

Precision Analysis of Bird Trend Monitoring in NWT Proposed National Wildlife Areas

-- Dave Huggard¹ and Craig Machtans², Feb. 25, 2013 --

¹Apophenia Consulting, 517 E. 10th Street, North Vancouver, BC, V7L 2E7

²Canadian Wildlife Service, Environment Canada, Box 2310, Yellowknife, NT, X1A 2P7
craig.machtans@ec.gc.ca

Precision Analysis of Bird Trend Monitoring in NWT Proposed National Wildlife Areas

1. Overview

The objective of this work was to predict how precise trend estimates would be for songbirds monitored in four proposed National Wildlife Areas (NWA) in the Northwest Territories, with different possible sampling intensities and durations of sampling. The information is intended to provide guidance for design of the bird monitoring program in each area, and to set expectations about how long it will take to obtain trend estimates at particular precision levels for different species. The NWAs are: Edézhíe (Horn Plateau), Samba K'e (Trout Lake), Ka'a'gee Tu (Kakisa) and Ts'ude niline Tu'eyeta (Ramparts)¹. Estimates of the expected precision of trends are based on analyzing data simulated for each target species and NWA with different densities of monitoring sites, monitoring schedules and years of monitoring. The parameters needed for these simulations – bird densities and variance components – were first estimated from field data from Fort Liard provided by CM and from Breeding Bird Survey (BBS) data from the western Canadian boreal and subboreal region (Section 3, with details in Appendix 1). Some additional data from the proposed areas themselves was also used, and all three main sources of data were provided in a consolidated format from the Boreal Avian Modeling database, National V3 Oct2012 (www.borealbirds.ca). Simulation methods are described in Section 4. Results presented in Section 5 include a summary of how the design factors affect expected precision of trends, and the specific precision values expected for each species over time. Full results are available as Excel spreadsheets.

2. Study areas and species

The locations of the four proposed NWAs and the Fort Liard study area are shown in Figure 2.1. The Fort Liard area is in the same geographic vicinity as the three southern NWAs, but is at lower elevation and is probably in more productive forest. The BBS routes are scattered across boreal northern Saskatchewan, Alberta and BC, southern NWT and Yukon and the subboreal central interior BC (Figure 2.2).

The study examined four densities of monitoring sites in the NWAs, based on uniform grids with 30km x 30km, 20km x 20km, 15km x 15km and 10km x 10km spacing. The number of monitoring sites at these spacings that fit into the NWAs is shown in Table 2.1 and illustrated for one area in Figure 2.3.

Table 2.1. Approximate number of monitoring sites in each NWA at 4 spacings.

Spacing (km)	Proposed Park			
	Horn Plateau	Trout Lake	Kakiska	Ramparts
30 x 30	20	11	8	13
20 x 20	45	28	22	27
15 x 15	74	43	35	49
10 x 10	161	90	75	113

¹ A fifth proposed NWA mentioned in the project statement of work, Kwets'ootl'àà, is a small area west of Yellowknife that is mostly in Great Slave Lake, and is not included here.

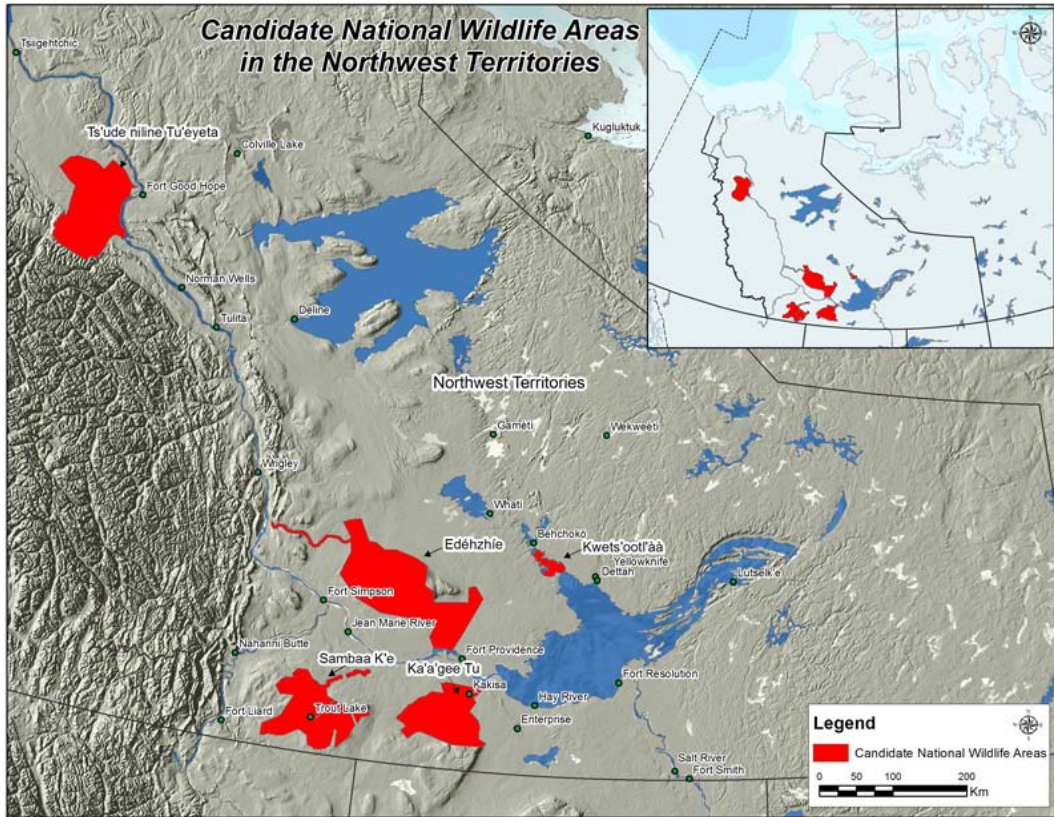


Figure 2.1. Locations of the four proposed NWAs and the Fort Liard study area in NWT. Map provided by Environment Canada.

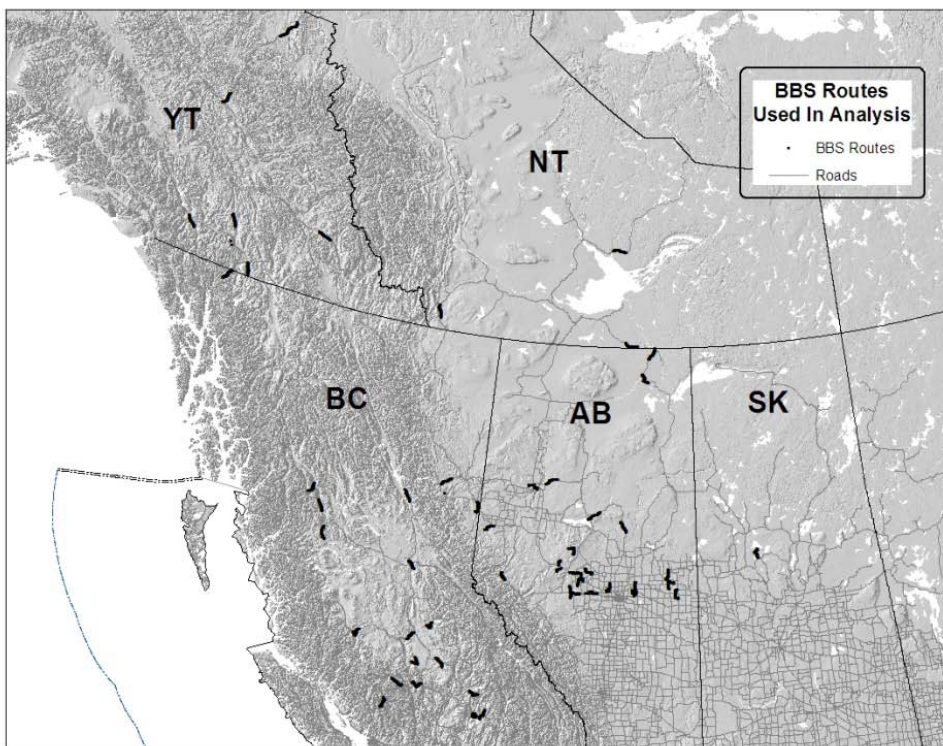


Figure 2.2. Location of BBS routes used in the analysis.

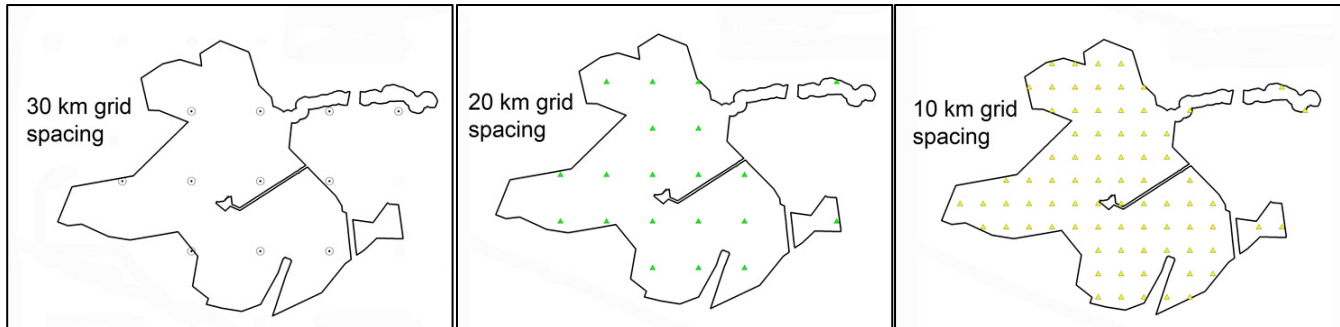


Figure 2.3. Examples of survey sites on 30km, 20km and 10km grids in Smbaa K'e proposed NWA. Number of sites used in the analysis is slightly higher for the 30km spacing, because some sites that fell just outside the area were moved into undersampled areas within the NWA.

