**Analysis of species accumulation curves, species richness estimators, and suggested changes in grid size for bird community sampling.**

Data analyzed are from five parks utilizing a 100 x 100 m grid:

GWCA (2016),

HEHO (2009),

HOME (2009),

LIBO (2015),

and PIPE (2009).

Notes: Non-breeding species (i.e., transients, winter residents, and migrants) were excluded from analyses. All other records associated with a plot were retained.

Species richness estimators included the Chao 2 and Jackknife 2, which have been used extensively in the literature, and the newer ICE and ACE, which were determined to perform well in a recent review (Reese et al. 2014).

The following questions were addressed:

**I. What do the species accumulation curves look like?**

**II. How do various species richness estimators compare?**

**III. Is there an effect of habitat type?**

**IV. What software can be used to calculate species richness estimators?**

**V. How does the frequency of sites at which a species is observed vary by species?**

**VI. What would be the expected effect of reducing sampling effort on observed species richness?**

**VII. How often are birds recorded from multiple plots?**

**VIII. What is the distribution of distances over which birds are observed?**

**I. What do the species accumulation curves look like?**

Pages 3-7 (analyses conducted in PC-ORD):

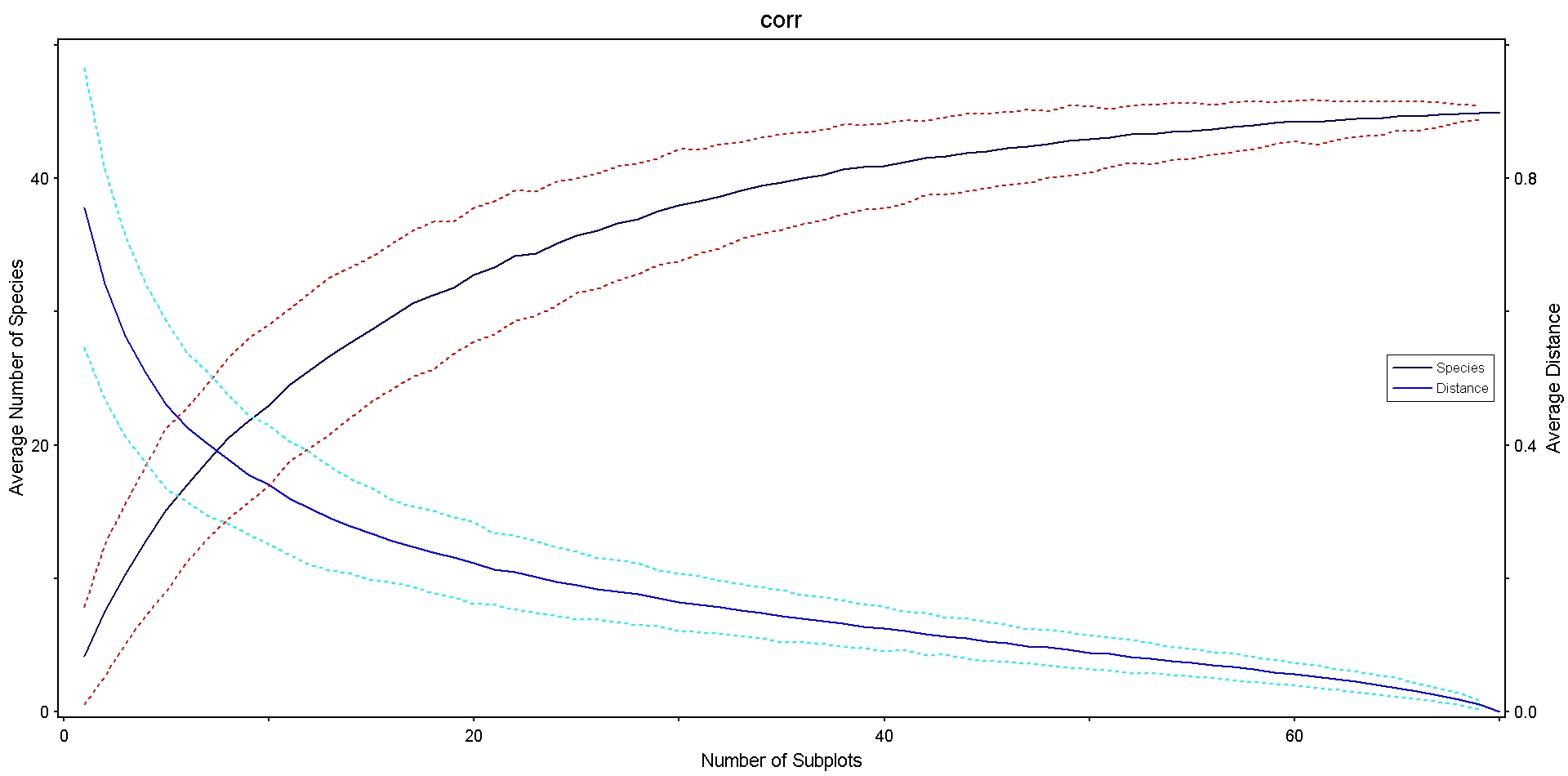
*Output consists of species accumulation curves, distance curves, Jackknife and Chao-2 estimators for total species richness and associated statistics.*

Species accumulation curves plot the cumulative number of species observed (on the Y axis) as a function of increasing number of plots (on the X axis). As the number of samples increases, the curve will begin to flatten until, at saturation, the slope becomes zero and no new species are found with the addition of further subsamples. The procedure takes repeated random samples of increasing size from the species presence-absence matrix and presents the average number of species occurring in a sample of a given size, along with error bars approximating the 95th percentile (based on two standard deviations from the mean).

Simultaneously, a distance matrix is calculated and the average distance among sample units is also calculated for each given sample size and plotted as a curve. When all sample units are included in the subsample, the positions of the centroids of the entire matrix and of the subsample are the same and hence their distance is zero.

The jackknife estimators and the Chao-2 estimator are nonparametric resampling procedures that adjust the richness estimate based on different ratios of the number of species that occur only once or twice in the matrix relative to the sample size. These estimators may improve accuracy and reduce bias for relatively restricted, homogeneous areas with few rare species. If, in such datasets, the curve does not flatten out and the jackknife estimates are considerably higher than the observed species richness, this could reflect either insufficient sampling or the presence of numerous rare species.

