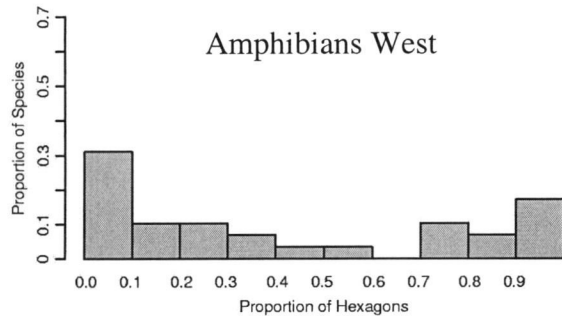
Reptiles, West, Random



Reptiles, East, Random



APPENDIX C. Histograms for east and west side taxonomic groups



APPENDIX D. Nested Accumulation Method Program

```
nrep <- 200 # desired number of repetitions of accumulation
set.seed (100)

hex.hex <- read.table ("hexhex.txt", header=TRUE) # adjacency table
hex.spp <- read.table ("hex_spp.txt", header=TRUE, sep=",") # species in hexagon data

row.names (hex.hex) <- hex.hex[,1]
hex.names <- row.names (hex.hex)
hex.hex <- hex.hex[,-1]

hex.spp <- hex.spp[hex.names,]
hex.spp <- hex.spp[, apply (hex.spp, 2, sum) != 0]

hex.hex <- as.matrix (hex.hex)
hex.hex[t (is.na (apply (hex.hex, 1, match, as.numeric
  (row.names (hex.hex))))))] <- 0

nhex <- nrow (hex.hex) # number of hexagons
nspp <- ncol (hex.spp) # number of species

nrep <- min (nrep, nhex) # don't do more reps than there are hexes
starts <- hex.names[sample (1:nhex, nrep)]
accum <- matrix (nspp, nrow=nrep, ncol=nhex) # accumulation totals for each rep

zeros <- apply (hex.spp, 1, sum)
zeros <- ifelse (zeros == 0, 1, 0)

for (i in 1:nrep) { # for each repetition

  hex.first <- starts[i] # random starting hex for this rep
  spp.set <- hex.spp[hex.first,] # which species found so far
  hex.set <- rep (0, nhex) # which hexes visited so far
  names (hex.set) <- hex.names
  hex.set[hex.first] <- 1
  hex.accum <- hex.first # list of visited hexagons, in order
  spp.accum <- sum (spp.set) # corresponding list of number of
    # accumulated species
  hex.queue <- hex.first # queue of hexagons to next visit

  while (length (hex.queue) > 0 && sum (spp.set) < nspp) {
    hex.center <- hex.queue[1]
    hex.queue <- hex.queue[-1]
    for (j in 1:6) {
      if (sum (spp.set) < nspp) {
        hex.next <- as.character (hex.hex[hex.center, j])
        if (hex.next != 0) {
          if (hex.set[hex.next] == 0) {
            hex.set[hex.next] <- 1
            spp.set <- as.integer (spp.set | hex.spp[hex.next,])
            hex.accum <- c(hex.accum, hex.next)
            spp.accum <- c(spp.accum, sum (spp.set))
            hex.queue <- c(hex.queue, hex.next)
          }
        }
      }
    }
  }
}
```

```

    }
  }
}
accum[i, 1:length(spp.accum)] <- spp.accum # accumulated richness
}

# hex.accum <- as.numeric (hex.accum)

meanspp <- apply (accum, 2, mean)
write.table (meanspp, "spp_accum.txt", quote = FALSE) # output table sum of number of species per
# accumulation area