The analysis was conducted for 51 species (50 birds plus red squirrel) that had enough records to analyse in the Fort Liard dataset. However, red squirrels were not recorded in the BBS data or in the available data from one-year surveys in the four NWAs, and were excluded from subsequent analysis. Several species did not occur in particular NWAs, and so have no estimates of precision for those NWAs.

3. Parameter estimates

3.1 Summary of analysis

The simulation to estimate expected precision of trend for a particular species requires a mean count, and several variance components. The mean count was estimated from single years of surveys that have been conducted in each NWA. The values for the three southern NWAs were averaged, due to the limited data available from the single year of surveys in each NWA. The mean counts of species in the northern Ts'ude niline Tu'eyeta NWA used only the single year of data for that area, because it is in a distinctly different geographic region. The simulation also uses a value for the proportion of sites at which a species never occurs (non-habitat). Because this requires multiple years of data to estimate reliably, the value for each species from the Fort Liard study area was used.

The variance components needed for the simulations were derived from Craig Machtans' Fort Liard dataset and the boreal BBS dataset. The Fort Liard study has survey results from two sessions per year at 58 sites (some yearly variation) over 14 years, with surveys in years 1, 2, 3, 4, 5, 8, 11 and 14.

The BBS provided yearly surveys over 15 years at 53 sites, with some missing years at particular sites. The following variance components were derived from these data sets (details of the General Linear Mixed Model (GLMM) variance analysis are in Appendix 1):

1. Yearly variation around trend. This is the yearly variation of the overall (detectable) population in a study area around the true trend. Yearly variation is a hybrid of Fort Liard and broader BBS estimates.
2. Autocorrelation of yearly variation. This measures the extent to which yearly variation is correlated, positively or negatively, with the yearly variation from previous year(s). The estimate comes only from BBS, because the five consecutive years at Fort Liard are too short to support those estimates.
3. Variation among sites. This represents inherent variability among sites for the species (e.g., due to habitat, site productivity, etc.). Estimates of site variation come from the Fort Liard study, which sampled sites in a way similar to that proposed for the NWA monitoring. The single years of pilot

monitoring in the NWA areas could also provide information on variation among sites. However, these estimates were much weaker than those from the multi-year Fort Liard study, and in the same general range, so only the Fort Liard values were used in the simulations. BBS route-level variation was not used, because of the much different sampling design for the BBS, and also the greater variety of habitats they represent.

4. Variation in the trend for individual sites (site*year variation). This variance component represents how much individual sites' trends differ from the overall trend (because of local disturbances, succession, etc.). This component is estimated from the BBS data only, because the models did not run reliably for the Fort Liard data (see Appendix 1). There could also be some covariation of site variance and site*year variation, but this was ignored here because the site and site*year variation are estimated from different datasets. This covariation was also low in most cases where it could be examined in the BBS data.
5. Variation among point count stations within a site. This value comes from the Fort Liard data and is needed to be able to simulate a different number of stations per site. It is also used to account for extra-Poisson variation at the site level.

The simulation modeling (next section) was all done on the log-scale, except the final step of assigning actual counts. The variance components were therefore also on the log-scale. This means that the same variance values could be used even where a species differed in absolute abundance among proposed NWAs.

Estimating several parameters for many species inevitably produces some results that are too extreme, just by chance with many uncertain values (Link and Sauer 1996). One way to reduce this problem is to use empirical Bayes smoothing, in which a particular parameter for a species is combined in a weighted average with the average value for a group of similar species. For this analysis, species were grouped into high, medium and low abundance species, with a fourth group for irruptive species (crossbills, siskins). Ideally, the weighting of the two estimates, the individual species' and the group average, is proportional to $1/\text{variance}$ of the estimate. However, the GLMM approach does not produce error estimates on the variance components. As an approximation, the weighting was inversely proportional to the number of observations of the species, scaled so that a species with 50 observations had equal weight for its estimate and for the group average. (A species with <50 observations would have more weight assigned to the group average; a species with >50 observation less weight to the group average).

3.2 Results – variance components

Mean counts in the two NWA areas and estimates of variance components are shown in Table 3.1. Yearly variance, which is a dominant factor in SE of trends, is lowest for three widespread easily-detected species (Swainson's thrush, red-eyed vireo and American robin), with values ranging up to 20 times as high for rarer species and for some species that may be in marginally productive habitat in these areas. Four of the 5 highest yearly variances are for irruptive species. Site variance shows a smaller range of values, and is less strongly tied to species' abundances. (Note, however, that relatively low site variance can occur for species that are absent from a high proportion of sites – they have more

Table 3.1. Mean counts in the two NWA regions (3 southern NWAs averaged together), proportion of sites where the species is (always) absent, and estimates of variance components used in the simulations. Table is sorted by yearly variance, a dominant factor in trend SE.

Species	Mean count		Proportion		Variance components			Station
	Southern NWAs	Ts'ude niline Tu'eyeta	0 sites	Yearly	AR1	Site	Site*Year	
SWTH	0.815	0.850	0.000	0.0053	-0.379	0.091	0.0041	0.008
REVI	0.126	0.026	0.017	0.0073	0.263	0.909	0.0055	0.070
AMRO	0.315	0.445	0.033	0.0073	-0.113	0.791	0.0027	0.174
YRWA	0.809	0.502	0.000	0.0120	0.053	0.271	0.0039	0.039
WETA	0.045	0.004	0.085	0.0128	0.284	0.984	0.0086	0.021
OVEN	0.102	0.000	0.102	0.0134	-0.279	1.553	0.0073	0.013
WTSP	0.680	0.093	0.017	0.0178	0.000	0.527	0.0168	0.125
LEFL	0.088	0.026	0.474	0.0192	0.004	0.675	0.0082	1.177
WAVI	0.020	0.031	0.228	0.0197	-0.346	0.987	0.0058	0.258
AMRE	0.056	0.013	0.254	0.0219	-0.123	0.821	0.0124	0.400
GRAJ	0.366	0.053	0.118	0.0231	0.093	0.823	0.0041	0.214
CORA	0.076	0.035	0.212	0.0246	-0.239	0.841	0.0093	0.295
AMCR	0.001	0.000	0.614	0.0268	0.389	0.149	0.0086	0.721
CHSP	0.819	0.507	0.034	0.0283	-0.371	0.543	0.0053	0.022
YWAR	0.090	0.463	0.304	0.0312	0.021	0.160	0.0072	0.727
OCWA	0.128	0.445	0.500	0.0316	-0.397	0.158	0.0093	0.718
YBSA	0.016	0.009	0.017	0.0343	-0.076	0.471	0.0129	0.059
LISP	0.314	0.595	0.500	0.0348	-0.242	0.228	0.0086	0.644
MAWA	0.191	0.009	0.000	0.0351	-0.010	0.386	0.0101	0.056
DEJU	0.339	0.291	0.406	0.0381	0.491	1.195	0.0053	0.164
COYE	0.107	0.004	0.856	0.0386	-0.174	0.156	0.0055	0.765
HETH	0.522	0.137	0.712	0.0403	0.126	2.214	0.0100	0.149
CAWA	0.010	0.000	0.288	0.0462	-0.363	0.957	0.0160	0.196
ALFL	0.441	0.441	0.596	0.0494	0.271	0.246	0.0088	0.772
RCKI	0.710	0.269	0.373	0.0502	0.224	1.098	0.0061	0.205
WEWP	0.012	0.009	0.414	0.0505	-0.242	0.158	0.0113	0.727
NOWA	0.113	0.405	0.569	0.0609	0.411	0.212	0.0094	0.920
BCCH	0.003	0.000	0.384	0.0678	-0.191	0.144	0.0089	0.728
RBGR	0.010	0.000	0.032	0.0688	-0.211	0.847	0.0110	0.166
HAFL	0.004	0.000	0.590	0.0704	-0.040	0.239	0.0169	0.759
TEWA	0.546	0.185	0.000	0.0708	0.224	0.175	0.0082	0.011
PAWA	0.616	0.207	0.400	0.0711	-0.431	2.154	0.0081	0.196
FOSP	0.156	0.762	0.300	0.0789	-0.389	1.379	0.0141	0.176
BBWA	0.013	0.000	0.051	0.0859	0.427	0.697	0.0277	0.079
BOCH	0.040	0.000	0.141	0.0873	0.433	0.120	0.0121	0.618
BHVI	0.061	0.000	0.214	0.0905	-0.363	0.128	0.0144	0.700
MOWA	0.004	0.000	0.595	0.0932	-0.098	0.145	0.0125	0.783
WIWR	0.000	0.000	0.338	0.1071	0.000	0.839	0.0100	0.200
YBFL	0.069	0.018	0.780	0.1148	0.097	1.331	0.0176	0.242
VATH	0.013	0.009	0.695	0.1149	-0.188	0.202	0.0074	0.645
BAWW	0.095	0.018	0.347	0.1153	0.144	0.206	0.0103	0.593
PUFI	0.000	0.004	0.240	0.1166	0.149	0.152	0.0153	0.691
PHVI	0.006	0.000	0.520	0.1278	-0.558	0.157	0.0129	0.720
RBNU	0.016	0.000	0.119	0.1416	0.479	0.415	0.0116	0.105
GCKI	0.006	0.000	0.514	0.1423	-0.113	0.158	0.0115	0.731
EVGR	0.001	0.000	0.584	0.2244	-0.074	0.154	0.0254	0.743
CMWA	0.060	0.000	0.336	0.3132	-0.181	0.915	0.0089	0.226
BOWA	0.039	0.009	0.881	0.4045	-0.010	0.144	0.0116	1.122
WWCR	0.118	0.093	0.000	0.4927	-0.008	0.375	0.0442	0.913
PISI	0.037	0.009	0.000	0.9123	0.215	0.408	0.0206	0.333

Notes: The counts are ordinal scale, but the variance components are on the log-scale. They are variances - take the square root for standard deviations. AR1 is the first-order autoregression coefficients of yearly variation.

consistent low values at sites where they occur. The simulation includes the extra binomial component of site-level variance from these species being absent at some proportion of sites.) The station variance tends to be lower, except for a few species that aggregate (crows, least flycatchers) or have particular habitat preferences within sites (yellowthroats, waterthrushes, etc.).