**GWCA (2016)**



**Species accumulation curve for 45 species at 70 sites at GWCA in 2016.**

Estimates of total number of species:

45.0000 = Number of species observed

49.9286 = First-order jackknife estimate

49.9990 = Second-order jackknife estimate

47.5000 = Chao2 estimate, classic form

46.6429 = Chao2 estimate, bias corrected form

11.2500 = Variance of Chao2 estimate, classic form

4.7504 = Variance of Chao2 estimate, bias corrected

Additional information used for jackknife and Chao estimates:

5 = Number of species with only 1 occurrence

5 = Number of species with only 2 occurrences

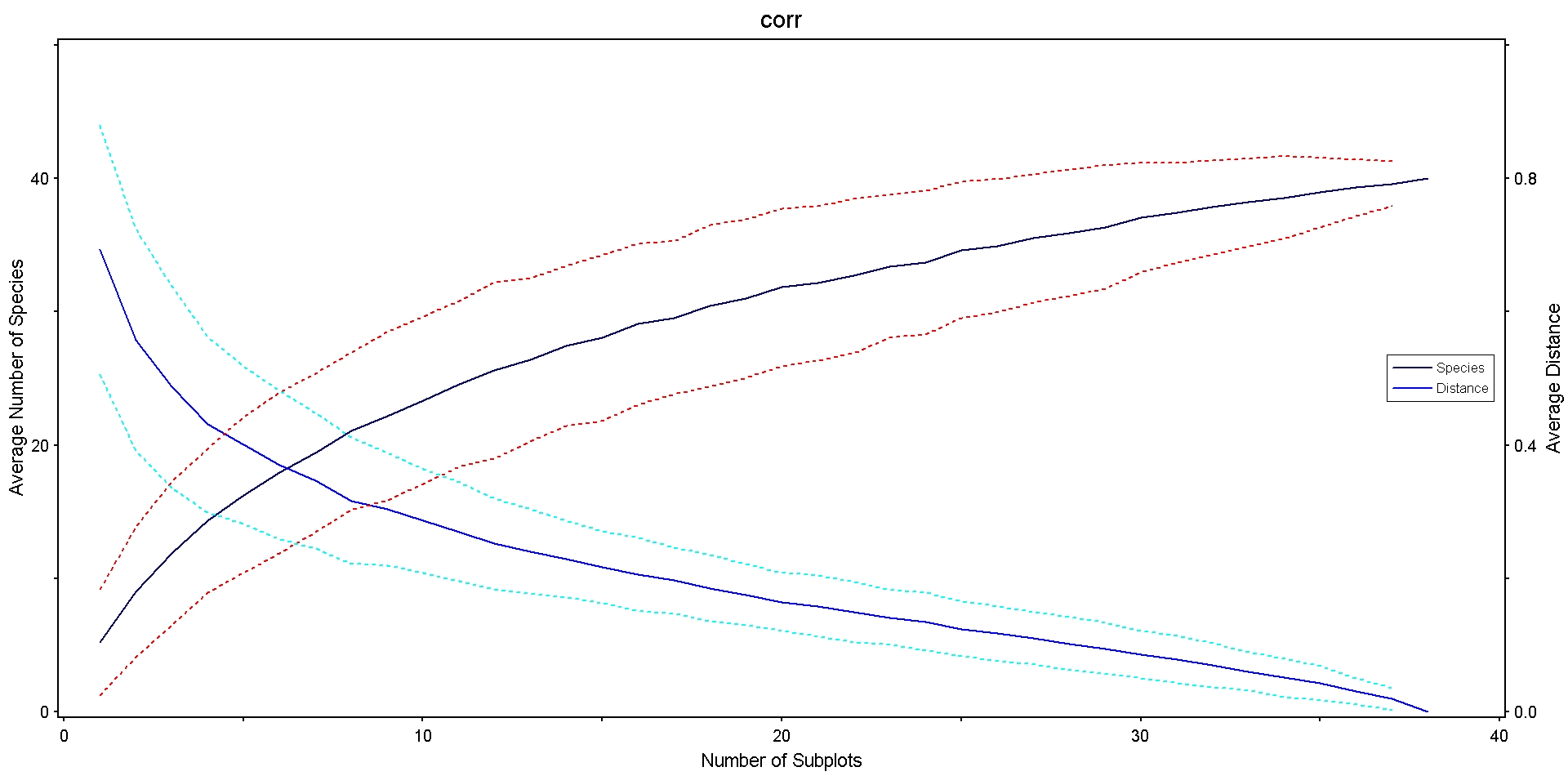
So: 22% of observed species found at only 1 or 2 sites.

Cumulative number of bird species recorded during breeding bird surveys at GWCA from 2008 through 2016: 91

Percentage of the cumulative number of bird species recorded during the above multiple year interval that the single year Chao2 estimator represents: 51%

Percentage of the cumulative number of bird species recorded during the above multiple year interval that the actual number observed represents: 49%

**HEHO (2009)**

****

**Species accumulation curve for 40 species at 38 sites at HEHO in 2009.**

Estimates of total number of species:

40.0000 = Number of species observed

52.6579 = First-order jackknife estimate

57.5996 = Second-order jackknife estimate

50.5625 = Chao2 estimate, classic form

48.4386 = Chao2 estimate, bias corrected form

69.3989 = Variance of Chao2 estimate, classic form

39.2463 = Variance of Chao2 estimate, bias corrected

Additional information used for jackknife and Chao estimates:

13 = Number of species with only 1 occurrence

8 = Number of species with only 2 occurrences

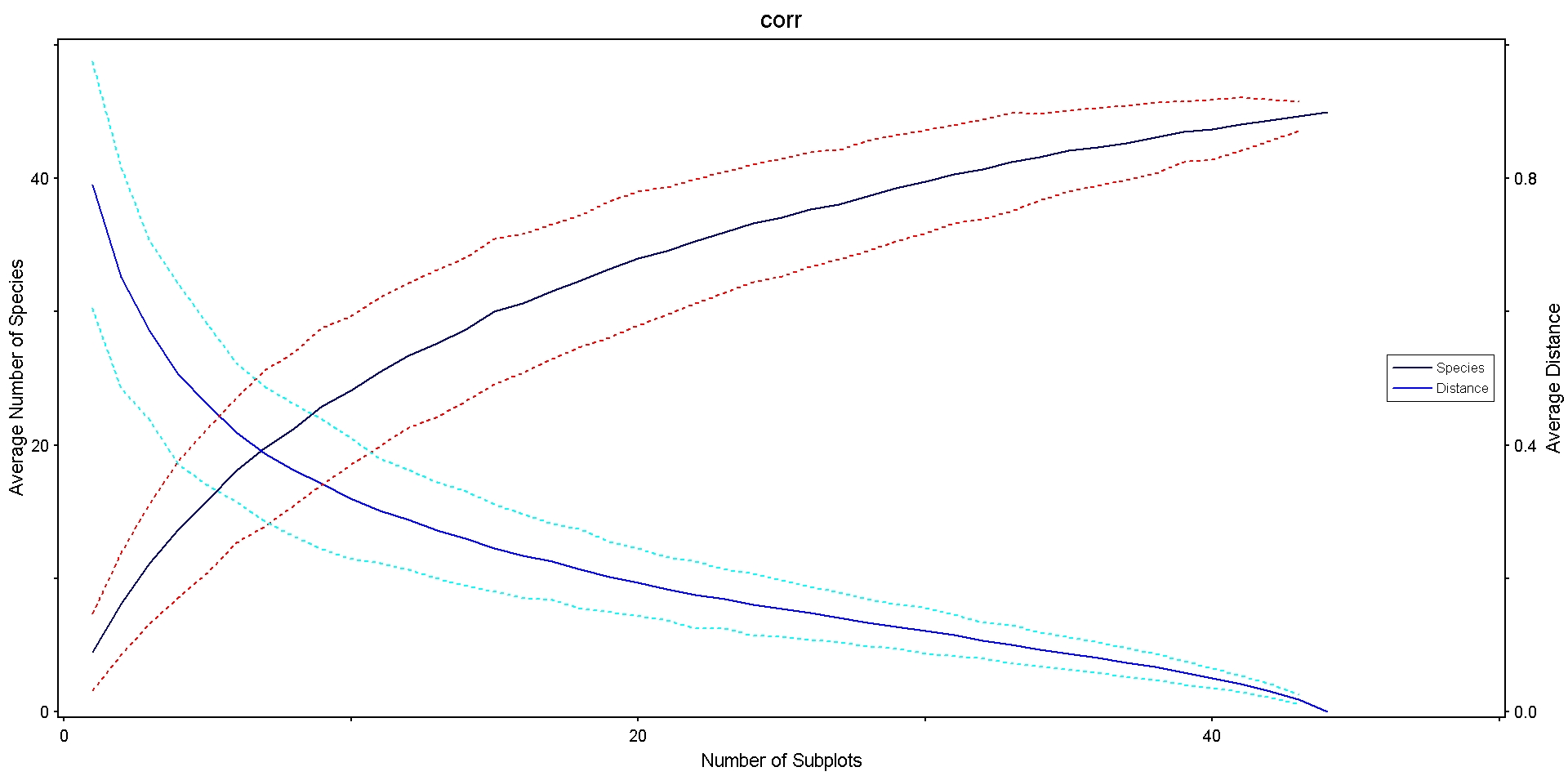
So: 53% of observed species found at only 1 or 2 sites.