```


APPENDIX E. Random Accumulation Method Program

```
nrep <- 200 # desired number of repetitions of accumulation
set.seed (100)

hex.spp <- read.table ("hex_spp.txt", header=TRUE, sep=",") # species in hexagon data
hex.hex <- read.table ("hexhex.txt", header=TRUE) # adjacency table
row.names (hex.hex) <- hex.hex[,1]
hex.names <- row.names (hex.hex)
# row.names (hex.spp) <- hex.spp[,1]
hex.hex <- hex.hex[,-1]

# if subset, reduce hex.spp to actual hexagons and species present
hex.spp <- hex.spp[hex.names,]
hex.spp <- hex.spp[, apply (hex.spp, 2, sum) != 0]

nhex <- nrow (hex.spp) # number of hexagons
nspp <- ncol (hex.spp) # number of species

nrep <- min (nrep, nhex) # don't do more reps than there are hexes
accum <- matrix (nspp, nrow=nrep, ncol=nhex) # accumulation totals for each rep

zeros <- apply (hex.spp, 1, sum)
zeros <- ifelse (zeros == 0, 1, 0)

for (i in 1:nrep) { # for each repetition

  hex.queue <- hex.names[sample (1:nhex, nhex)] # random search
  spp.set <- rep (0, nspp) # which species found so far
  hex.set <- rep (0, nhex) # which hexes visited so far
  names (hex.set) <- hex.names
  spp accum <- NULL # number of accumulated species
  j <- 0

  while (sum (hex.set) < nhex && sum (spp.set) < nspp) {
    j <- j + 1
    hex.next <- hex.queue[j]
    if (zeros[hex.next] == 0) {
      hex.set[hex.next] <- 1
      spp.set <- as.integer (spp.set | hex.spp[hex.next,])
      spp accum <- c(spp accum, sum (spp.set))
    }
  }
  accum[i, 1:length(spp accum)] <- spp accum # accumulated richness
}

meanspp <- apply (accum, 2, mean)
write.table (meanspp, "spp_accum.txt", quote = FALSE)
```

APPENDIX F. Regression Curve Fitting Program

```
sa.curve <- function (a, z, C) C * (a^z)

mean.read <- read.table ("spp_accum.txt", header=TRUE, sep=" ")
meanspp <- mean.read[,1]

postscript ("spp_geog_method.eps", horizontal=FALSE, onefile=FALSE)
par (mfrow=c(1,2), mgp=c(2,0.7,0), mar=c(3,3,3,1), omi= c(4, 0.5, 3, 0.5))
for (i in c(50, 200)) {
  nhex <- seq (i)
  fit <- lm (log (meanspp[nhex]) ~ log (nhex))
  C <- exp (fit$coefficients[1])
  z <- fit$coefficients[2]
  rsq <- summary (fit)$r.squared
  plot (meanspp[nhex], pch=20, xlab="Number of Hexagons", ylab="Number of Species")
  points (sapply (nhex, sa.curve, z, C), col="red", pch=20)
  mtext (paste ("z =", round(z,2), "\nc =", round(C,2), "\nR-squared =", round(rsq,2)),
    side=1, line=-3)
}
mtext ("Species Name, Geographic Extent, Accumulation Method", side=3, line=-1, cex=1.5,
outer=TRUE)
graphics.off ()
```

