Statistical Variations in Bird Survey Data From 2000 – 2007

Rachael Millings

Office of Science, Science Undergraduate Laboratory Internship

State University of New York at Stony Brook

Stony Brook, New York

Brookhaven National Laboratory

Upton, New York

August 15, 2008

Written under the oversight of Dr. Timothy M. Green and Jennifer Higbie in the Environmental & Waste Management Services division at Brookhaven National Laboratory.

Student: ____
Signature

Research Advisor: ____
Signature

Research Advisor: ____
Signature

ABSTRACT

Statistical Variations in Bird Survey Data From 2000 to 2007. RACHAEL MILLINGS (State University of New York at Stony Brook, Stony Brook, NY 11794) TIMOTHY M. GREEN AND JENNIFER HIGBIE (Brookhaven National Laboratory, Upton, NY 11973).

Since 2000, the Environmental & Waste Management Services (EWMS) Division at Brookhaven National Laboratory (BNL) has conducted monthly bird surveys from March through October using a point-count method. The purpose of these surveys is to identify new species in the area; monitor changes in species populations, as well as changes in parasitic species populations; create accurate habitat models; and evaluate the effect of land management on various species. To accomplish this, statistical analysis of the data collected from 2000 through 2007 was performed using Statistix 9, an analytical statistics program. The data were organized into smaller files by year, family, and species. Biodiversity indices were calculated for each year. Next, descriptive statistics files, such as scatter plots and pie charts were created to gain a visual perspective of how the number of birds observed varied with respect to weather and temporal variables, as well as how biodiversity indices varied by year. A one-way analysis of variance (ANOVA) test and/or a Kruskal-Wallis nonparametric ANOVA test was then performed on selected data sets to determine the significance of any variation from the mean. Finally, a Geographical Information System was used to identify any correlation between species population variations and habitat availability. The results showed that the brown-headed cowbird (*Molothrus ater*) and the ovenbird (*Seiurus aurocapillus*) experienced population changes during the 2000-2007 period (α – 0.05). Both increases and decreases in the total sum of birds and bird species observed with respect to season were determined. ANOVA testing also showed that the diversity of the Biology Fields and Peconic River transects was greater than that of the South and Z-Path transects (α – 0.05). According to the Pearson correlation test, a strong correlation of

0.6581 exists between the species richness of a transect and vegetation type representing more that 30 percent of the station area. Using these results, the EWMS division at BNL can make informed management decisions concerning the experimental design of the survey method and the oversight of bird species populations and habitats.

1. INTRODUCTION

Bird species populations play a significant role in defining the ecological character of an ecosystem. Statistical variations in the populations are often good indicators of less obvious changes in other factors affecting the ecosystem. They also provide a relatively accurate source of data from which to determine the biodiversity of the ecosystem.

The United States Geographical Survey (USGS), an organization that monitors changes in American natural resources, potential hazards and the environment[1], has recognized the importance of bird populations to American ecosystems. Subsequently, in the early 1960's, the USGS began organizing regular surveys of breeding birds along established survey routes. These surveys are conducted mostly by qualified birdwatchers using the point-count method to identify the species and number of the birds observed. The data is collected and statistically analyzed to provide an accurate understanding of the behavior of breeding bird population in North America.

In May 2000, Timothy Green, of the Environmental & Waste Management Services (EWMS) division at the Brookhaven National Laboratory (BNL), began conducting similar bird surveys on a monthly basis from the spring through the fall of each year. The purpose of the surveys was to identify new species in the area; monitor changes in species populations, as well as changes in parasitic species populations; create accurate habitat models; and evaluate the effect of land management on various species. Statistical analysis was required to assist in accurately evaluating the data collected from these surveys, and in making informed management decisions concerning the results.

2. METHOD OF ANALYSIS

The bird surveys were conducted using a point count method. Six transects, or permanent point count survey routes, were established through varying habitats on the BNL. Throughout these transects, 28 point count locations, or stations, were spaced approximately 1000 ft. apart to ensure that no observation would be repeated at an adjacent location. Each transect was regularly visited during the last week of the month from March through October initially (2000-2002), then from April through September (2003-2007). At the beginning of each transect survey, standard weather data and the start time was recorded using a Kestral 4000 portable weather station. Each station was visited for a 5-minute count period during which all birds identified by call or by sight were recorded by all observers (usually two). For the purposes of this project, the transects will be identified in this paper by name - Biology Fields (BF), East Trenches (ET), North (NT), Peconic River (PR), South (ST), and Z-Path (ZP). The stations were identified by the initials of the respective transect and a number.

The analysis was primarily conducted using a statistical analysis program, Statistix 9. The data were imported from Excel into Statistix 9, and reformatted as necessary after all entry errors were corrected. Then smaller data files were created to isolate observations within a particular group, i.e., species, family, or year.