Cumulative number of bird species recorded during breeding bird surveys at HEHO from 2005 through 2013: 61

Percentage of the cumulative number of bird species recorded during the above multiple year interval that the single year Chao2 estimator represents: 79%

Percentage of the cumulative number of bird species recorded during the above multiple year interval that the actual number observed represents: 66%

**HOME (2009)**



**Species accumulation curve for 45 species at 44 sites at HOME in 2009.**

Estimates of total number of species:

45.0000 = Number of species observed

57.7045 = First-order jackknife estimate

60.7902 = Second-order jackknife estimate

53.4500 = Chao2 estimate, classic form

51.9298 = Chao2 estimate, bias corrected form

46.0103 = Variance of Chao2 estimate, classic form

26.9313 = Variance of Chao2 estimate, bias corrected

Additional information used for jackknife and Chao estimates:

13 = Number of species with only 1 occurrence

10 = Number of species with only 2 occurrences

So: 51% of observed species found at only 1 or 2 sites.

Cumulative number of bird species recorded during breeding bird surveys at HOME from 2009 through 2013: 71

Percentage of the cumulative number of bird species recorded during the above multiple year interval that the single year Chao2 estimator represents: 73%

Percentage of the cumulative number of bird species recorded during the above multiple year interval that the actual number observed represents: 63%

**LIBO (2015)**

****

**Species accumulation curve for 28 species at 35 sites at LIBO in 2015.**

Estimates of total number of species:

28.0000 = Number of species observed

34.8000 = First-order jackknife estimate

36.8244 = Second-order jackknife estimate

32.9000 = Chao2 estimate, classic form

31.4000 = Chao2 estimate, bias corrected form

28.3220 = Variance of Chao2 estimate, classic form

12.7581 = Variance of Chao2 estimate, bias corrected

Additional information used for jackknife and Chao estimates:

7 = Number of species with only 1 occurrence

5 = Number of species with only 2 occurrences

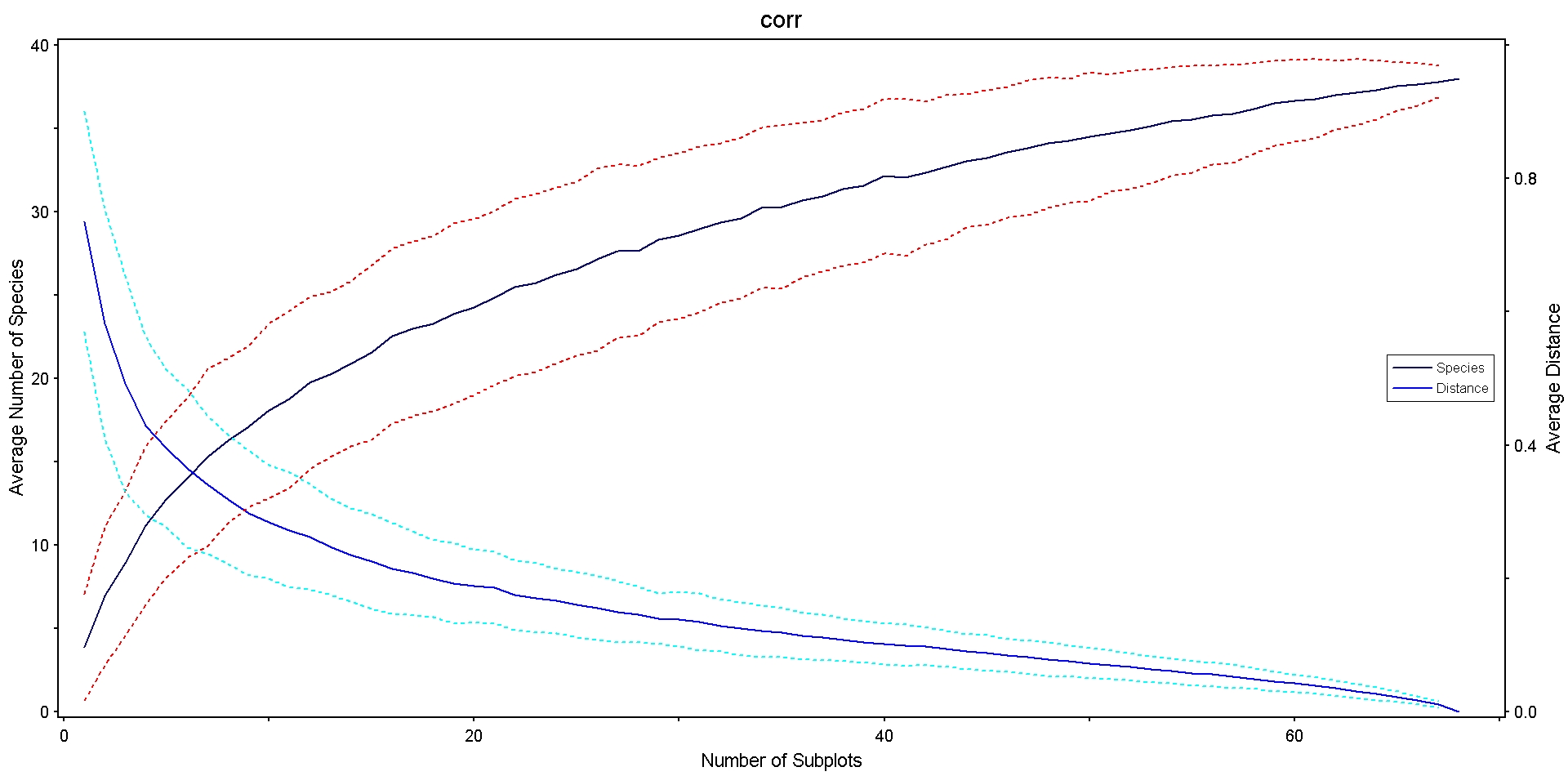
So: 43% of observed species found at only 1 or 2 sites.