APPENDIX G. Histogram Creation Program

```
# Amphibians
hex.hex <- read.table ("west_hexhex.txt", header=TRUE)
hex.spp <- read.table ("hex_amph.txt", header=TRUE, sep=",")
row.names (hex.hex) <- hex.hex[,1]
hex.names <- row.names (hex.hex)
hex.hex <- hex.hex[,-1]
# reduce hex.spp to actual hexagons and species present
hex.spp <- hex.spp[hex.names,]
hex.spp <- hex.spp[, apply (hex.spp, 2, sum) != 0]
amph.west <- apply (hex.spp,2,sum)

hex.hex <- read.table ("east_hexhex.txt", header=TRUE)
hex.spp <- read.table ("hex_amph.txt", header=TRUE, sep=",")
row.names (hex.hex) <- hex.hex[,1]
hex.names <- row.names (hex.hex)
hex.hex <- hex.hex[,-1]
# reduce hex.spp to actual hexagons and species present
hex.spp <- hex.spp[hex.names,]
hex.spp <- hex.spp[, apply (hex.spp, 2, sum) != 0]
amph.east <- apply (hex.spp,2,sum)

# Birds
hex.hex <- read.table ("west_hexhex.txt", header=TRUE)
hex.spp <- read.table ("hex_bird.txt", header=TRUE, sep=",")
row.names (hex.hex) <- hex.hex[,1]
hex.names <- row.names (hex.hex)
hex.hex <- hex.hex[,-1]
# reduce hex.spp to actual hexagons and species present
hex.spp <- hex.spp[hex.names,]
hex.spp <- hex.spp[, apply (hex.spp, 2, sum) != 0]
bird.west <- apply (hex.spp,2,sum)

hex.hex <- read.table ("east_hexhex.txt", header=TRUE)
hex.spp <- read.table ("hex_bird.txt", header=TRUE, sep=",")
row.names (hex.hex) <- hex.hex[,1]
hex.names <- row.names (hex.hex)
hex.hex <- hex.hex[,-1]
# reduce hex.spp to actual hexagons and species present
hex.spp <- hex.spp[hex.names,]
hex.spp <- hex.spp[, apply (hex.spp, 2, sum) != 0]
bird.east <- apply (hex.spp,2,sum)

# Fish
hex.hex <- read.table ("west_hexhex.txt", header=TRUE)
hex.spp <- read.table ("hex_fish.txt", header=TRUE, sep=",")
row.names (hex.hex) <- hex.hex[,1]
hex.names <- row.names (hex.hex)
hex.hex <- hex.hex[,-1]
# reduce hex.spp to actual hexagons and species present
hex.spp <- hex.spp[hex.names,]
hex.spp <- hex.spp[, apply (hex.spp, 2, sum) != 0]
fish.west <- apply (hex.spp,2,sum)
```

```

hex.hex <- read.table ("east_hexhex.txt", header=TRUE)
hex.spp <- read.table ("hex_fish.txt", header=TRUE, sep=",")
row.names (hex.hex) <- hex.hex[,1]
hex.names <- row.names (hex.hex)
hex.hex <- hex.hex[,-1]
# reduce hex.spp to actual hexagons and species present
hex.spp <- hex.spp[hex.names,]
hex.spp <- hex.spp[, apply (hex.spp, 2, sum) != 0]
fish.east <- apply (hex.spp,2,sum)

# Mammals
hex.hex <- read.table ("west_hexhex.txt", header=TRUE)
hex.spp <- read.table ("hex_mamm.txt", header=TRUE, sep=",")
row.names (hex.hex) <- hex.hex[,1]
hex.names <- row.names (hex.hex)
hex.hex <- hex.hex[,-1]
# reduce hex.spp to actual hexagons and species present
hex.spp <- hex.spp[hex.names,]
hex.spp <- hex.spp[, apply (hex.spp, 2, sum) != 0]
mamm.west <- apply (hex.spp,2,sum)

hex.hex <- read.table ("east_hexhex.txt", header=TRUE)
hex.spp <- read.table ("hex_mamm.txt", header=TRUE, sep=",")
row.names (hex.hex) <- hex.hex[,1]
hex.names <- row.names (hex.hex)
hex.hex <- hex.hex[,-1]
# reduce hex.spp to actual hexagons and species present
hex.spp <- hex.spp[hex.names,]
hex.spp <- hex.spp[, apply (hex.spp, 2, sum) != 0]
mamm.east <- apply (hex.spp,2,sum)

# Reptiles
hex.hex <- read.table ("west_hexhex.txt", header=TRUE)
hex.spp <- read.table ("hex_rept.txt", header=TRUE, sep=",")
row.names (hex.hex) <- hex.hex[,1]
hex.names <- row.names (hex.hex)
hex.hex <- hex.hex[,-1]
# reduce hex.spp to actual hexagons and species present
hex.spp <- hex.spp[hex.names,]
hex.spp <- hex.spp[, apply (hex.spp, 2, sum) != 0]
rept.west <- apply (hex.spp,2,sum)

hex.hex <- read.table ("east_hexhex.txt", header=TRUE)
hex.spp <- read.table ("hex_rept.txt", header=TRUE, sep=",")
row.names (hex.hex) <- hex.hex[,1]
hex.names <- row.names (hex.hex)
hex.hex <- hex.hex[,-1]
# reduce hex.spp to actual hexagons and species present
hex.spp <- hex.spp[hex.names,]
hex.spp <- hex.spp[, apply (hex.spp, 2, sum) != 0]
rept.east <- apply (hex.spp,2,sum)

amph.west <- amph.west / max (amph.west)
aw <- cut (amph.west, breaks=(0:10)/10, labels=FALSE)
awp <- table (aw) / sum (table (aw))