For population analysis, all species with 50 or more observations, and all families with 75 or more observations were separated for investigation. Using descriptive and summary statistics, i.e. error bar charts, pie charts, and scatter plots, a visual representation was created of the number of birds sighted per observation within each of the separated species and families with respect to the average dew point, relative humidity, and temperature, the month, and the year. For further analysis, each species data file was truncated to include only those observations taken in the April-September period, since several years did not have observations taken in March and/or October. To determine if the mean number of birds per observation had changed significantly over the eight-year period, an analysis of variance (ANOVA) test was performed for each species. For species with a normal distribution of the relevant data (number of birds per observation), a one-way ANOVA test was used, while a Kruskal-Wallis ANOVA test was used for species that did not have a normal distribution of the relevant data. For each species with a normal distribution, a means bar plot was also created to provide a visual representation of the data. Finally, to determine if population statistics would support the general monthly behavior of breeding birds in the northern American states, the mean monthly sum of bird species observed for each year was graphed as a scatter plot for annual comparison, and a one-way ANOVA test was performed on the mean total number of birds per observation with respect to month.

For biodiversity and habitat analysis, a Geographical Information System (GIS) was used to identify the types of vegetation within the 500-foot radius of each transect station. These data were recorded in a separate data file, along with the number of species observed per transect, and the number of birds observed within a species, with respect to transect. Using this data, six biodiversity indices (species richness; Simpson's index, Simpson's diversity index, Simpson's reciprocal index, Shannon-Wiener diversity index, and Simpson's evenness index) were calculated for each transect station, and descriptive statistics were used to visually compare the results with respect to transect station. Then the data were grouped by transect, and a one-way ANOVA test was performed for each biodiversity index with respect to transect. Six biodiversity indices (those indicated above) were also calculated for each of the eight years. Descriptive graphs and charts were created to give a rough estimation of any annual changes. Due to the distribution of the data, ANOVA tests with respect to year could not be performed. Finally, the percentage of each type of vegetation was calculated for each transect station to aid further analysis. The vegetation types considered were scarlet oak / heath forest; pitch pine / white oak forest; pitch pine / mixed oak / heath forest; planted white pine forest; red maple / mesic heath forest; red maple / scarlet oak / mesic heath forest; successional areas; grassland; cattail marsh; water; disturbed areas; buildings; and roads (Figure 1). The transect stations were numbered from 1 (smallest species richness) to 28 (greatest species richness), and the vegetation types were placed in three numbered groups (1 through 3) with 1 containing types most present in stations with small species richness and 3 containing types most present in stations with large species richness. Then a Pearson correlation test was performed to determine the strength of the correlation between the types of vegetation within a transect station and the station's species richness (sum of species observed during the eight year period).

3. INTERPRETATION OF ANALYSIS RESULTS

A. Population

Of the 34 species separated for analysis, 2 showed significant changes in population size at $\alpha = 0.05$ during the eight-year period. Both changes showed a negative growth rate, which, in the case of *Molothrus ater* (brown-headed cowbird) a parasitic species, was a desired occurrence (Figure 2 and Chart 1).

Annual comparison of the mean monthly sum of bird species observed for each year revealed that, in general, the mean sum of species observed each month gradually increased from 0 in March to approximately 50 in May and June, before gradually decreasing as October approached (Figures 3 and 4). This was expected, since breeding birds are mostly in the northern part of America during mid-spring through early fall.

The ANOVA test performed on the mean total number of birds per observation with respect to month showed that the largest number of birds (approximately 2.5 – 4) were observed in the migratory months of March, September and October, while the mean number of the other months gradually decreased to approximately 1.75 in the nesting months of May and June as those months were approached from either March or October (Figure 5 and Chart 2). This trend also supported the general behavior of breeding bird populations in northern American states.

B. Habitat

The transect station biodiversity index varied with respect to the station. ANOVA tests performed on data from each transect station with respect to one of six biodiversity indices revealed that the diversity of the Biology Fields and the Peconic River transects was greater than that of the South and Z-Path transects at the $\alpha = 0.05$ (Figure 6, Chart 3, and Table 1). The diversity of the East Trenches and North transects were both smaller than that of the Biology Fields at the same level of significance.

Using Pearson's correlation test to compare the species richness of a transect station to the vegetation within the transect station, and considering only those types representing 30 percent or more of the station area, it was found that a strong correlation of 0.6581 exists. This indicates that stations with 30 percent or more of the area represented by water, cattail marshes, red maple / mesic heath, and red maple / scarlet oak / mesic heath forests are more likely to have a greater species richness than those stations with 30 percent or more of the area represented by scarlet oak / heath, pitch pine / white oak, pitch pine / mixed oak / heath, or white pine forests.