Cumulative number of bird species recorded during breeding bird surveys at LIBO from 2007 through 2015: 50

Percentage of the cumulative number of bird species recorded during the above multiple year interval that the single year Chao2 estimator represents: 63%

Percentage of the cumulative number of bird species recorded during the above multiple year interval that the actual number observed represents: 56%

**PIPE (2009)**

****

**Species accumulation curve for 38 species at 68 sites at PIPE in 2009.**

Estimates of total number of species:

38.0000 = Number of species observed

48.8382 = First-order jackknife estimate

51.8659 = Second-order jackknife estimate

45.5625 = Chao2 estimate, classic form

44.0212 = Chao2 estimate, bias corrected form

43.0708 = Variance of Chao2 estimate, classic form

24.1371 = Variance of Chao2 estimate, bias corrected

Additional information used for jackknife and Chao estimates:

11 = Number of species with only 1 occurrence

8 = Number of species with only 2 occurrences

So: 50% of observed species found at only 1 or 2 sites.

Cumulative number of bird species recorded during breeding bird surveys at PIPE from 2009 through 2013: 71

Percentage of the cumulative number of bird species recorded during the above multiple year interval that the single year Chao2 estimator represents: 62%

Percentage of the cumulative number of bird species recorded during the above multiple year interval that the actual number observed represents: 54%

**II. How do various species richness estimators compare?**

Pages 9-13 (analyses conducted in Estimate S):

*Output consists of species accumulation curves, as well as Jackknife 2, Chao-2, ACE, and ICE estimators for each sample size.*

The Jackknife 2 and Chao-2 species richness estimators have been used frequently in numerous studies. The ACE and ICE species richness estimators are relatively new estimators, but have been shown to provide less biased and more accurate estimates than the more popular Jackknife 2 and Chao-2 estimators (Reese et al. 2014).



**Species accumulation curve for 45 species at 70 sites at GWCA in 2016, along with 4 species richness estimators: Jackknife 2, Chao 2, ICE and ACE.**



**Species accumulation curve for 40 species at 38 sites at HEHO in 2009, along with 4 species richness estimators: Jackknife 2, Chao 2, ICE and ACE.**

  
**Species accumulation curve for 45 species at 44 sites at HOME in 2009, along with 4 species richness estimators: Jackknife 2, Chao 2, ICE and ACE.**



**Species accumulation curve for 28 species at 35 sites at LIBO in 2015, along with 4 species richness estimators: Jackknife 2, Chao 2, ICE and ACE.**



**Species accumulation curve for 38 species at 68 sites at PIPE in 2009, along with 4 species richness estimators: Jackknife 2, Chao 2, ICE and ACE.**

**III. Is there an effect of habitat type?**

A. PIPE dataset:

To evaluate the potential effect of habitat type, 8 sites at PIPE which were in or near primarily wooded areas (#5, 9, 21, 27, 28, 33, 34, 40) were excluded and the analyses were re-run using only grassland sites.

Hypothesis: The relatively few wooded sites contain species that prefer wooded areas and are not found in the grasslands, and these species are rare in the park because of the small wooded area. Such truly rare species will inflate species richness estimators. Thus excluding wooded sites should result in species estimators that are closer to observed species richness.

Results:

Sample size decreased from 68 to 60.

Species richness decreased from 38 to 33.

Chao 2 estimate was 33% higher than observed species richness (compared to 16% higher for all sites).

ICE estimate was 42% higher than observed species richness (compared to 24% higher for all sites).

*Removing wooded sites at PIPE did not result in observed species richnesses that were closer to estimated species richnesses, but the opposite was true.*



**Species accumulation curve for 33 species at 60 sites at PIPE in 2009, along with 4 species richness estimators: Jackknife 2, Chao 2, ICE and ACE. Only grassland sites were included.**

Estimates of total number of species:

33.0000 = Number of species observed

43.82 = First-order jackknife estimate

50.65 = Second-order jackknife estimate

43.82 = Chao2 estimate, bias corrected form

11 = Number of species with only 1 occurrence

4 = Number of species with only 2 occurrences

So: 45% of observed species found at only 1 or 2 sites.

B. HOME dataset:

To evaluate the potential effect of habitat type at HOME, only sites which were in grassland areas were included (sites 1-30).

Hypothesis: By evaluating only grassland sites, only species that prefer grasslands would be included. By evaluating a potentially smaller species pool, this may result in species estimators that are closer to observed species richness.

Results:

Sample size decreased from 44 to 30.

Species richness decreased from 45 to 33.

Chao 2 estimate was 24% higher than observed species richness (compared to 16% higher for all sites).

ICE estimate was 61% higher than observed species richness (compared to 27% higher for all sites).

*Removing wooded sites at HOME did not result in observed species richnesses that were closer to estimated species richnesses, but the opposite was true.*



**Species accumulation curve for 33 species at 30 sites at HOME in 2009, along with 4 species richness estimators: Jackknife 2, Chao 2, ICE and ACE. Only grassland sites were included.**

Estimates of total number of species:

33.0000 = Number of species observed

45.57 = First-order jackknife estimate

49.59 = Second-order jackknife estimate

40.54 = Chao2 estimate, bias corrected form

13 = Number of species with only 1 occurrence

9 = Number of species with only 2 occurrences

So: 67% of observed species found at only 1 or 2 sites.

**IV. What software can be used to calculate species richness estimators?**

**PC-ORD** has a “Species-area Analysis” option that creates a smoothed species accumulation curve graph, and calculates Jackknife 1, Jackknife 2, and Chao 2species richness estimators based on the final sample size. PC-ORD is relatively expensive but several of us have this software already.

**EstimateS** is freely available and calculates a diversity of species richness estimators and other diversity statistics. Every estimator and statistic is calculated for the full sample size and all smaller sample sizes, making the graphs presented above possible. EstimateS does not have built-in graphing functions, but graphs can be easily prepared with a graphing program.

**V. How does the frequency of sites at which a species is observed vary by species?**

Pages 20-24:

*Output consists of species abundance curves (actually frequency rather than abundance).*

The number of sites at which each species was observed is graphed, with the species ordered by most to least common. Many species are rare, being observed at only a few sites.



**The number of sites at which each species was observed at GWCA in 2016. Species are ordered by relative abundance, from most to least common.**

Number of sites at which each of the species of regional conservation concern were observed: \*

BEVI – 3

CEWA – 0

HESP – 0

KEWA – 0

LOSH – 0

RHWO – 0

\* 0 indicates the species has been observed in this park during monitoring in some other year.



**The number of sites at which each species was observed at HEHO in 2009. Species are ordered by relative abundance, from most to least common.**

Number of sites at which each of the species of regional conservation concern were observed: \*

DICK – 6

FISP – 2

GRSP – 0

YSFL– 0

\* 0 indicates the species has been observed in this park during monitoring in some other year.



**The number of sites at which each species was observed at HOME in 2009. Species are ordered by relative abundance, from most to least common.**

Number of sites at which each of the species of regional conservation concern were observed: \*

BEVI – 0

RHWO – 7

\* 0 indicates the species has been observed in this park during monitoring in some other year.