4. Simulation of trend precision

Monitoring data was simulated separately for each species and NWA, for:

1. True trend: -8, -5, -2, 0, 2, 5, and 8%/yr.
2. Site spacing: 30km x 30km, 20km x 20km, 15km x 15km and 10km x 10km. Because the NWAs differ in size, these spacings produce a different number of sites in each NWA (Table 2.1).
3. Monitoring schedule: Nine revisit designs were examined. In five of these, all sites in a NWA are visited in the same year, at 1-5 year intervals. In the other four designs, the NWA is visited each year, with 1/2, 1/3, 1/4 or 1/5 of sites surveyed each year, so that all the sites are visited within 2, 3, 4 or 5 years.
4. Monitoring duration: 11, 16, 21, 31yr. These values were used (rather than 10, 15, etc) so that there would be 3, 4, etc. visits with the design in which all sites in a NWA were revisited every 5 years (i.e., visits in years 1, 6, 11, 16...)

All simulations used 9 point count stations at a site.

For a particular combination of these factors, the steps in the simulations for each species and NWA were:

1. Number of sites occupied. This was drawn from a binomial distribution with the number of sites being simulated and the observed proportion of sites occupied. Sites where the species is not present are assumed to contain no suitable habitat for the species.
2. Mean abundance at occupied sites. The mean count of the species in the NWA or group of NWAs (Table 3.1) includes sites where the species was absent, so this number is adjusted to the mean count *where the species is present* (at least in some years) by dividing the value reported in Table 3.1 by (1 – proportion of sites where the species is always absent). The value was converted to the log scale for subsequent analysis.
3. True yearly means. The true trend was applied for years 1 to 31, as a linear change on the log scale: $\log(\text{Abundance}) = \log(\text{Initial abundance}) + \text{Year} * \log[(100 + \text{Trend})/100]$, where trend is in %/yr. The trend is therefore exponential growth or decay on the ordinal scale. Note that while yearly variation is added below, there is no “process noise” (Humbert et al. 2009) here – the true rate of population change stays the same. This is most likely to be valid for shorter monitoring durations.

The “true yearly mean” here can be interpreted as the population size averaging across sources of yearly fluctuations: demographic stochasticity, movements of species in or out of the area and variation in detectability.

Subsequent steps were repeated for each of 100 Monte Carlo simulations:

4. Variation of yearly means. Autocorrelated stochastic yearly variation was added around the true trend, using R module *arima.sim*. Note that yearly variance calculated in the GLMM analysis did not include autocorrelation – the AR coefficients were calculated separately. Because autocorrelation can increase or decrease the variance, the yearly variance used in *arima.sim* was first adjusted so

that the variance of the resulting autocorrelated values would equal the estimated yearly variance. This was done empirically, by calculating the ratio of yearly variance to the variance of a large number of simulated autocorrelated values with the estimated AR coefficients for the species.

Yearly variation can be positive or negative. Positive variation in a year would occur from better than average reproduction or survival in the previous year, more birds settling within the area (rather than, for example, migrating further or elsewhere) or conditions that produce better than average detectability. We cannot distinguish these sources of yearly variation in the available data.

5. Site variation. Variation around the yearly mean was added for each site’s mean from a normal distribution (on the log scale) with mean 0 and variance equal to the site variance. The deviation from the yearly mean stays the same (on the log scale) for a given site. I.e., this variation represents permanent variation in site quality.
6. Variation in trend for each site. This component was added to the yearly values for each site by adding the results of a truncated-normal random-walk (on the log scale) with mean 0 and variance for each time step calculated to produce the estimated site*year variance at 14 years (the length of each time period used in the BBS variance partitioning analysis). This random-walk approach was used instead of a simple normally distributed linear term (on the log scale) for site-trend variation, because the exponential growth implied by a log-linear term produced some sites with unrealistically huge numbers of birds in the simulations with long durations. A random-walk is also more appropriate biologically, because particularly high growth or loss at a particular site cannot

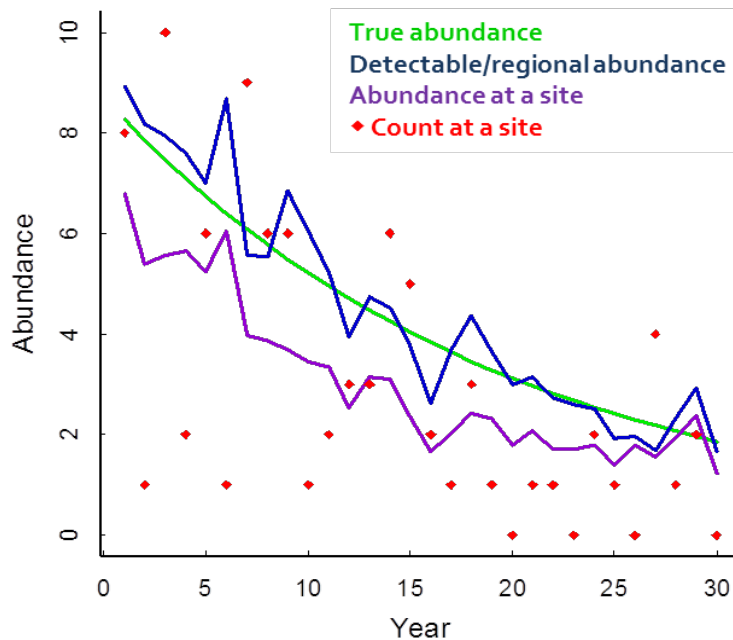


Figure 4.1 Example of steps in simulation: 1) True abundance (green) is the exponential change from current abundance (here, -5%/yr). It can be interpreted as the abundance, averaging regional demographic variability and average detectability. 2) Changes in detectable abundance in the monitoring region (blue) adds autocorrelated yearly variation. This may be above “true abundance” if yearly regional population or detectability are above the long-term average. 3) The mean for a particular site (violet) differs from the regional mean (in this case, a moderate negative difference) and variation in the trend at the site is added as a random walk. 4) The actual count for a year at the site (red points) is the sum of Poisson counts at the 9 point count stations, including normal variation among the stations.

persist indefinitely. Instead, the local abundance adjusts to the improved or deteriorated conditions, then responds to subsequent improvements or disturbances, as represented by the random walk.

7. Adjustment of mean for addition of log-normal error. Adding normally distributed error on the log scale increases the mean when the values are back-transformed to the ordinal scale (e.g. Thomas and Martin 1996). This upward bias in the mean needs to be removed so that simulated counts are based on the correct mean. This was done by adjusting the mean of the back-transformed site values for each year to the back-transformed year mean. With the log-scale values, this is done by adding $\log(\exp(\text{Year}_i)/\text{mean}(\exp(\text{Site}_{j,i})))$ to the log-scale $\text{Site}_{j,i}$, where Year_i is the yearly mean for year i (step3) and $\text{Site}_{j,i}$ is the value for the j^{th} site in that i^{th} year. This removes the upward bias, while preserving the log-scale variation among years and sites.
8. Station variation. Station-level variance was added to the site value for each of the 9 stations at a site, using normal variation on the log scale. An adjustment similar to step 7 was again applied to ensure that the ordinal-scale mean of the 9 stations equalled the site value.
9. Actual count. The actual count at each station was simulated using a Poisson distribution based on the station value. The station values were then summed to the site level, which was used as the sample unit. The station variance added in step 8 accounts for any extra-Poisson variation (aggregation) at the site level.
10. Subsetting for multi-year revisit cycles. For revisit cycles of more than one year or the designs where a fraction of sites were surveyed each year, only results for the appropriate years and sites were used in the trend analysis (next step).
11. Regressions for trend. Trend was calculated for all simulations using a simple log-link Poisson model, with the site as the sample unit. The slope coefficient is the log-scale linear trend, which was converted back to ordinal scale percent annual change as $100*(e^{\text{slope}}-1)\%$. The standard deviation of the ordinal scale %/yr trends from the 100 Monte Carlo iterations gives the expected standard error of the trend estimate for a particular species and combination of factors.

5. Results

5.1 Design factors

5.1.1 Effect of true trend

The expected SE of the trend estimate is slightly higher when the true trend of the population is extremely negative or positive (Figure 5.1). A strong negative trend, particularly if it lasts for a long time (21, 31 years) leads to low mean counts of a species, which increases the relative effect of Poisson count error. The extreme positive trend exacerbates the effects of the annual variability, producing somewhat more variation in the trend estimates. In both cases, the increase in uncertainty with extreme trends is relatively small, so subsequent results are shown only for the scenario with 0%/yr true trend.