Descriptive statistics revealed that the annual bird biodiversity is high. Both diversity and evenness indices indicate a relatively diverse and even population (Figure 7 and Table 2). While

ANOVA tests could not be performed on the annual distribution of biodiversity indices, a cursory evaluation of the indices does not reveal any drastic annual changes. Further work is required to determine the existence of annual changes with more statistically rigorous analysis.

4. FUTURE WORK & CONCLUSION

The goal of this project was to identify statistical variations in bird survey data collected over an eight-year period at BNL. The results of statistical analysis performed on the data with reference to population and habitat revealed that the bird species population is relatively stable, with few changes, and quite diverse. Also, species richness has a strong positive correlation to wetlands, indicating that water is a significant factor to consider when examining bird biodiversity on the BNL site.

Future work should include adapting the survey method to include density estimation during observation, e.g., distance sampling, using double observers, or calculating a detection probability. Such adaptations would greatly enhance the reliability of the data as accurate samples of the population[2]. Using time series analysis to analyze regular annual changes and combining variables during analysis, viz., examining changes in habitat with respect to year *and* transect station, would also increase the ability of the analysis results to accurately predict future trends.

5. REFERENCES

1. Purpose Statement. United States Geographical Survey http://usgs.gov

2. Statistical Approaches to the Analysis of Point Count Data: A Little Extra Information Can Go a Long Way. Farnsworth et al. USDA Forest Service Gen. Tech. Rep. PSW-GTR-191. 2005.

FIGURES

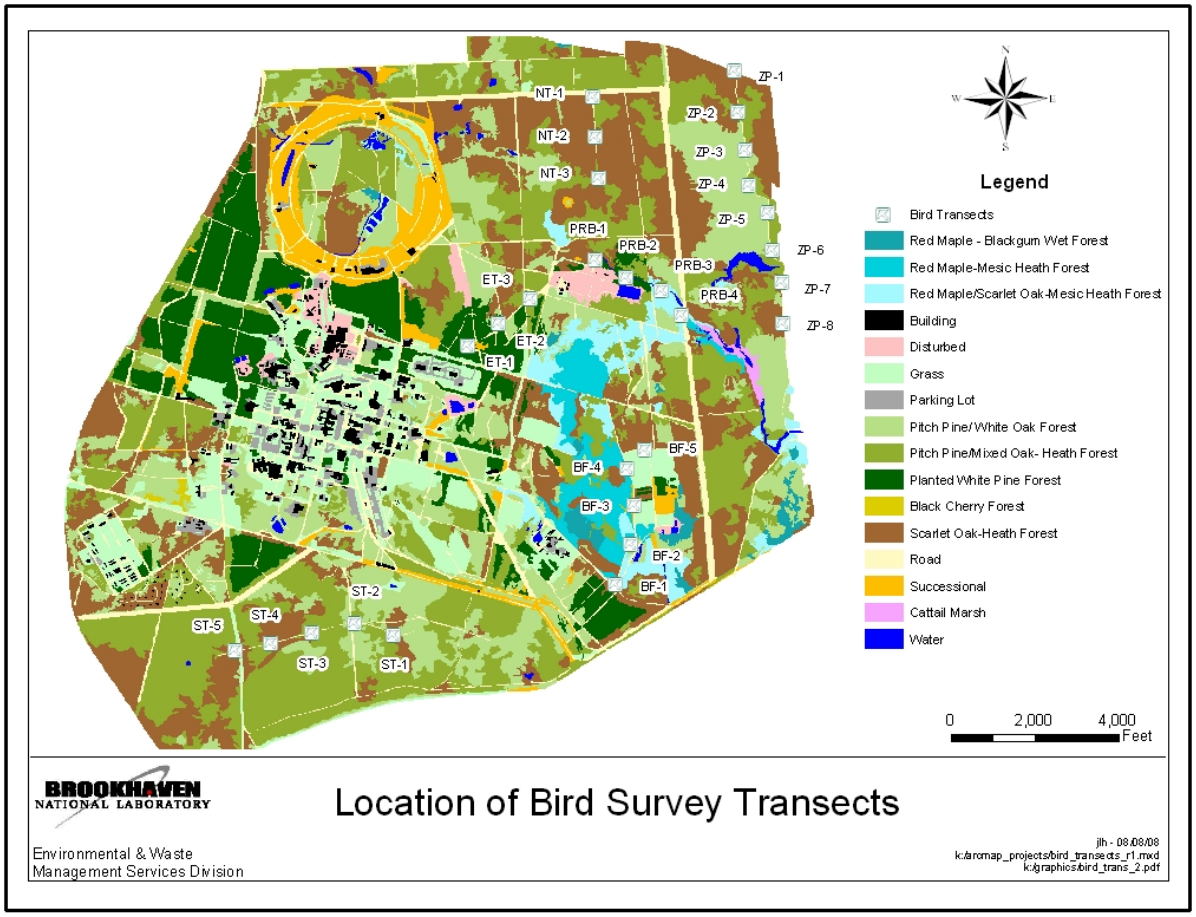

Figure 1: This map outlines the distribution of various types of vegetation within the Brookhaven National Laboratory site; of particular significance to this project are the types of vegetation within the bird transect boundaries.