**The number of sites at which each species was observed at LIBO in 2015. Species are ordered by relative abundance, from most to least common.**

Number of sites at which each of the species of regional conservation concern were observed: \*

KEWA – 4

RHWO – 0

WOTH – 5

\* 0 indicates the species has been observed in this park during monitoring in some other year.



**The number of sites at which each species was observed at PIPE in 2009. Species are ordered by relative abundance, from most to least common.**

Number of sites at which each of the species of regional conservation concern were observed: \*

BLTE – 0

DICK – 2

GRSP – 15

LEBI – 0

SWHA – 0

UPSA – 1

\* 0 indicates the species has been observed in this park during monitoring in some other year.

**VI. What would be the expected effect of reducing sampling effort on observed species richness?**

*From Dave (plots to be sampled if grid size is changed from a 100 x 100 to a 200 x 200 m spacing):*

Subset of plots sampled for birds on parks with plots on 100 x 100 m sampling grids. Plots listed by parks are on 200 x 200 m grids.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Plot I.D. | | X Coordinate (Easting) | Y Coordinate (Northing) |  |  |
| George Washington Carver National Monument, Missouri - GWCA | | | | | |
| GWCATweety2 | | 379211.842 | 4094759.924 |  |  |
| GWCATweety4 | | 379494.685 | 4094759.924 |  |  |
| GWCATweety11 | | 379070.421 | 4094618.503 |  |  |
| GWCATweety13 | | 379353.263 | 4094618.503 |  |  |
| GWCATweety15 | | 379636.106 | 4094618.503 |  |  |
| GWCATweety22 | | 379211.842 | 4094477.082 |  |  |
| GWCATweety24 | | 379494.685 | 4094477.082 |  |  |
| GWCATweety31 | | 379070.421 | 4094335.660 |  |  |
| GWCATweety33 | | 379353.263 | 4094335.660 |  |  |
| GWCATweety35 | | 379636.106 | 4094335.660 |  |  |
| GWCATweety42 | | 379211.842 | 4094194.239 |  |  |
| GWCATweety46 | | 379070.421 | 4094052.817 |  |  |
| GWCATweety47 | | 379353.263 | 4094052.817 |  |  |
| GWCATweety48 | | 379636.106 | 4094052.817 |  |  |
| GWCATweety54 | | 379211.842 | 4093911.396 |  |  |
| GWCATweety56 | | 379494.685 | 4093911.396 |  |  |
| GWCATweety62 | | 379070.421 | 4093769.975 |  |  |
| GWCATweety64 | | 379353.263 | 4093769.975 |  |  |
| GWCATweety66 | | 379636.106 | 4093769.975 |  |  |
|  |  | |  |  |  |
| Herbert Hoover National Historic Site, Iowa - HEHO | | | | | |
| HEHOTweety 01 | | 636971.393 | 4614410.700 |  |  |
| HEHOTweety 02 | | 637005.863 | 4614129.966 |  |  |
| HEHOTweety 03 | | 636882.731 | 4613972.364 |  |  |
| HEHOTweety 04 | | 637286.597 | 4614164.436 |  |  |
| HEHOTweety 05 | | 637163.465 | 4614006.834 |  |  |
| HEHOTweety 06 | | 637040.333 | 4613849.232 |  |  |
| HEHOTweety 07 | | 637567.332 | 4614198.906 |  |  |
| HEHOTweety 08 | | 637444.200 | 4614041.303 |  |  |
| HEHOTweety 09 | | 637321.067 | 4613883.701 |  |  |
| HEHOTweety 31 | | 637128.995 | 4614287.568 |  |  |
| HEHOTweety 35 | | 637409.730 | 4614322.038 |  |  |
|  | |  |  |  |  |
| Homestead National Monument of America, Nebraska - HOME | | | | | |
| HOMETweety 03 | | 684124.165 | 4462154.122 |  |  |
| HOMETweety 11 | | 684124.165 | 4461871.279 |  |  |
| HOMETweety 13 | | 684407.008 | 4461871.279 |  |  |
| HOMETweety 21 | | 683982.744 | 4461729.858 |  |  |
| HOMETweety 23 | | 684265.587 | 4461729.858 |  |  |
| HOMETweety 25 | | 684548.429 | 4461729.858 |  |  |
| HOMETweety 37 | | 683912.033 | 4461941.990 |  |  |
| HOMETweety 39 | | 683558.480 | 4461871.279 |  |  |
| HOMETweety 47 | | 683699.901 | 4461729.858 |  |  |
|  | |  |  |  |  |
| Lincoln Boyhood National Memorial, Indiana - LIBO | | | | | |
| LIBOTweety 03 | | 500432.278 | 4219553.960 |  |  |
| LIBOTweety 08 | | 500432.278 | 4219271.117 |  |  |
| LIBOTweety 14 | | 500432.278 | 4218988.274 |  |  |
| LIBOTweety 19 | | 500290.856 | 4218846.853 |  |  |
| LIBOTweety 26 | | 500149.435 | 4218705.432 |  |  |
| LIBOTweety 28 | | 500432.278 | 4218705.432 |  |  |
| LIBOTweety 32 | | 500008.014 | 4218564.010 |  |  |
| LIBOTweety 34 | | 500290.856 | 4218564.010 |  |  |
|  | |  |  |  |  |
| Pipestone National Monument, Minnesota - PIPE | | | | | |
| PIPETweety 03 | | 714311.551 | 4877308.484 |  |  |
| PIPETweety 05 | | 714594.394 | 4877308.484 |  |  |
| PIPETweety 11 | | 714170.130 | 4877167.063 |  |  |
| PIPETweety 13 | | 714452.973 | 4877167.063 |  |  |
| PIPETweety 15 | | 714735.815 | 4877167.063 |  |  |
| PIPETweety 23 | | 714028.708 | 4877025.642 |  |  |
| PIPETweety 25 | | 714311.551 | 4877025.642 |  |  |
| PIPETweety 27 | | 714594.394 | 4877025.642 |  |  |
| PIPETweety 29 | | 714877.237 | 4877025.642 |  |  |
| PIPETweety 37 | | 714170.130 | 4876884.220 |  |  |
| PIPETweety 40 | | 714735.815 | 4876884.220 |  |  |
| PIPETweety 47 | | 714028.708 | 4876742.799 |  |  |
| PIPETweety 50 | | 714877.237 | 4876742.799 |  |  |
| PIPETweety 56 | | 714170.130 | 4876601.378 |  |  |
| PIPETweety 58 | | 714452.973 | 4876601.378 |  |  |
| PIPETweety 64 | | 714028.708 | 4876459.956 |  |  |
| PIPETweety 66 | | 714877.237 | 4876459.956 |  |  |

**If grid dimensions are increased from 100 x100 m to 200 x 200 m, then:**

* At GWCA: Number of sample plots decreases from 70 to 19 (reduction by 73%).
* At HEHO: Number of sample plots decreases from 38 to 11 (reduction by 71%).
* At HOME: Number of sample plots decreases from 48 to 9 (reduction by 81%).
* (I think this wasn’t done correctly; a few too many sites were deleted.)
* At LIBO: Number of sample plots decreases from 35 to 8 (reduction by 77%).
* At PIPE: Number of sample plots decreases from 68 to 17 (reduction by 75%).