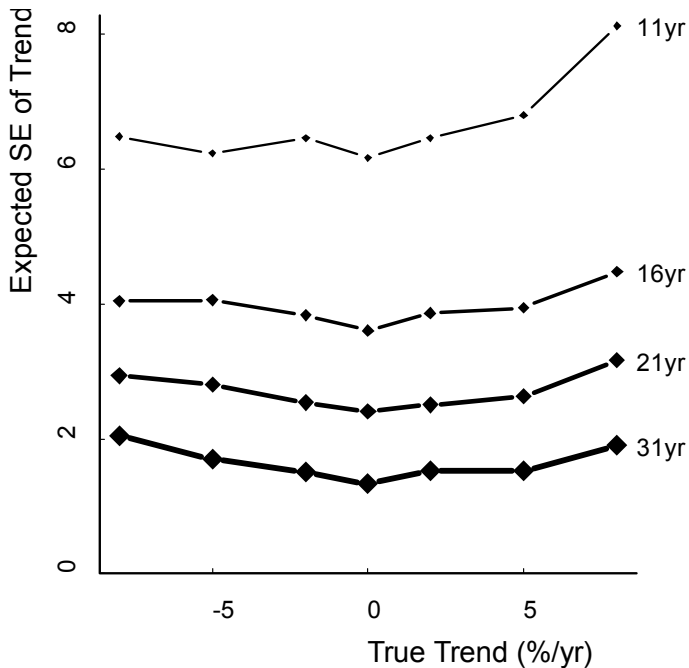


Figure 5.1. Effect of true trend (%/yr) on expected SE of trend estimate (%/yr), after 11, 16, 21 and 31 years. Values shown are the median across all species, for Sambaa K’e (Trout Lake) NWA with 43 monitoring sites (15x15km spacing) monitored every fifth year.

5.1.2 Number of sites and remeasurement schedule

Small numbers of monitoring sites in a NWA increase the expected SE considerably, but then there are diminishing returns for increasing the number of sites (Figure 5.2). The example in Figure 5.2 is for Sambaa K’e (Trout Lake), so there are only 11 sites at 30x30km spacing, increasing to 28, 43 and 90 at 20x20, 15x15 and 10x10 spacing, respectively. This figure is only meant to give an overview of the effect of number of sites. Actual decisions about how many sites to survey should examine the detailed results, because: 1) This figure is for the median across all species. Individual species, particularly rare ones, may show greater increases in precision with higher numbers of sites. 2) These results are for 21 years, with 5-year revisit period, which means only 5 surveys are included. Annual variability is a dominant factor with few years surveyed, limiting any gains in precision from more sites. More sites may have a greater benefit for longer monitoring durations, and particularly for schedules with more frequent monitoring (such as the schedules with annual monitoring of ½ to 1/5 the sites; red lines in Figure 5.2). In those cases, the frequent monitoring reduces the effects of the annual variability, so that the benefits of more sites is more apparent (the red lines continue to decline at denser spacing).

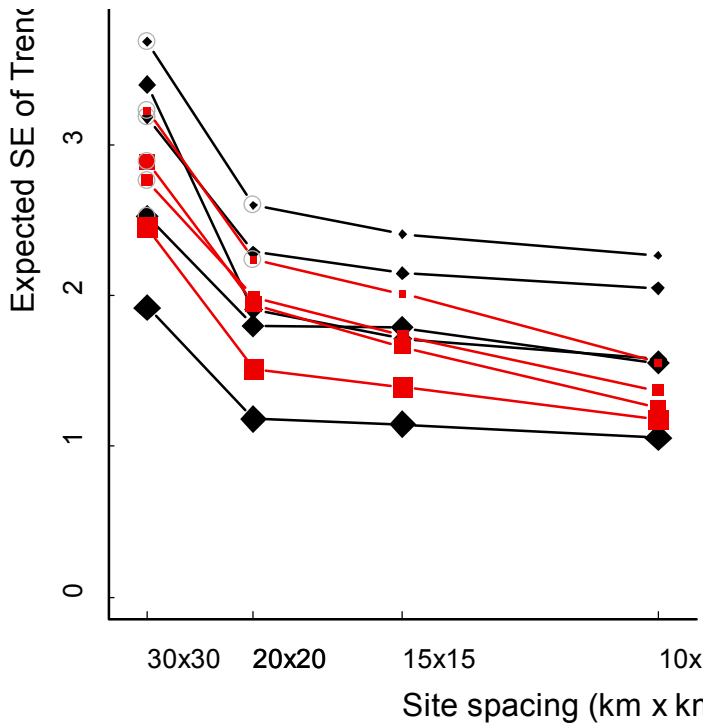


Figure 5.2. Effect of site spacing (number of sites) on expected precision of trend (SE; %/yr). Black diamonds (◆) are for surveys of all sites, ranging from every year (largest symbols) to every 5th year (smallest symbols). Red squares (■) are for ½ sites monitored each year (large symbol) to 1/5 sites monitored each year (smallest symbol). The values are the medians across all species, for Samba K'e (Trout Lake) NWA with 21 years of monitoring. Thin grey circles around some points indicate where >6 species failed to run; values may be underestimates in these cases (because the species that fail to run are rare and would likely have increased the median).

5.1.3 Duration and remeasurement schedule

Increased monitoring duration sharply decreases the expected SE of the trend estimate, as is typically found in this type of analysis (Figure 5.3). The SE decreases more than (inverse) proportionally to duration – doubling duration reduces the expected SE by more than 50%.

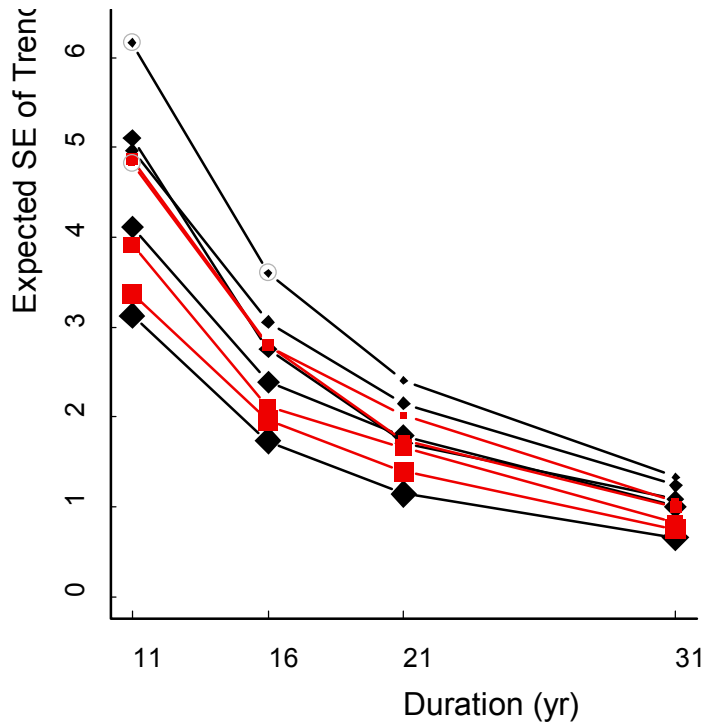


Figure 5.3. Effect of monitoring duration (years) on expected precision of trend (SE; %/yr). Black diamonds (◆) are for surveys of all sites in a year, ranging from every year (largest symbols) to every 5th year (smallest symbols). Red squares (■) are for ½ sites monitored each year (large symbol) to 1/5 sites monitored each year (smallest symbol). The values are the medians across all species, for Samba K’e (Trout Lake) NWA with 21 years of monitoring. Thin grey circles around some points indicate where >6 species failed to run; values may be underestimates in these cases (because the species that fail to run are rare and would likely have increased the median).

Monitoring every fifth year (small black points in Figure 5.3) for 21 years (5 visits) produces a substantially lower SE than monitoring every year (largest black points) for 11 years (11 visits), emphasizing that duration itself is a dominant factor, not just the total number of visits. The trade-off between total estimated cost and expected SE is shown further for 5 remeasurement designs after 11, 16, 21 and 31 years in Figure 5.4. The values are for a single WMA, Samba K’e (Trout Lake). The estimated cost was based simply on \$3000 per surveyed site (Figure 5.4a), or \$3000 per surveyed site *plus* overhead of \$20,000 per park per year (Figure 5.5b). The same overhead applies in a year whether all sites are surveyed in a WMA in that year or just a fraction. These costs are rough estimates, meant to illustrate the cost-versus-SE trade-off for different designs.

Expected SE of the trend declines with duration of monitoring, and total cost increases. The decline is sharpest with the designs that measure all sites every 5th year or 1/5 of sites each year. The design measuring all sites each year obtains lower SE’s more quickly than the less frequent or less intense monitoring, but at much greater cost. Interpolating between points on Figure 5.4a, obtaining a median SE of 2%/yr by monitoring all sites every 5th year would take 26 years² and cost \$550,000, compared to 15 years at a cost of \$1,250,000 by monitoring all sites every year.

² 23 years by interpolation, but the 5-year monitoring would not happen until year 26.

The overhead component is important in determining which design is most cost efficient. Without overhead, visiting 1/5 of sites each year is most efficient. With overhead, visiting all sites every 5th year becomes more efficient (because the overhead cost of sending a crew to the park is only incurred in 1 year out 5, rather than every year). A more detailed cost-efficiency analysis would need better estimates of the different cost components, and consideration of the value of obtaining precise trend estimates sooner rather than later.