```

```

amph.east <- amph.east / max (amph.east)
ae <- cut (amph.east, breaks=(0:10)/10, labels=FALSE)
aep<- table (ae) / sum (table (ae))

bird.east <- bird.east / max (bird.east)
be <- cut (bird.east, breaks=(0:10)/10, labels=FALSE)
bep <- table (be) / sum (table (be))

bird.west <- bird.west / max (bird.west)
bw <- cut (bird.west, breaks=(0:10)/10, labels=FALSE)
bwp <- table (bw) / sum (table (bw))

fish.west <- fish.west / max (fish.west)
fw <- cut (fish.west, breaks=(0:10)/10, labels=FALSE)
fwp <- table (fw) / sum (table (fw))

fish.east <- fish.east / max (fish.east)
fe <- cut (fish.east, breaks=(0:10)/10, labels=FALSE)
fep<- table (fe) / sum (table (fe))

mamm.west <- mamm.west / max (mamm.west)
mw <- cut (mamm.west, breaks=(0:10)/10, labels=FALSE)
mwp <- table (mw) / sum (table (mw))

mamm.east <- mamm.east / max (mamm.east)
me <- cut (mamm.east, breaks=(0:10)/10, labels=FALSE)
mep<- table (me) / sum (table (me))

rept.west <- rept.west / max (rept.west)
rw <- cut (rept.west, breaks=(0:10)/10, labels=FALSE)
rwp <- table (rw) / sum (table (rw))

rept.east <- rept.east / max (rept.east)
re <- cut (rept.east, breaks=(0:10)/10, labels=FALSE)
rep<- table (re) / sum (table (re))

# then put in array of dimension 5, 2, 10, i.e., 5 taxa by 2 regions
# by 10 deciles

taxa <- c("amph", "bird", "fish", "mamm", "rept")
regions <- c("west", "east")
deciles <- array (0, dim = c(5,2,10),
  dimnames = list(taxa, regions, as.character(seq(10))))

deciles[1,1,as.integer(names(awp))] <- awp
deciles[1,2,as.integer(names(aep))] <- aep
deciles[2,1,as.integer(names(bwp))] <- bwp
deciles[2,2,as.integer(names(bep))] <- bep
deciles[3,1,as.integer(names(fwp))] <- fwp
deciles[3,2,as.integer(names(fep))] <- fep
deciles[4,1,as.integer(names(mwp))] <- mwp
deciles[4,2,as.integer(names(mep))] <- mep
deciles[5,1,as.integer(names(rwp))] <- rwp

```

```

deciles[5,2,as.integer(names(rep))] <- rep

ymax <- round (max (deciles) + 0.05, 1)

postscript (file="hist.eps", horizontal=FALSE, onefile=FALSE,
  paper="special", width=8.5, height=11)
par (mfrow=c(5,2), mgp=c(2,0.7,0), mar=c(3,4,1,1), tcl=0)
for (i in 1:5) for (j in 1:2) {
  barplot (deciles[i,j,], ylim=c(0,ymax), col="gray", axes=FALSE,
    space=0,
    names.arg=c("0.0 0.1","0.1 0.2","0.2 0.3","0.3 0.4","0.4 0.5",
      "0.5 0.6","0.6 0.7","0.7 0.8","0.8 0.9","1.0"),
    xlab="Proportion of Hexagons",
    ylab="Proportion of Species")
  # axis (side=1, at=c(0,2,4,6,8,10), labels=c(0,.2,.4,.6,.8,1.0))
  axis (side=2, at=seq(0,ymax,by=0.1), tcl=-0.5,
    labels=as.character(seq(0,ymax,by=0.1)))
  text (5, ymax, paste(taxa[i], regions[j]), cex=1.5)
}
graphics.off ()

```