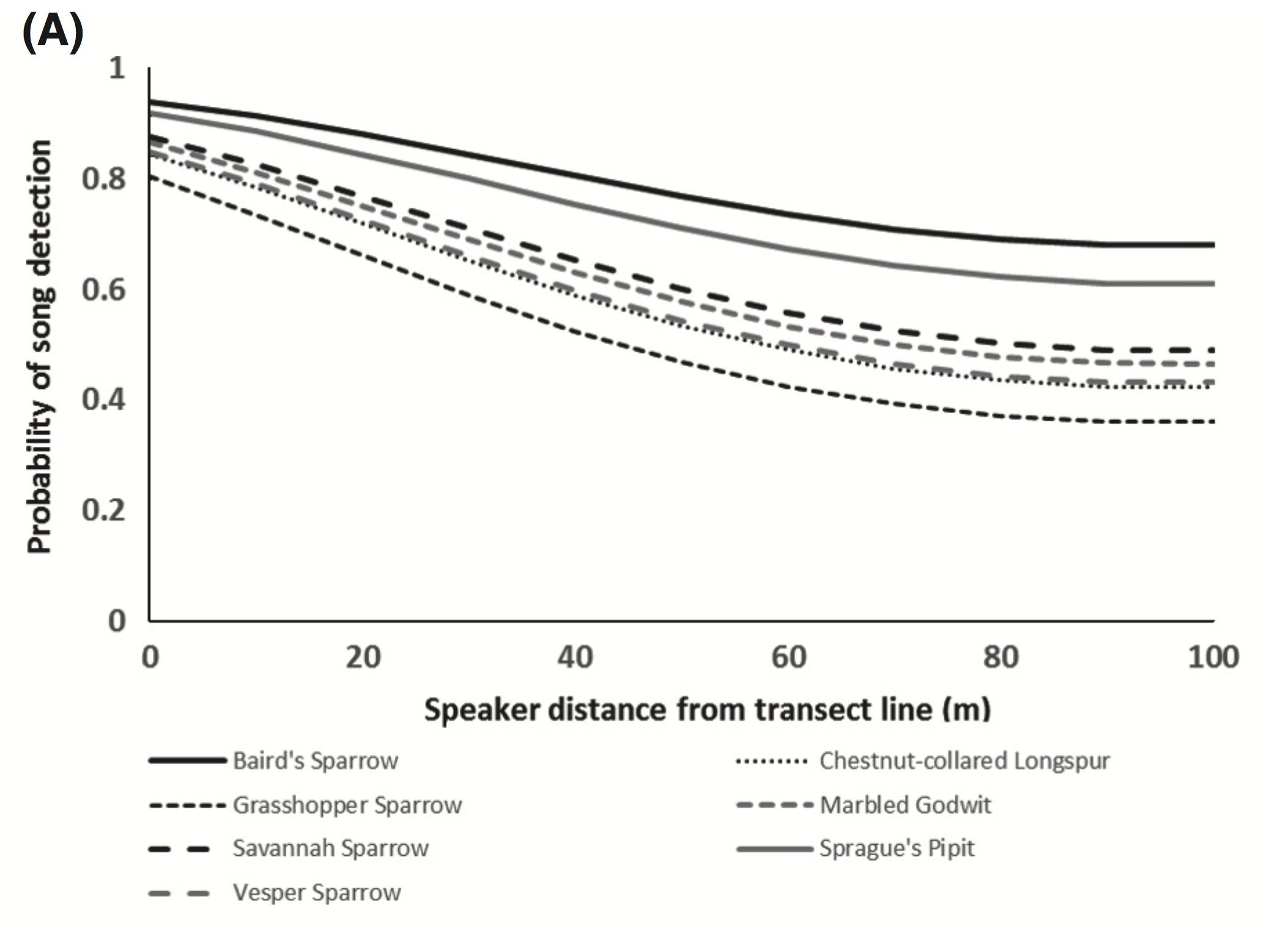
**And, based on the most recent bird surveys:**

* At GWCA: Number of observed species would decrease from 45 to 32 (estimated) (reduction by 29%).
* At HEHO: Number of observed species would decrease from 40 to 24 (estimated) (reduction by 40%).
* At HOME: Number of observed species would decrease from 45 to 23 (estimated) (reduction by 27%).
* At LIBO: Number of observed species would decrease from 28 to 13 (estimated) (reduction by 54%).
* At PIPE: Number of observed species would decrease from 38 to 23 (estimated) (reduction by 39%).

*Overall, sampled plot number decreases by 71-77%, while number of observed species decreases by 27-54%.*

Note: The above assumes that all species observed at all plots were recorded in the database. If species were not recorded because they were considered to be previously observed from another plot, then it is likely that additional records would be obtained from these plots using a larger grid size, and the decrease in number of observed species be less than that reported here.

In a recent study of grassland birds (Koper et al. 2016), detection probabilities of birds averaged about ~ 60% at 50 m and ~50% at 100 m. Thus it is likely that some, but not all individuals currently observed would still be observed at the greater plot spacing. This would vary by species (see graph below).



**Probability of song detection in southern Alberta grasslands, 2013–2014. Detection probability varied with both species and speaker distance.**

From Koper et al. 2016. This study took place in native mixed-grass prairie sites in south-central Alberta, Canada (approximately 50° 53° 58.091′′ N, 112° 26° 35.456′′ W).

**VII. How often are birds recorded from multiple plots?**

There are two ways of accessing how frequently the same birds may be recorded from more than one plot:

*A. Based on identification of having been previously recorded:*

In the dataset, there is a column that indicates whether birds were previously observed at another plot. The following are the number of times birds were recorded as being already observed at another plot (the “IsPreviousPlot” column in the dataset is checked):

GWCA (2016): 41 of 383 records (11%)

HEHO (2009): 8 of 262 records (3%)

HOME (2009): 0 of 258 records (0%)

LIBO (2015): 3 of 135 records (2%)

PIPE (2009): 7 of 348 records (2%)

If flyovers are excluded, the results are:

GWCA (2016): 41 of 344 records (12%)

HEHO (2009): 7 of 182 records (4%)

HOME (2009): 0 of 218 records (0%)

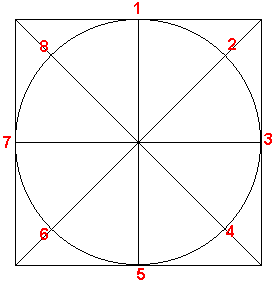
LIBO (2015): 3 of 130 records (2%)

PIPE (2009): 7 of 234 records (3%)

*B. Based on recorded distances:*

Alternatively, because of the 100 x 100 m grid spacing, a circle encompassing the maximum area of each sampling point (so that the edges of the circle would touch the edges of neighboring circles) would have a radius of 50 m. If squares of 100 m sides were superimposed on top of the circles, so that the sides of all squares were immediately adjacent to all neighboring squares, the area of the square would represent the overall area closest to each sampling point.

In the figure below, the line labeled “1” would have a radius of 50 m. The line labeled “2” would be 141.4 m long, or 70.7 m from the center to the corner. Thus any birds observed at distances >50 m may be closer to the center of other sampling points, and all birds observed at distances >70 m must be closer to the center of other sampling points, or outside (or near) park boundaries.