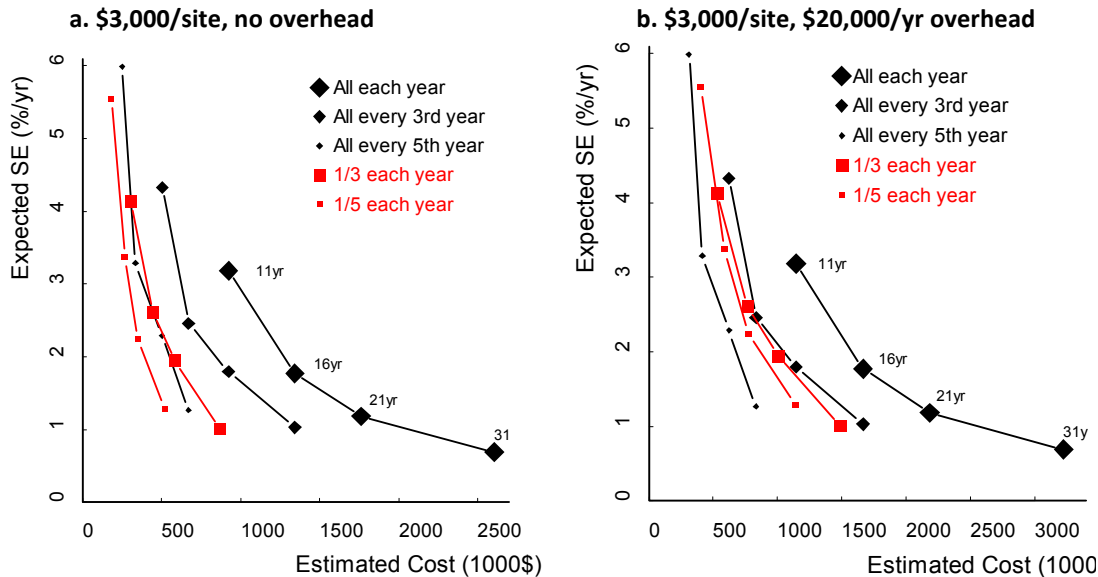


Figure 5.4. Simple example of change in expected SE of trend as a function of total monitoring cost, for 5 remeasurement designs, assuming: a) \$3000/site with no overhead; b) \$3000/site + \$20,000 overhead for each year that a park is visited. Results are the median expected SE across all species, for 20km x 20km spacing in Sambaa K'e (Trout Lake) WMA.

5.2. Expected precision for individual species

The above results were based on the median expected SE across all species, and may give some guidance for monitoring design. However, it is also important to consider the variation among individual species, for designing the monitoring and for setting expectations about how many species will be monitored to some target precision with a given effort and duration of monitoring. Figure 5.4 shows the variation among species in expected SE as a function of duration of monitoring, for different spacings of sites, and for all sites measured every fifth year (Figure 5.4a) and for 1/5 of sites measured each year (Figure 5.4b). Each species shows the same general response to the design variables, but there is at least an order of magnitude difference in the actual expected SE values.

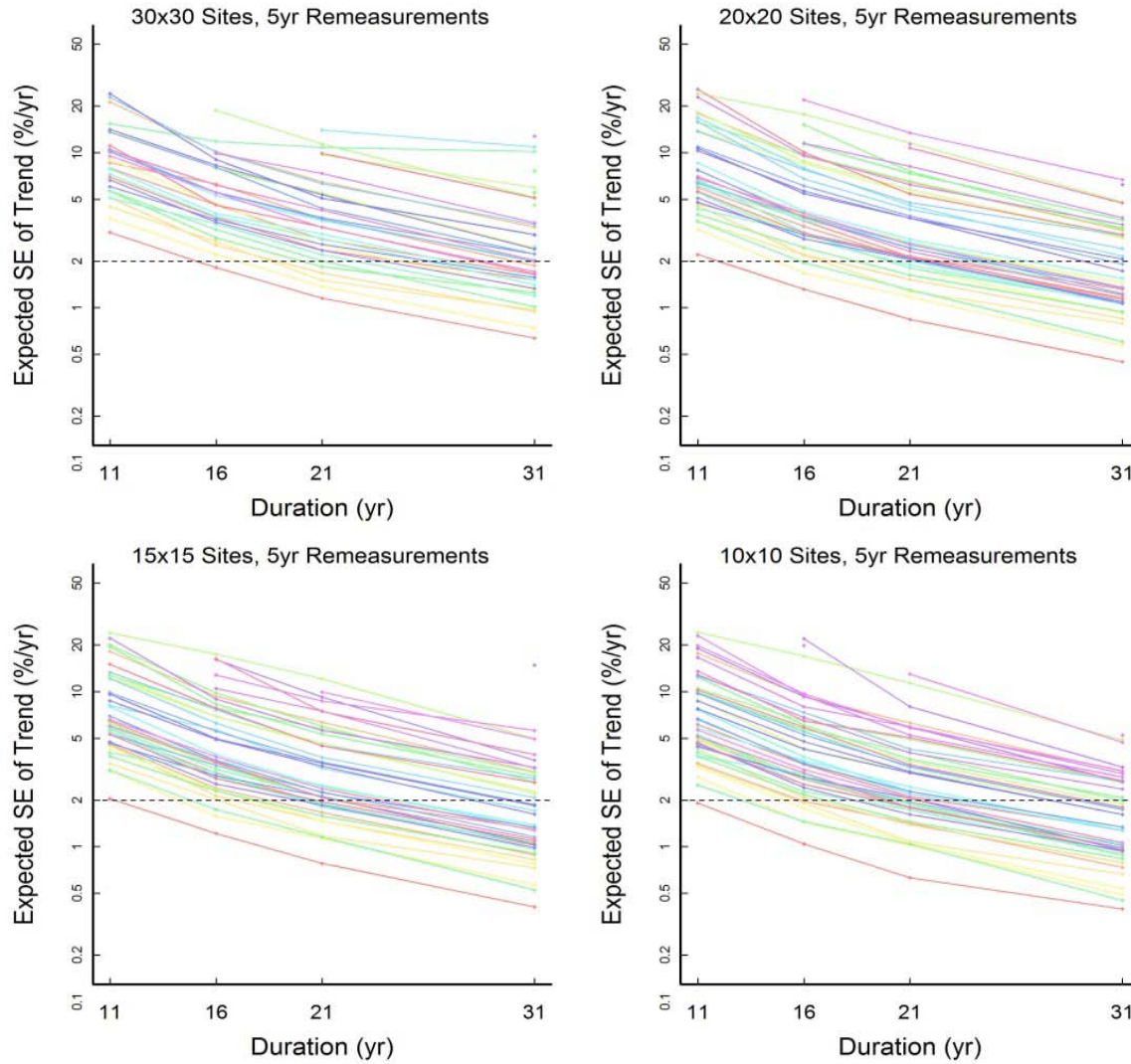


Figure 5.4a. Expected SE for individual species as a function of monitoring duration, with different spacings of sites, for all sites monitored every fifth year. Each coloured line is a species. Some lines are short where the model did not run (for rare species at shorter durations, where all counts can be 0).