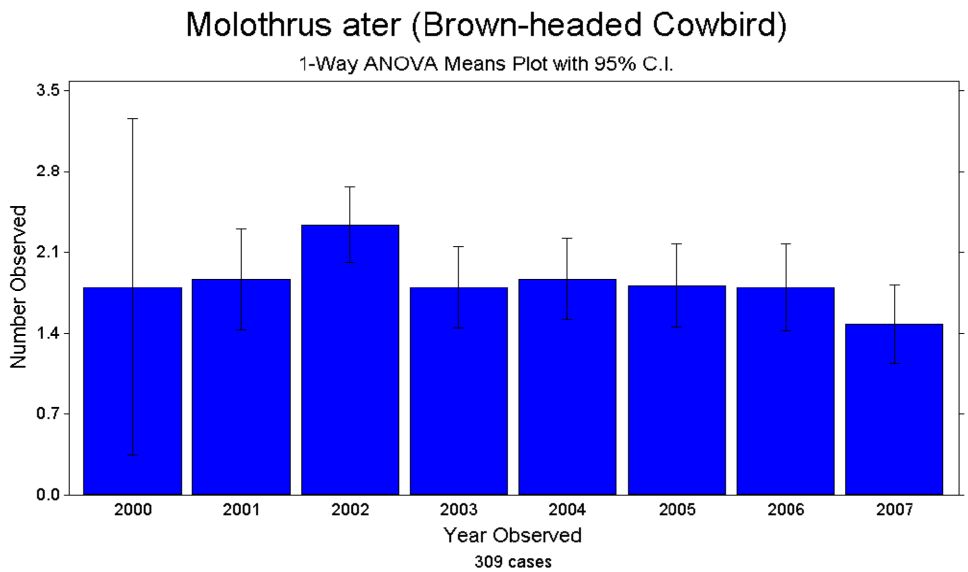

Figure 2: This plot gives a visual representation of the significant decrease in the number of brown-headed cowbirds per observation from 2002 to 2007 at $\alpha = 0.05$. Since this species is parasitic, a decrease in population is desired.

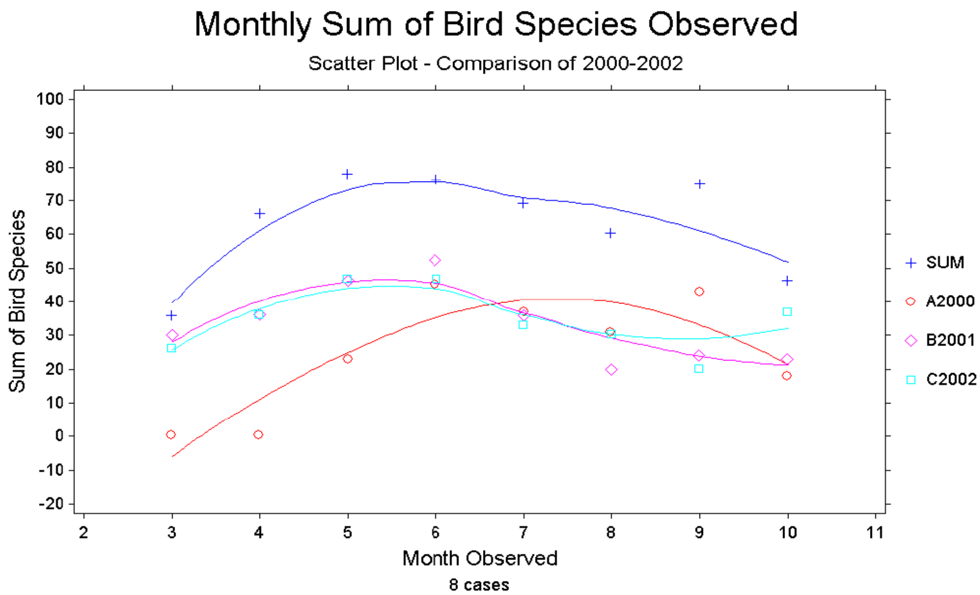

Figure 3: This plot compares the monthly sum of bird species for 2000 - 2002, as well as the monthly sum for the eight-year period. The first three years are somewhat dissimilar to the subsequent ones, due to the irregularity of surveys conducted in March and October.