The following are the number of observations of birds at distances greater than 50 and 70 m, respectively. Flyovers were not included in these calculations:

GWCA (2016): 155 of 344 records > 50 m (45%); 97 of 344 records > 70 m (28%)

HEHO (2009): 90 of 182 records > 50 m (49%); 63 of 182 records > 70 m (35%)

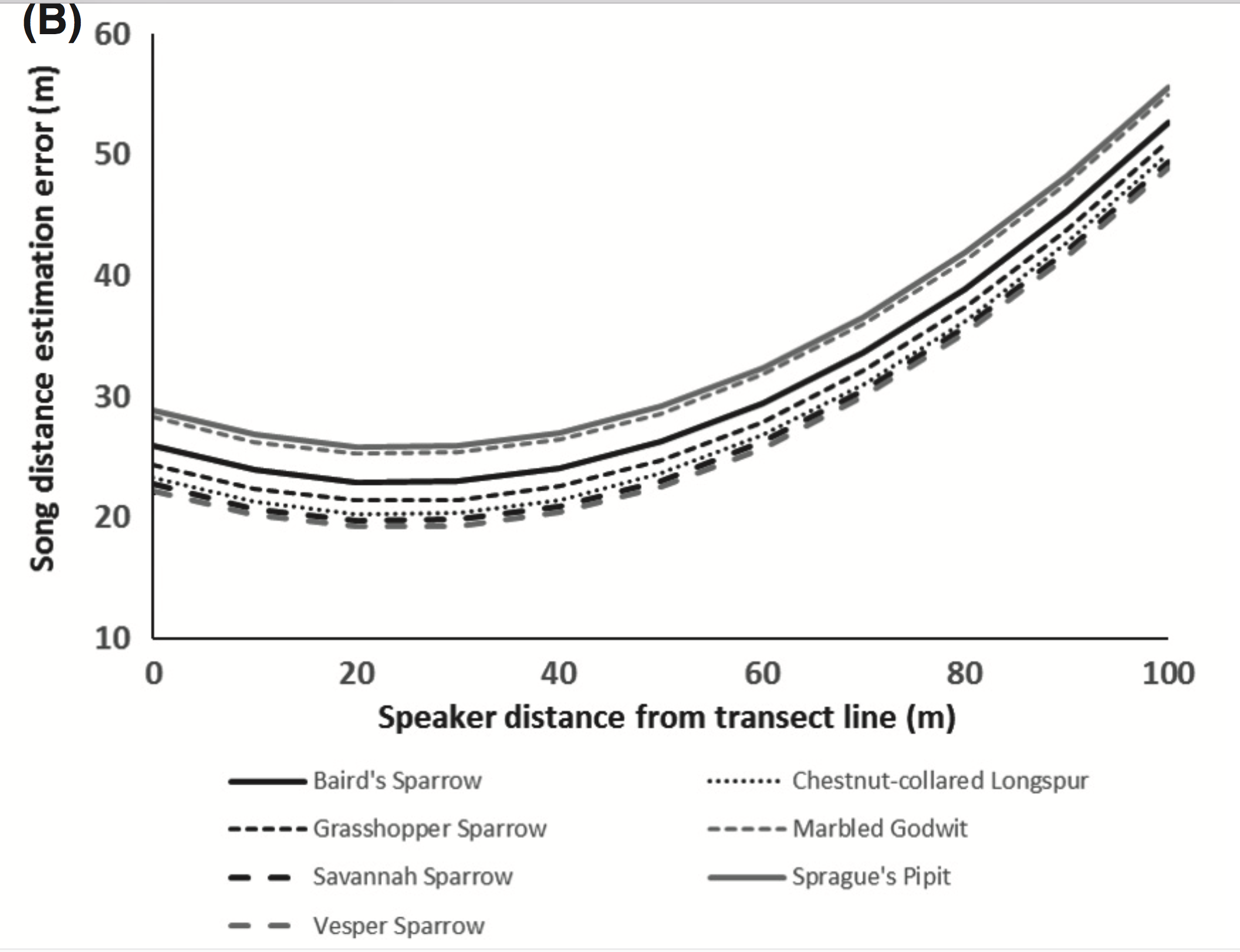
HOME (2009): 80 of 218 records > 50 m (37%); 47 of 218 records > 70 m (22%)

LIBO (2015): 41 of 130 records > 50 m (32%); 17 of 130 records > 70 m (13%)

PIPE (2009): 103 of 234 records > 50 m (44%); 58 of 234 records > 70 m (25%)

*Overall, the percentage of birds recorded at >50 m ranged from 32-49%, while the percentage of birds recorded at >70 m ranged from 13-35%.*

This analysis, of course, assumes all distances to be accurate. There is likely to exist observer error in distance estimates. In a study of observer error in distance estimation of calls of grassland birds, distances were off by ~25 m at an actual distance of 50 m from the source of the calls, and off by ~35 m at an actual distance of 70 m (see below):



**Song distance estimation error in southern Alberta grasslands, 2013–2014.**

From Koper et al. 2016. This study took place in native mixed-grass prairie sites in south-central Alberta, Canada (approximately 50° 53° 58.091′′ N, 112° 26° 35.456′′ W).

**VIII. What is the distribution of distances over which birds are observed?**

Histograms of distances for each park are presented on this and the following pages:

****

**Histogram of the distances at which birds were observed at GWCA in 2016.**

****

**Histogram of the distances at which birds were observed at HEHO in 2009.**

****

**Histogram of the distances at which birds were observed at HOME in 2009.**

****

**Histogram of the distances at which birds were observed at LIBO in 2015.**

****

**Histogram of the distances at which birds were observed at PIPE in 2009.**

**Conclusions:**

Typically, the minimum sample size is determined by evaluating the species accumulation curve and selecting a point at which the curve begins to level off. In this region of the curve, adding additional samples results in very few additional species being discovered. Additionally, estimated species richness is close to observed species richness (>90% of species actually present are detected). Of the five data sets evaluated, this appears to be the case only for GWCA. Sample size could be reduced somewhat at GWCA and most species would still be detected. Reduction of sample size at other parks, however, would result in fewer species being discovered. This is because many species are present at only one or two sites.

An alternative strategy, if only species richness is of importance, is to select sample size based on where the species richness estimator curves begin to flatten out. This would, in theory, allow the estimation of species number without information loss in this variable, but the identities of a smaller number of species would be known. In other words, we would not know which rare species may or may not be present in any given year. This would also result in a slight increase in the standard errors of the population trend estimation using the TRIM method for the more common species (i.e., greater uncertainty in determining trend).

The Chao 2, ICE and ACE species richness estimators performed similarly and yielded precise results in most cases. The Jackknife 2 provided higher richness estimates in most cases. Because true richness is unknown, it is not possible to evaluate the accuracy of these estimators. Based on these results, either the Chao 2, ICE or ACE species richness estimators would be reasonable to employ. The most important consideration, after choosing one, is to use it consistently. The freeware program EstimateS can be used to calculate all these estimators.