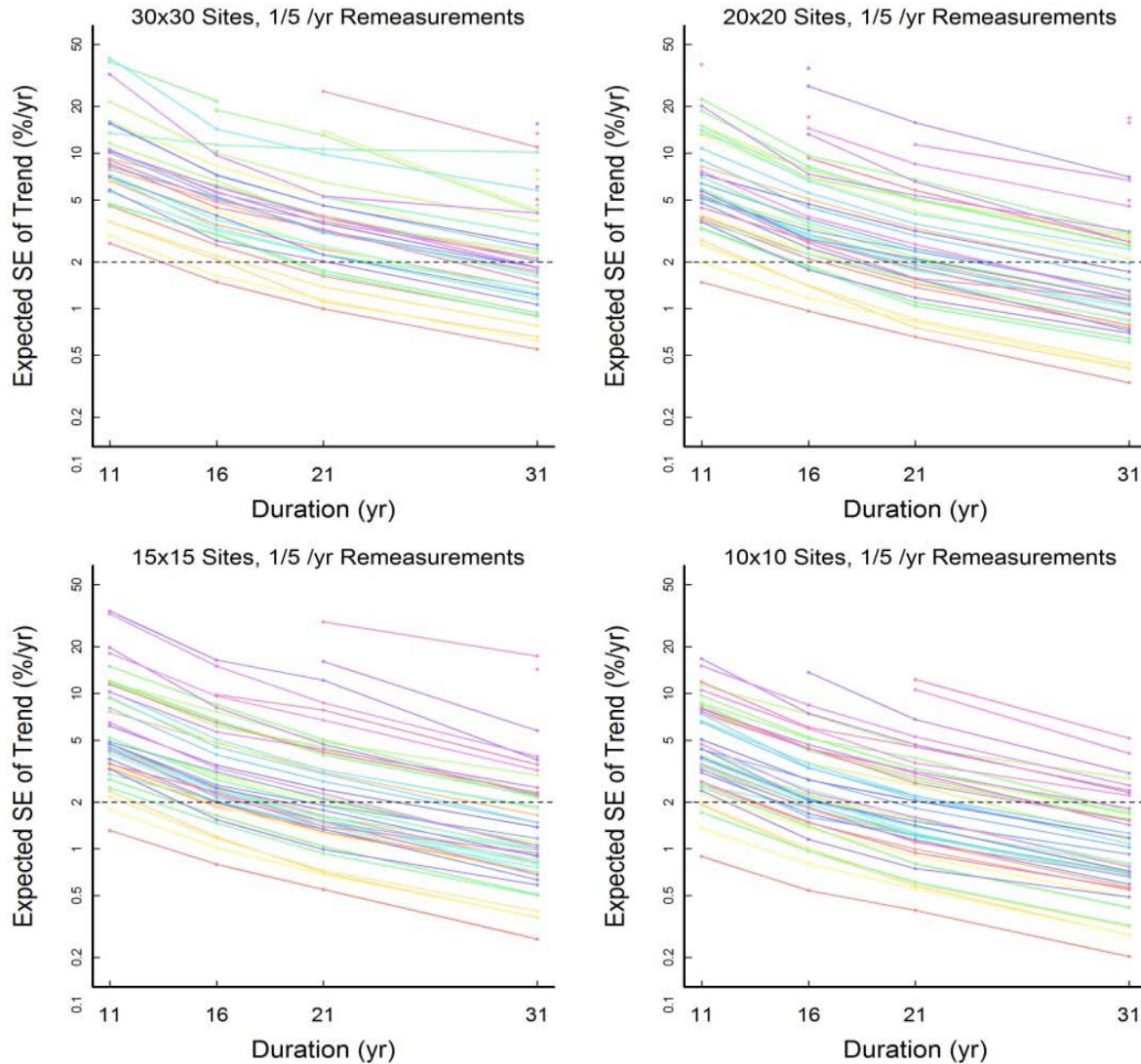


Figure 5.4b. Expected SE for individual species as a function of monitoring duration, with different spacings of sites, for 1/5 of sites monitored each year. Each coloured line is a species. Some lines are short where the model did not run (for rare species at shorter durations, where all counts can be 0).

The expected SE for each species in each NWA under all combinations of the design variables are available in an Excel spreadsheet (**SE of trend Complete summary table.xlsx**³). These results were also summarized as the number of years before the SE of the trend for a species is expected to reach <2%/yr (**Years to SE 2 for each species.xlsx**). This value was chosen because it would produce 87% confidence intervals that do not overlap 0 for a trend of $\pm 3\%/yr$, which may be a reasonable basis for a management decision. A SE of 2%/yr would also produce 95% confidence intervals that do not overlap 0 for a trend of $\pm 4\%/yr$.

An example of the years that each species is expected to take to reach a SE of 2%/yr is given in Table 5.1 for the schedule with all sites monitored every fifth year. Species are sorted by the expected

³ Blank cells in these tables indicate that the trend models did not run on the simulated data, usually because there were too many years with 0 records of the species. SE's of trend can be assumed to be very high for these blank cells, and time to reach a SE of 2%/yr much >31 years.

years in the Edézhíe (Horn Plateau) are with 15x15km site spacing. This also sorts the other columns, more-or-less. With this 5-year remeasurement schedule, about 20 species are expected to reach the 2%/yr SE value in <21 years in the southern NWAs, with the denser site spacings. About ten additional species would reach this SE level in 22-30 years. The numbers of species are lower in the Ts'ude niline Tu'eyeta (Ramparts) area.

The results were further summarized as the number of species reaching the 2%/yr SE level in each NWA, with all the design combinations (Table 5.2). These tables show the strong effect of monitoring duration on the number of species with precise trend estimates, the effect of increasing density of sites particularly with shorter monitoring durations, and the substantial effect of more frequent monitoring. There is also a modest increase in the number of species reaching the target SE if some sites are monitored every year, compared to all sites monitored every few years.

Table 5.1. Expected years to SE <2%/yr for each species and NWA, with four spacings of monitoring sites, based on all sites monitored every 5 years.

Species	Edehzie (Horn Plateau)				Sambaa K'e (Trout Lake)				Ka'a'gee Tu (Kakisa)				Ts'ude niline Tu'eyeta			
	30x30	20x20	15x15	10x10	30x30	20x20	15x15	10x10	30x30	20x20	15x15	10x10	30x30	20x20	15x15	10x10
SWTH	14	11	<11	<11	15	12	11	<11	16	13	12	11	15	13	12	<11
YRWA	15	14	14	13	17	15	15	14	18	15	14	13	19	16	15	14
AMRO	17	15	14	13	21	16	15	13	23	16	14	13	18	15	14	13
REVI	21	15	15	14	25	20	18	15	29	21	19	15	>31	>31	26	20
WTSP	17	16	15	15	19	17	16	16	19	17	16	16	29	22	20	18
GRAJ	18	16	15	15	20	19	18	17	23	19	18	16	>31	26	22	19
OVEN	25	18	17	16	30	23	19	16	>31	21	19	16				
CHSP	18	18	17	17	19	17	17	16	21	20	20	19	20	19	18	18
LEFL	26	21	19	17	>31	25	22	18	>31	23	21	19	>31	>31	28	22
LISP	20	19	19	19	24	22	20	19	27	21	21	19	20	19	19	18
OCWA	24	20	19	18	29	22	21	20	>31	24	21	20	22	20	20	19
WETA	29	21	19	16	>31	26	25	21	>31	29	24	19	>31	>31	>31	>31
DEJU	22	20	19	19	27	20	20	19	27	22	21	21	26	22	20	20
CORA	27	23	19	18	31	25	23	21	>31	26	24	21	>31	>31	27	24
MAWA	22	20	20	19	28	22	21	20	30	22	21	20	>31	>31	>31	29
HETH	25	21	20	20	>31	21	21	20	29	21	19	18	>31	28	21	20
YWAR	26	22	21	20	>31	24	22	20	>31	26	23	20	20	19	19	19
AMRE	30	22	21	20	>31	26	22	20	>31	29	26	21	>31	>31	>31	26
RCKI	24	22	21	20	23	20	20	20	27	22	21	21	27	23	21	20
ALFL	24	21	21	21	27	22	22	21	27	24	21	21	24	22	21	21
COYE	26	26	22	21	29	26	24	22	27	27	24	21	>31	>31	>31	>31
NOWA	29	24	22	22	>31	25	24	24	>31	28	26	23	26	23	22	22
TEWA	24	24	24	23	25	23	22	22	28	27	26	26	27	25	25	24
PAWA	26	25	25	25	28	26	25	25	29	27	25	24	29	27	25	26
FOSP	27	26	25	24	>31	27	25	24	>31	28	26	25	27	27	26	25
WAVI	>31	30	26	20	>31	>31	29	23	>31	>31	30	26	>31	31	26	20
BHVI	31	29	28	27	>31	30	29	28	>31	30	28	27				
YBFL	>31	30	28	27	>31	>31	>31	29	>31	>31	>31	30	>31	>31	>31	>31
YBSA	>31	>31	29	25	>31	>31	>31	29	>31	>31	>31	31	>31	>31	>31	29
BAWW	30	30	29	28	>31	31	30	29	>31	30	29	28	>31	>31	>31	>31
BOCH	>31	31	30	29	>31	>31	30	28	>31	>31	>31	28				
AMCR	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31				
BBWA	>31	>31	>31	29	>31	>31	>31	30	>31	>31	>31	>31				
BCCH	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31				
BOWA	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31
CAWA	>31	>31	>31	29	>31	>31	>31	31	>31	>31	>31	>31				
CMWA	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31				
EVGR	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31				
GCKI	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31				
HAFL	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31				
MOWA	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31				
PHVI	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31				
PISI	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31
RBGR	>31	>31	>31	29	>31	>31	>31	>31	>31	>31	>31	>31				
RBNU	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31				
VATH	>31	>31	>31	29	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31
WEWP	>31	>31	>31	26	>31	>31	>31	30	>31	>31	>31	>31	>31	>31	>31	30
WWCR	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31	>31

Table 5.2. Number of species expected to meet a target SE of <2%/yr for each NWA, monitoring schedule, site spacing and duration of monitoring.

Edehzie (Horn Plateau; 48 species)

Spacing	Duration	All sites surveyed every					Fraction of sites surveyed each year			
		Year	2 years	3 years	4 years	5 years	1/2	1/3	1/4	1/5
30x30	11yr	5	2	1	1	0	4	2	1	1
	16yr	23	12	6	6	2	12	10	5	5
	21yr	30	24	20	13	8	27	21	19	12
	31yr	40	32	29	27	27	36	32	30	30
20x20	11yr	10	3	2	1	0	6	5	4	2
	16yr	26	21	9	7	6	25	20	10	10
	21yr	34	28	25	20	14	30	27	27	24
	31yr	44	38	33	31	30	41	40	35	35
15x15	11yr	14	5	3	1	1	9	5	5	5
	16yr	28	21	12	11	6	25	24	16	17
	21yr	36	30	26	22	19	33	30	28	28
	31yr	45	42	37	35	31	44	41	40	38
10x10	11yr	18	5	4	2	1	13	10	8	5
	16yr	31	22	15	14	8	27	27	24	25
	21yr	41	34	29	23	22	37	34	33	30
	31yr	47	44	41	38	36	45	45	44	44

Sambaa K'e (Trout Lake; 48 species)

Spacing	Duration	All sites surveyed every					Fraction of sites surveyed each year			
		Year	2 years	3 years	4 years	5 years	1/2	1/3	1/4	1/5
30x30	11yr	4	1	0	0	0	2	1	0	0
	16yr	15	7	3	2	1	6	4	3	2
	21yr	25	16	15	9	5	19	13	9	7
	31yr	35	29	26	25	18	32	27	25	21
20x20	11yr	9	2	1	1	0	6	3	2	2
	16yr	25	16	6	7	3	16	14	8	7
	21yr	32	25	25	16	9	27	24	23	18
	31yr	41	33	32	28	27	39	33	32	31
15x15	11yr	11	3	1	2	0	7	6	2	2
	16yr	26	20	8	7	3	25	21	9	10
	21yr	35	26	25	20	14	30	26	26	23
	31yr	44	37	35	31	29	40	39	34	33
10x10	11yr	17	4	3	3	1	10	8	5	5
	16yr	28	21	11	11	6	26	26	18	18
	21yr	37	29	26	24	18	34	31	29	27
	31yr	47	43	40	37	34	45	43	43	40

Table 5.2. Continued.