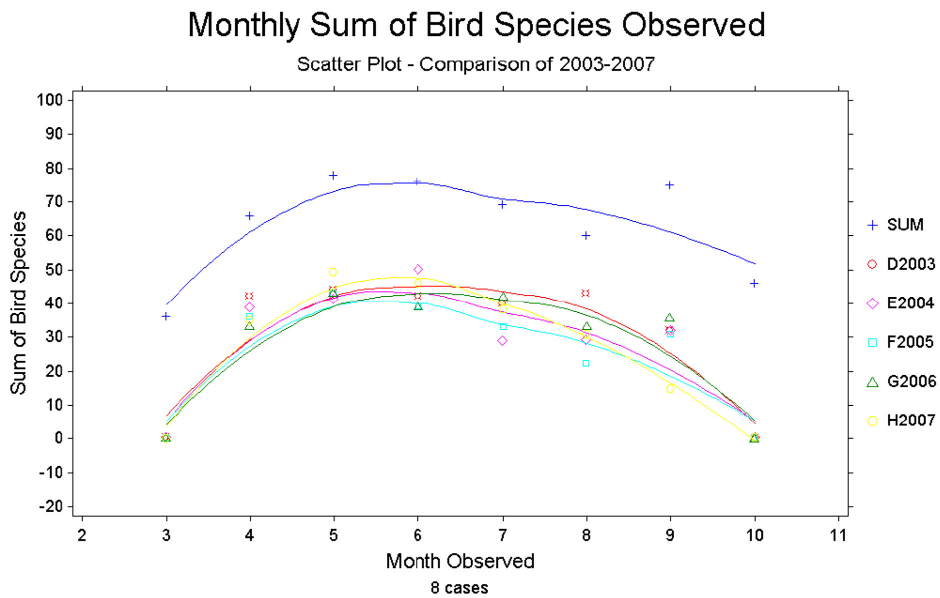

Figure 4: This plot compares the monthly sum of bird species for 2003 - 2007, as well as the monthly sum for the eight-year period. These years are quite similar to each other and to the eight-year sum, as expected.

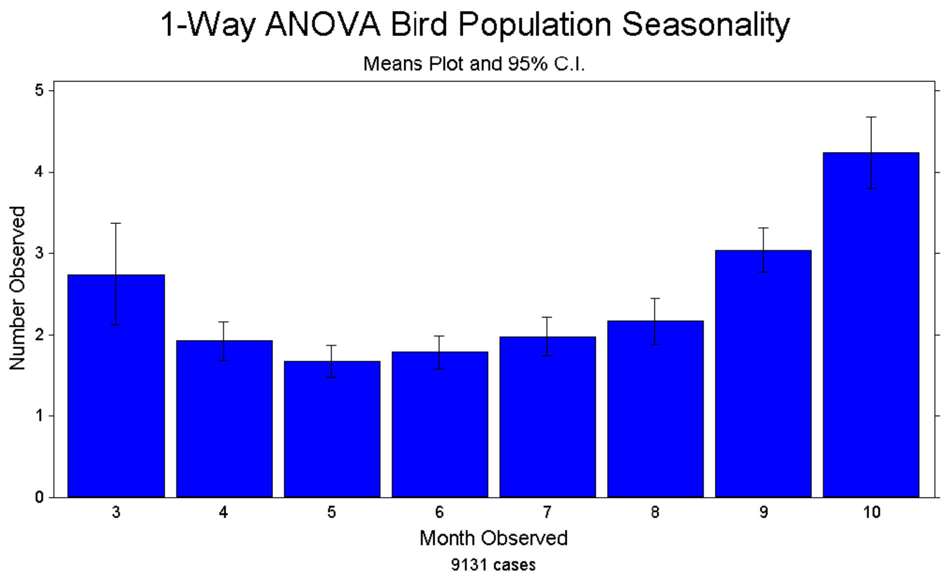

Figure 5: This plot compares the total mean number of birds per observation by month. Since more birds are seen during migratory periods, and less during nesting periods, this plot displays an expected trend.