It is possible that if the species accumulation curve does not flatten out and species richness estimators are considerably higher than observed species richness, this could be indicative of the presence of numerous rare species rather than insufficient sampling per se. One way to evaluate this is to compare the cumulative species lists recorded from multiple years of sampling in a park to that of the species richness estimators, or actual number of species observed, for a single year. If the cumulative species richness is similar to the actual number of observed species in a given year, and similar to or even lower than the richness estimators, this is consistent with the presence of numerous rare species that are nevertheless detected. If the cumulative species list is considerably higher than the richness estimator, however, this is consistent with insufficient sampling. The percentage of the cumulative number of bird species recorded during surveys over multiple years that the Chao2 estimator represented in a single year ranged from 51 to 79%. The percentage of the cumulative number of bird species recorded during surveys over multiple years that the actual number observed represented in a single year ranged from 49 to 66%. Thus, a fairly large proportion of species known to breed in the park could have been present in a given year and not detected.

Habitat variability could, and almost certainly does, result in greater species diversity. Thus, all else being equal, parks with woodlands and grasslands should have more species than parks with only grasslands. I evaluated the hypothesis that rare species associated with the less common woodland habitats in two of the parks were responsible for the observed patterns in species accumulation as a function of sample size. Deleting woodland sites did not change the observed patterns. Many bird species were still rare even when only grassland sites were considered. Additionally, reduction of sample size in and of itself will result in more species being present at only one or two sites.

If grid dimensions were to be increased from 100 x100 m to 200 x 200 m at these parks, sampled plot numbers would decrease by 71-77%, while number of observed species would decrease by 27-54%. This is assuming, however, that all species observed at the plots were recorded as such in the database. If species were not recorded because they were considered to be already observed from another plot, then it is likely that additional records would be obtained from these plots using a larger grid size, and the decrease in number of observed species be less than that reported here.

If the frequency of recording the same individual multiple times is determined based on the column in the dataset that indicates whether birds were previously observed at another plot, the number of records of previously counted birds is small, ≤ 12% overall, and ≤ 4% at four of the five parks.

If the frequency of recording the same individual multiple times is determined based on distances recorded in relation to the plot spacing, however, this percentage is much higher: The percentage of birds recorded at >50 m ranged from 32-49%, while the percentage of birds recorded at >70 m ranged from 13-35%. Birds were frequently recorded at estimated distances >100m, and occasionally at distances >200 m.

**Recommendations in relation to grid size:**

For sampling birds, most authorities recommend a minimum distance of 250 m between point count stations, perhaps farther in open grassland habitats (e.g., Ralph et al. 1995). Given the small size of these parks, however, such spacing would result in very few sample sites, and an underestimation of species richness. The 100 x 100 m grids currently in use, however, have the potential to frequently result in multiple observations of the same individual, which can inflate estimates of abundance.

There are (at least) four potential options:

(1) *Continue monitoring at the current (100 x100 m) grid size.* The main advantage of this option is that it will provide the most information on species composition, including the identities of the rarer species. The main disadvantage is that many individuals apparently will be counted twice (or more). Such multiple observations would not necessarily be a problem for species in which the distance method is not used, as in such cases only relative abundance information is obtained. In the extreme, however, many multiple observations would result in greater relative abundance estimates than the absolute population abundance, which could provide erroneous inferences. Such multiple observations should have little or no effect on the estimation of trend, as long as the degree of overestimation is similar from year to year.

In theory, observing the same individual twice is not a problem in the distance method, as long as the observations are made from different points, which would be the case here. In fact, most authorities believe that double counting is not a serious problem, and pales in comparison to the problems associated with variability in detection among observers and errors in estimating accurate distances. Rosenstock et al. (2002) state that “Repeated observations of the same individual on multiple visits or during the same visit do not violate the assumption of independence, as long as individuals are not detected more than one time from the same line or point in a given survey. Buckland et al. (1993) state that double-counting generally is of little consequence, particularly “if such events are relatively infrequent. They do warn that “Bias is likely to be small unless repeated counting is common during a survey”, but do not quantify “common”. In the only study that I could find in which this was actually quantified, Alldredge et al. (2008) found that a double counting rate of 8% was associated with overestimates of population size using distance sampling.

(2) *Increase grid size from 100 x100 m to 200 x 200 m*. The main advantage of this option is that fewer individuals would be counted multiple times, although this would still occur occasionally. The main disadvantage is that less information on species identities and composition would be obtained, with many of the rarer species not being observed. The large reduction in sample size (~75%) translates into as many as 50% fewer species observed, and for most parks even the species richness estimators would underestimate true species richness with such a small sample size.

(3) *Continue monitoring at the current grid size but change how data are summarized and analyzed.* Use all data, as is done currently, to determine species richness and composition. For abundance estimation, use only observations with distances less than ~ 70 m, in an attempt to exclude individuals not associated with the sample point. This should have no effect on the estimation of trend, or on relative abundance estimates. Truncating distances may affect estimates obtained by the distance method, however. The conventional wisdom is that data closest to the observer should have the greatest impact on density estimates. Stanbury and Gregory (2009), however, found that truncating distances at 100 m, which reduced their sample by a third, resulted in significant overestimates by the distance method. They stated that the reason for this was “unclear”.

(4) *Employ a grid size intermediate between 100 x100 m and 200 x 200 m*. It is possible to reduce the number of sample sites by some amount less than 75%, which would reduce the problem of multiple observations of the same individual. However, this would also decrease the amount of information on species identities and composition. Given the competing goals of knowledge of species composition versus accurate abundance estimation, any increase in one from changing the sample size would result in a decrease in the other.

Ultimately, the question revolves around a trade-off: Is it more important to obtain information on the identities of the rare species, or on the abundance or relative abundance of the more common species?

**Literature Cited:**

Alldredge, M. W., K. Pacifici, T. R. Simons, and K. H. Pollock. 2008. A Novel Field Evaluation of the Effectiveness of Distance and Independent Observer Sampling to Estimate Aural Avian Detection Probabilities. Journal of Applied Ecology 45:1349-1356.

Buckland, S. T., D. R. Anderson, K. P. Burnham, and J. L. Laake. 1983. Distance Sampling: Estimating abundance of biological populations. Chapman and Hall, London.

Koper, N., L. Leston, T. M. Baker, C. Curry and P. Rosa. 2016. Effects of ambient noise on detectability and localization of avian songs and tones by observers in grasslands. Ecology and Evolution 6: 245–255.

Ralph, C. J., S. Droege, and J. R. Sauer. 1995. Managing and Monitoring Birds Using Point Counts: Standards and Applications. Pages 161-169 *in* C. J. Ralph, J. R. Sauer, and S. Droege, editors. Monitoring Bird Populations by Point Counts. Gen. Tech. Rep. PSW-GTR-149. Pacific Southwest Research Station, Forest Service, US. Department of Agriculture, Albany, CA.

Reese, G. C., K. R. Wilson and C. H. Flather. 2014. Performance of species richness estimators across assemblage types and survey parameters . Global Ecology and Biogeography, 23, 585–594.

Rosenstock, S. S., D. R. Anderson, K. M. Giesen, L. Tony, and M. F. Carter. 2002. Landbird Counting Techniques: Current Practices and an Alternative. The Auk 119:46-53.

Stanbury, A. and R. Gregory. 2009. Exploring the effects of truncated, pooled and sexed data in distance sampling estimation of breeding bird abundance. Bird Study 56:298-309.