Ka'a'gee Tu (Kakisa; 48 species)

Spacing	Duration	All sites surveyed every					Fraction of sites surveyed each year			
		Year	2 years	3 years	4 years	5 years	1/2	1/3	1/4	1/5
30x30	11yr	3	0	0	0	0	1	0	0	0
	16yr	13	5	3	2	0	5	5	2	1
	21yr	24	13	9	5	3	15	8	8	5
	31yr	32	26	23	21	16	30	24	22	18
20x20	11yr	9	2	1	1	0	4	2	2	1
	16yr	24	12	7	4	3	11	12	5	5
	21yr	29	24	21	15	9	27	22	20	15
	31yr	40	33	28	27	27	38	32	30	31
15x15	11yr	10	3	2	1	0	5	4	2	2
	16yr	25	17	8	6	3	20	15	9	9
	21yr	31	27	25	18	14	28	27	25	23
	31yr	44	38	32	30	28	39	34	33	33
10x10	11yr	16	4	2	2	0	10	7	5	4
	16yr	27	22	13	10	6	25	24	16	17
	21yr	36	30	27	22	19	35	28	28	27
	31yr	45	42	37	35	31	44	42	40	39

Ts'ude niline Tu'eyeta (Ramparts; 33 species)

Spacing	Duration	All sites surveyed every					Fraction of sites surveyed each year			
		Year	2 years	3 years	4 years	5 years	1/2	1/3	1/4	1/5
30x30	11yr	6	1	0	0	0	1	1	0	0
	16yr	12	8	3	4	1	8	7	3	3
	21yr	17	14	12	7	6	15	10	10	9
	31yr	24	20	18	17	15	21	18	17	16
20x20	11yr	6	3	1	1	0	5	4	1	1
	16yr	14	10	5	5	3	12	8	7	7
	21yr	21	19	16	12	7	17	17	15	14
	31yr	29	23	21	21	18	26	23	22	20
15x15	11yr	8	3	3	2	0	7	6	3	2
	16yr	20	13	6	6	3	16	13	9	11
	21yr	25	21	20	14	10	22	20	18	17
	31yr	33	29	27	22	21	29	28	25	23
10x10	11yr	11	3	3	2	1	8	7	6	5
	16yr	24	16	10	10	3	21	18	13	13
	21yr	30	24	22	19	14	28	23	22	21
	31yr	33	31	30	28	25	33	33	30	29

Acknowledgements

Funding was provided by Environment Canada. This report used resources provided by the Boreal Avian Modelling (BAM) Project, an international research collaboration for the ecology, management and conservation of boreal birds. We acknowledge the BAM Project funding and data partners, and Steering and Technical Committee members who helped make this project possible (for complete list, see <http://www.borealbirds.ca/index.php/acknowledgements>).

References

- Humbert, J.-Y., L.S. Mills, J.S. Horne, and B. Dennis. 2009. A better way to estimate population trends. *Oikos* 118:1940-1946.
- Link, W.A. and J.R. Sauer. 1996. Extremes in ecology: avoiding the misleading effects of sampling variation in summary analyses. *Ecology* 77:1633-1640.
- Thomas, L. and K. Martin. 1996. The importance of analysis method for Breeding Bird Survey population trend estimates. *Conservation Biology* 10:479-490.

Appendix 1. Details of the analysis to estimate variance components

A1.1 Species and sites used

The variance analyses were done using a Poisson models for each of the 51 species. Only the main sites from the Fort Liard area were used, not the seismic study, to allow consistent years of records for most sites. The main sites include 82% of the records from this area. For BBS, sites in the boreal forest of Saskatchewan, Alberta, BC, Yukon and NWT, or in the BC Interior, were selected. Only years 1997-2011 were used, as there were few routes prior to that, and only routes with at least 6 years of sampling within that period were selected. The routes also had to have at least one survey in 1997 or 1998 and one in 2010 and 2011, to ensure that surveys covered the entire period, and to reduce confounding of routes and year. Because the interest was only in the yearly components of variation, the BBS data were rolled up to route-level counts.

A1.2 Proportion of sites with no detections in any year (“0 sites”)

General linear mixed models (GLMM) were used to partition the variance in the Fort Liard and BBS datasets. For most of the rarer species, there were sites in the Fort Liard data and routes in the BBS where the species never occurred in any year of sampling. These zero sites are a challenge for log-link GLMMs, often greatly exaggerating yearly or site variances. The proportion of zero sites was therefore calculated, then these sites were excluded from the GLMM for that species. (Note that this does not exclude sites where the species did not occur in only some years, or even in all but one year.) The GLMM results therefore apply to sites where the species is present at least occasionally. The 0 sites are included in the simulations simply as the proportion of sites where the species never occurs, with binomial error around the proportion.

A1.3 Quadratic trend removed

The 14-year study span with 8 sample years in the Fort Liard data, and the 15 years in the BBS data support using quadratic trends, for the purpose of estimating how much yearly variation there is around the trend line⁴. All but 4 of the 23 most common species at Fort Liard had more support for a

⁴ This doesn't mean that quadratic trends need to be simulated in the precision analysis, because the issue there is estimating precision around the “true” trend, which can be assumed to be linear for simplicity in those simulations. Quadratic trends or moderately-flexible splines should be considered when actually analysing the Fort Liard data.

quadratic than a linear trend over that period. Using only a linear trend in the variance-components model would over-estimate the amount of year-to-year variation. cursory examination showed up to an order magnitude greater yearly variation if a linear trend was assumed for the many species that show apparent quadratic trends.

A1.4 Refining the GLMM model

The initial model I tried for Fort Liard was:

$$\text{FL Abundance} \sim \text{Yr} + \text{I}(\text{Yr}^2) + \text{Visit} + (1|\text{Yr}) + (\text{Yr}|\text{Site}) + (1|\text{Station:Site})$$

That is, abundance is modeled as a quadratic trend plus a fixed effect of visit (2 visits per year), with random variation in the yearly mean, the mean and linear yearly trend within site and for stations within sites. The initial model for BBS was:

$$\text{BBS Abundance} \sim \text{Yr} + \text{I}(\text{Yr}^2) + (1|\text{Yr}) + (\text{Yr}|\text{Route})$$

There is only one BBS visit per year, and stops were summed within routes.

However, there were two problems with applying these models to the available data, which led to modified models.

Problem 1: The estimates of the site*year variance (the difference in trends among sites) for different species in the Fort Liard data were highly erratic, and affected estimates of other variance components. I think this is because, for most species in this dataset, there are many sites with 0 values in particular years, which caused overwhelming apparent variation in trend among sites on the log- or logit-scales. The Fort Liard GLMM was therefore run without that term:

$$\text{FL Abundance} \sim \text{Yr} + \text{I}(\text{Yr}^2) + \text{Visit} + (1|\text{Yr}) + (1|\text{Station/Site})$$

Note that the (1|Station:Site) term became (1|Station/Site) so that the site intercept term is still estimated. The site*year variance component comes only from the BBS dataset, where the large number of stops/route and more years of sampling reduce the problem.

Problem 2: A second issue was a general problem of GLMM analyses when there are relatively few levels of one variable that is confounded with another variable that has many levels. In this case, year has few levels and site or route have many. The two variables are somewhat confounded, because not all sites or routes were surveyed each year. This problem is mainly in the first year with the Fort Liard data, as some of the otherwise consistently measured sites were not started until the second year. The problem is much larger with BBS, where few routes were surveyed in all years (despite the selection criteria mentioned above). When there are many levels of one variable and few of a second confounded

a. Fort Liard

b. BBS

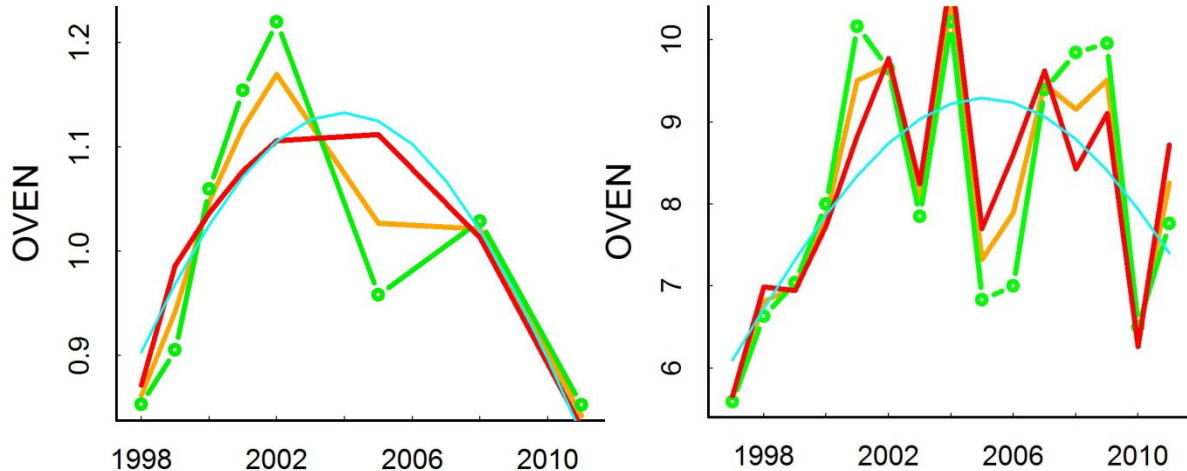


Figure A1.1. Illustration of the allocation of annual variability to site or route variability, particularly for the BBS dataset. Green points are the observed mean count of ovenbirds each year (for sites or routes that had at least one record; per station for Fort Liard, per route for BBS). Thin blue line is the simple GLM quadratic trend. Red line is the yearly fitted values for a GLMM with a random-effects site or route term but no year term. Orange line adds a random-effects term for year. For Fort Liard, the site term modifies the quadratic trend slightly, mainly by allowing a close fit to the first year (when an incomplete subset of sites were surveyed). For BBS, the more erratic surveying of routes allows the GLMM with route to follow the observed annual variations much more closely. In the Fort Liard example, the year term is substantial, moving the orange curve towards the more extreme yearly observations, while in the BBS example, the year term can explain little that is not already covered by the route term. The difference between the orange and green lines is attributed to Poisson count error in the GLMM.

variable, GLMMs tend to assign all the shared variation to the variable with many levels. In this case, that would result in overestimating the BBS route variance at the expense of underestimating the yearly variation. The issue is illustrated with ovenbirds in Figure A1.1. Ovenbirds are common and well-surveyed; the problem is more extreme with rare species.