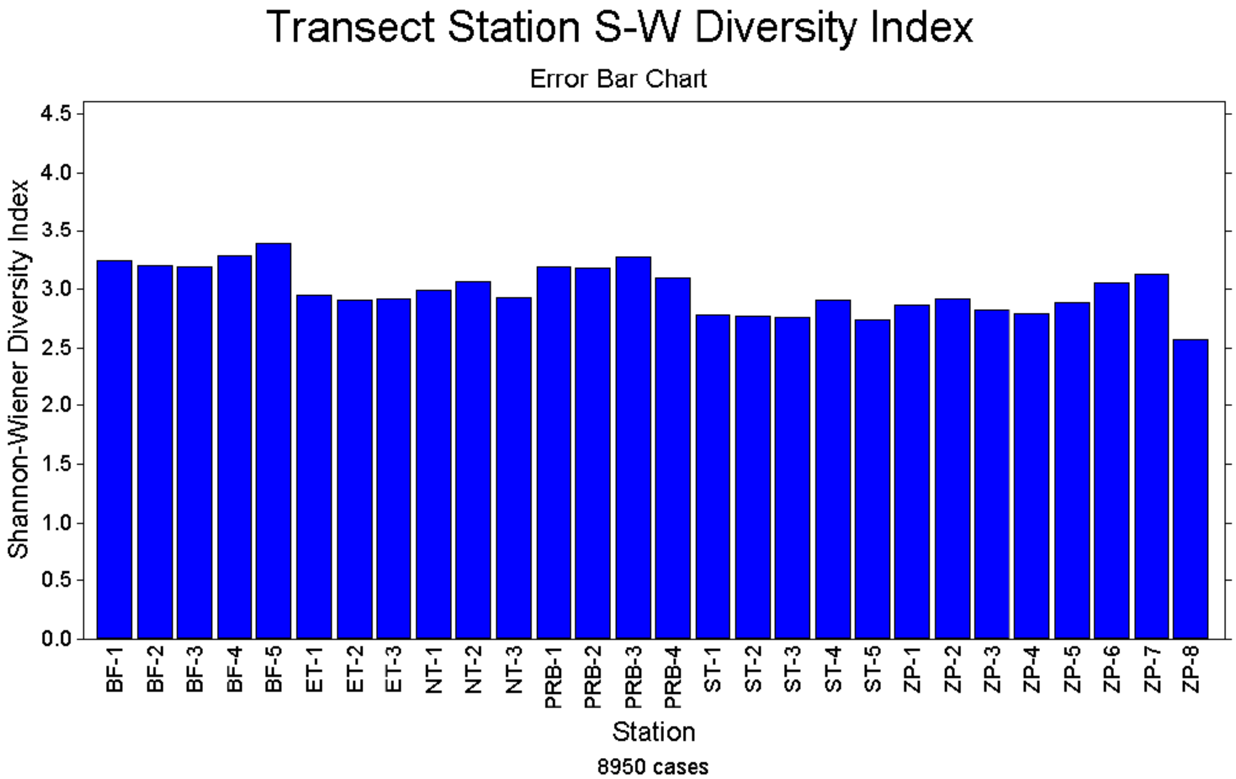

Figure 6: This chart gives a visual representation of the data. The basis for the hypothesized trend noted under Table 1 can be clearly seen.

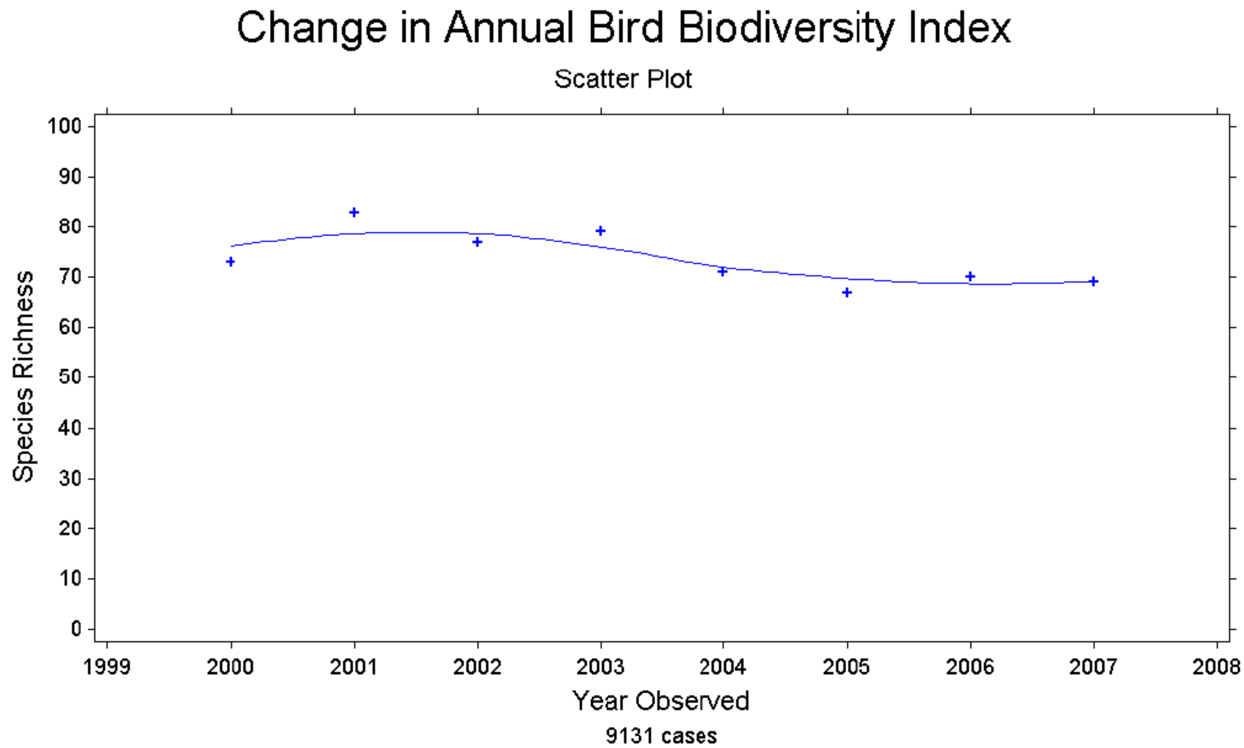

Figure 7: A total of 115 species were observed during the eight-year period. Since a number of those species were observed less than eight times (or less than once a year) each year was moderately species rich.

CHARTS

1. *Molothrus ater* (Brown-headed cowbird)

Tukey HSD All-Pairwise Comparisons Test of NUMBER by YEAR

YEAR	Mean	Homogeneous Groups
2002	2.3333	A
2004	1.8667	AB
2005	1.8140	AB
2000	1.8000	AB
2003	1.8000	AB
2006	1.8000	AB
2001	1.7000	AB
2007	1.4792	B

Alpha - 0.05 Standard Error for Comparison - 0.2299 TO 0.5539
Critical Q Value - 4.285
Critical Value for Comparison - 0.6965 TO 1.6783

There are 2 groups (A and B) in which the means
are not significantly different from one another.

2. *Seiurus aurocapillus* (Ovenbird)

Tukey HSD All-Pairwise Comparisons Test of NUMBER by YEAR

YEAR	Mean	Homogeneous Groups
2001	2.0571	A
2006	1.8600	A
2002	1.7551	AB
2000	1.7273	AB
2004	1.6250	AB
2003	1.6111	AB
2005	1.6000	AB
2007	1.3191	B

Alpha - 0.05 Standard Error for Comparison - 0.1560 TO 0.2683
Critical Q Value - 4.285
Critical Value for Comparison - 0.4727 TO 0.8129

There are 2 groups (A and B) in which the means
are not significantly different from one another.

Chart 1: This chart outlines the ANOVA test results for the two species that had significant changes in their populations – *Molothrus ater* and *Seiurus aurocapillus*. The *Molothrus ater* species had a decrease from 2002 to 2007, while the *Seiurus aurocapillus* species had an increase from 2000 to 2001, a decrease from 2001 to 2002, an increase from 2005 to 2006, and a decrease from 2006 to 2007. All population changes are at the 0.05 level of significance.

```
Tukey HSD All-Pairwise Comparisons Test of NUMBER by SEASON

SEASON    Mean  Homogeneous Groups
    10  4.2376  A
     9  3.0396   B
     3  2.7462   BC
     8  2.1623    CD
     7  1.9745    CD
     4  1.9212    CD
     6  1.7845    CD
     5  1.6752     D

Alpha - 0.05  Standard Error for Comparison - 0.1441 TO 0.3880
Critical Q Value - 4.285
Critical Value for Comparison - 0.4366 TO 1.1757

There are 4 groups (A, B, etc.) in which the means
are not significantly different from one another.
```

Chart 2: This chart outlines the results of the ANOVA test performed on the monthly sum of birds observed during the eight-year period. The analysis shows that the sum of birds is smallest in May and increases as both the spring and fall migration months are approached.

```
Tukey HSD All-Pairwise Comparisons Test of SRICHNESS by TRANSECT

TRANSECT      Mean  Homogeneous Groups
Biology Fi  60.000  A
Peconic Ri  52.250  AB
North Tran  42.667   BC
East Trenc  41.333   BC
Z-Path      38.250    C
South Tran  37.400    C

Alpha - 0.05    Standard Error for Comparison - 2.9642 TO 4.2454
Critical Q Value - 4.406
Critical Value for Comparison - 9.2340 TO 13.225

There are 3 groups (A, B, etc.) in which the means
are not significantly different from one another.


Tukey HSD All-Pairwise Comparisons Test of SNWRINDEX by TRANSECT

TRANSECT      Mean  Homogeneous Groups
Biology Fi  3.2588  A
Peconic Ri  3.1837  AB
North Tran  2.9912   BC
East Trenc  2.9223   BC
Z-Path      2.8784    C
South Tran  2.7879    C

Alpha - 0.05    Standard Error for Comparison - 0.0638 TO 0.0914
Critical Q Value - 4.406
Critical Value for Comparison - 0.1987 TO 0.2846

There are 3 groups (A, B, etc.) in which the means
are not significantly different from one another.
```

Chart 3: This chart contains the results of the ANOVA test performed on the transect station species richness and Shannon-Wiener diversity index. As hypothesized, the diversity indices of the Biology Fields and the Peconic River transects are significantly greater than those of the South and Z-Path transects.

‘

TABLES

Index / Transect Station	Species Richness	Simpson's Index	Simpson's Diversity Index	Simpson's Reciprocal Index	Shannon-Wiener Index	Simpson's Evenness Index
BF-1	53	0.0557487	0.9442513	17.937632	3.2429636	0.3384459
BF-2	51	0.0574765	0.9425235	17.398427	3.196578	0.3411456
BF-3	69	0.0810073	0.9189927	12.344571	3.189136	0.1789068
BF-4	63	0.0503331	0.9496669	19.867625	3.2820005	0.3153591
BF-5	64	0.0506679	0.9493321	19.736363	3.3832621	0.3083807
ET-1	46	0.0736543	0.9263457	13.576934	2.9516046	0.2951507
ET-2	36	0.0698256	0.9301744	14.321386	2.9006378	0.3978163
ET-3	42	0.0708125	0.9291875	14.121797	2.9146382	0.3362333
NT-1	45	0.0747323	0.9252677	13.381088	2.9892803	0.2973575
NT-2	42	0.0639532	0.9360468	15.63643	3.0608635	0.372296
NT-3	41	0.0857899	0.9142101	11.656387	2.9234245	0.2843021
PR-1	54	0.0643468	0.9356532	15.540781	3.189181	0.2877923
PR-2	45	0.0582698	0.9417302	17.161558	3.1777928	0.381368
PR-3	57	0.0564814	0.9435186	17.704933	3.2769355	0.3106129
PR-4	53	0.0774905	0.9225095	12.904815	3.09072	0.2434871
ST-1	38	0.0931188	0.9068812	10.738966	2.7789834	0.2826044
ST-2	37	0.0896579	0.9103421	11.153502	2.766118	0.301446
ST-3	37	0.0936375	0.9063625	10.679484	2.7547488	0.2886347

ST-4	39	0.0771449	0.9228551	12.962611	2.905493	0.3323747
ST-5	36	0.0882752	0.9117248	11.328214	2.7341617	0.3146726
ZP-1	38	0.0808511	0.9191489	12.368415	2.863346	0.3254846
ZP-2	43	0.0766422	0.9233578	13.047635	2.9185126	0.3034334
ZP-3	32	0.0770107	0.9229893	12.9852	2.8178248	0.4057875
ZP-4	30	0.0770619	0.9229381	12.976586	2.791071	0.4325529
ZP-5	37	0.0709856	0.9290144	14.087363	2.8840792	0.3807396
ZP-6	43	0.0671868	0.9328132	14.883886	3.0551022	0.3461369
ZP-7	46	0.0578739	0.9421261	17.278952	3.1293452	0.3756294
ZP-8	37	0.1390007	0.8609993	7.1942106	2.5678665	0.1944381
Diversity / Evenness	-	0 = most diverse; 1 = least diverse	1 = most diverse; 0 = least diverse	Positive infinity = most diverse; 0 = least diverse	4.6 = most diverse; 0 = least diverse	0 = most even; 1 = least even

Table 1 – From a brief perusal of this data, a trend in both the diversity and the evenness of a transect station can be hypothesized. The stations situated in the Biology Fields and the Peconic River transects appear to be the most diverse, while the South and Z-Path transects appear to be the least diverse.

Index / Year	Species Richness	Simpson's Index	Simpson's Diversity Index	Simpson's Reciprocal Index	Shannon-Wiener Index	Simpson's Evenness Index
2000	73	0.0525074	0.9474926	19.044947	3.376841	0.2608897
2001	83	0.0506175	0.9493825	19.755994	3.36946784	0.238024
2002	77	0.0756099	0.9243901	13.225785	3.19419197	0.1717634
2003	79	0.0573658	0.9426342	17.432003	3.30203596	0.2206583
2004	71	0.0558043	0.9441957	17.919782	3.25075604	0.2523913
2005	67	0.0574202	0.9425798	17.415472	3.27311523	0.2599324
2006	70	0.0745622	0.9254378	13.411621	3.11647739	0.1915946
2007	69	0.0520216	0.9479784	19.222782	3.33274301	0.278591
Diversity / Evenness	-	0 = most diverse; 1 = least diverse	1 = most diverse; 0 = least diverse	Positive infinity = most diverse; 0 = least diverse	4.6 = most diverse; 0 = least diverse	0 = most even; 1 = least even

Table 2 – All six biodiversity indices reveal a diverse bird population. The species richness is moderately high (see Fig. 5), as is the Shannon-Wiener diversity index. The Simpson diversity indices point to a high diversity, which is balanced by moderate evenness.