I decided to run the Fort Liard and BBS models as stated above, including the (Year|Route) term in the BBS model as the only available estimate of year*site variation. However, to deal with this problem for the BBS yearly variance, I used the variance of the mean fitted estimates for each year (orange line in Figure A1.1b) around the simple quadratic trend (blue line in Figure A1.1b), without the additional adjustment for route. This is effectively ascribing all the extra-Poisson variation in yearly means in the BBS data to true yearly variation, and none of that variation to the fact that a different subset of routes was sampled in different years. The selection of routes that were sampled in at least 6 years and across the time span was the only way that route*year confounding was “controlled”.

A1.5 No observer random effect in Fort Liard model

The GLMM model does not include a random effect for observer. Variation among observers and among years is almost completely confounded in the Fort Liard design, because almost all observers surveyed only one year (Table A1.1). This means that there is almost no way of telling whether variation among observers is because they were inherently different or because they were sampling in years that were different. A random-effects term for observer reduces much of the year-to-year variation around the quadratic trend (black versus red line in Figure A1.2), and this is essentially the same variation that a

random-effects term for year removes (red versus green line in Figure A1.2). For Swainson’s thrush, 80.6% of the residual variation after the quadratic relationship is shared by observer and year, 14.2% is purely attributable to observer and 5.2% purely to year. That shared 80.6% means that observer variation and residual yearly variation are very confounded here.

Table A1.1. Stations surveyed by each observer in different years, Fort Liard dataset.

Observer	1998	1999	2000	2001	2002	2005	2008	2011
Ob3	133	101	64	82	71			39
Ob8	125	63						
Ob1		103						
Ob2		17						
Ob7		69						
Ob12		1						
Ob11			84			110		
Ob4			90					
Ob5			116	95	88			
Ob10				76				
Ob6				101				
Ob16					103			
Ob17					92			
Ob19						128		
Ob18						110	103	
Ob20							48	
Ob21							96	
Ob22							77	
Ob23							21	
Ob24								84
Ob25								69
Ob27								131

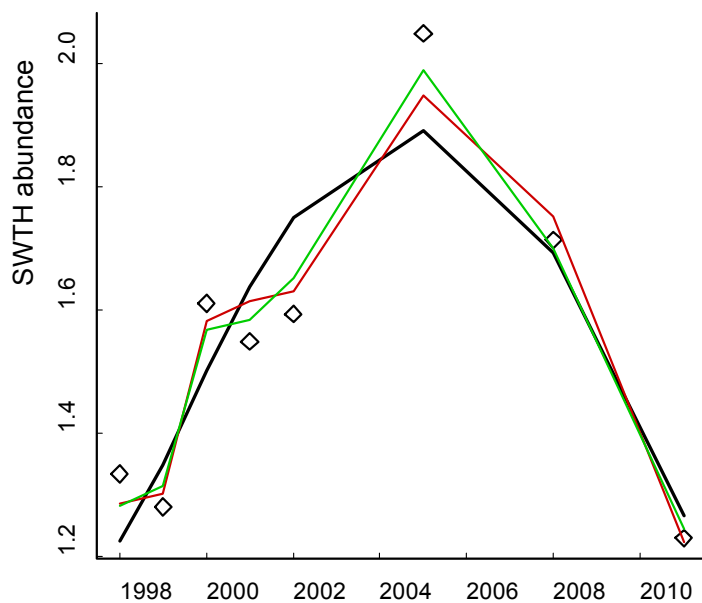


Figure A1.2. Example showing how confounded observer variation and year-to-year variation are with the Fort Liard dataset, for Swainson's thrush. Points are annual mean counts. Black line is quadratic model (with random effects for site and station, which only vary slightly from year-to-year). Red line is with an additional random-effects term for observer. Green line is with an additional random-effects term for year (i.e., year-to-year variation, given the quadratic trend).

Despite the confounding, an argument can be made that annual variation is the main driver here: the mean number of birds per plot for the one observer who did surveys in 6 years in this main part of the study was highly correlated ($r=0.92$) with the mean number of birds per plot of all the other observers across those 6 years. That supports the view that the variation is real annual variation – in the number of birds or their detectability – and that much of the variation among observers was due to their surveying in different years. (Note that this doesn't preclude an additional observer effect, or that a model with an observer term would get a lower (better) AIC score than one without the term).

A strong practical argument for not removing observer variation from annual variation here is that the surveys in a given NWA will most likely be several years apart, so that it is unlikely to have the same observers for more than one survey per NWA in any case. Hence, any observer variation will be compounded with temporal variation anyway.

A1.6 Additional dataset in NWA areas

Additional datasets have been collected in four of the proposed protected areas. These are helpful for providing local estimates of the mean abundance of the species, which is a main determinant of expected precision. Because the 3 southern NWAs are in the same general region and have only a single year of sampling, species abundance were averaged across these areas.

The datasets from the NWA areas were also examined for the possibility of providing an additional estimate of site variance, specific to the local area. With a criteria of 20+ observations to run the simple Poisson variance model, there were relatively few species that could be analysed in each area, and with few exceptions the results all showed estimates of no site variance. The reason is that with only one year of data at 3 stations, it is very unlikely to have more variation between sites than can be accommodated by Poisson error. The eight species-area combinations that could be analyzed and showed non-0 site variance had estimates in the general ballpark of the results of the Fort Liard analysis (0.027 – 0.424). As a result, I decided to use the area-specific means for the species from this dataset, but not the site variances.

A1.7 Combining BBS and Fort Liard yearly variances

Yearly variances were available from the Fort Liard and BBS analyses. They were combined in a precision-weighted mean for the final estimate. GLMM analyses do not provide an estimate of the precision of the variance components⁵. Precision was therefore assumed to be proportional to the

⁵ The two generic methods for establishing precision non-parameterically, the bootstrap and jackknife, do not work here. Bootstrapping is generally inappropriate for variance estimates, because it involves repeating some samples, hence reducing the variance. Jackknifing, as I tried it, produced negative variance estimates, because estimates are sensitive to some individual sites. The jackknife problem might be solvable, or there could be other options, but I did not pursue these for this purpose.

mean number of detections per year. The basis for this is that Poisson variance is equal to the mean, and inverse-variance weighted means are a standard Bayesian way of combining two uncertain estimates. For example, if the Fort Liard data averaged 100 detections of a species per year and the BBS 200 per year, the weights for that species would be 1/3 for the Fort Liard estimate and 2/3 for the BBS estimate.

This “yearly-detections-weighted” mean is important for dealing with another issue in the GLMM analysis. When species are rare, Poisson error becomes increasingly large, relatively. As a result, estimates of yearly variation become increasingly erratic, because the annual variation in the mean is swamped by the Poisson error in the counts. The pattern is that many rare species are assigned a yearly variation at or very near 0, while some are given high annual variation (Figure A1.3). This is not biological reality, it is simply a reflection of the luck-of-the-draw – whether the dominant random count variation happened to remain within the feasible range of Poisson variation, or whether it exceeded it, so that the model had to assign (high) variability to the year term. The weighting reduces the problem because few species are very uncommon in both the Fort Liard and BBS data.

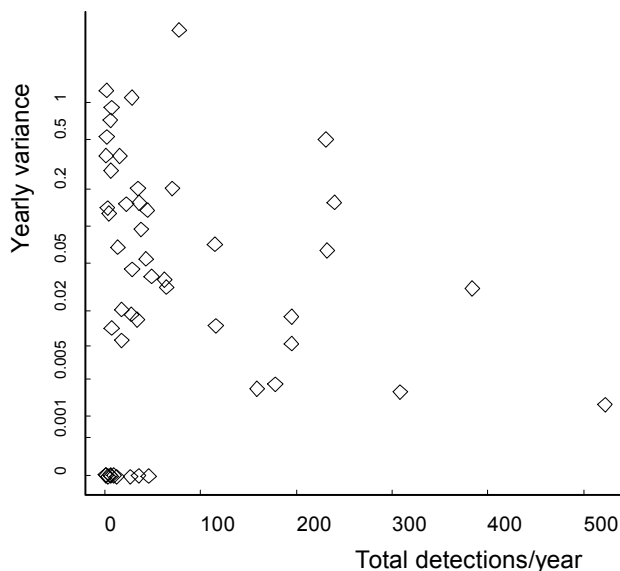


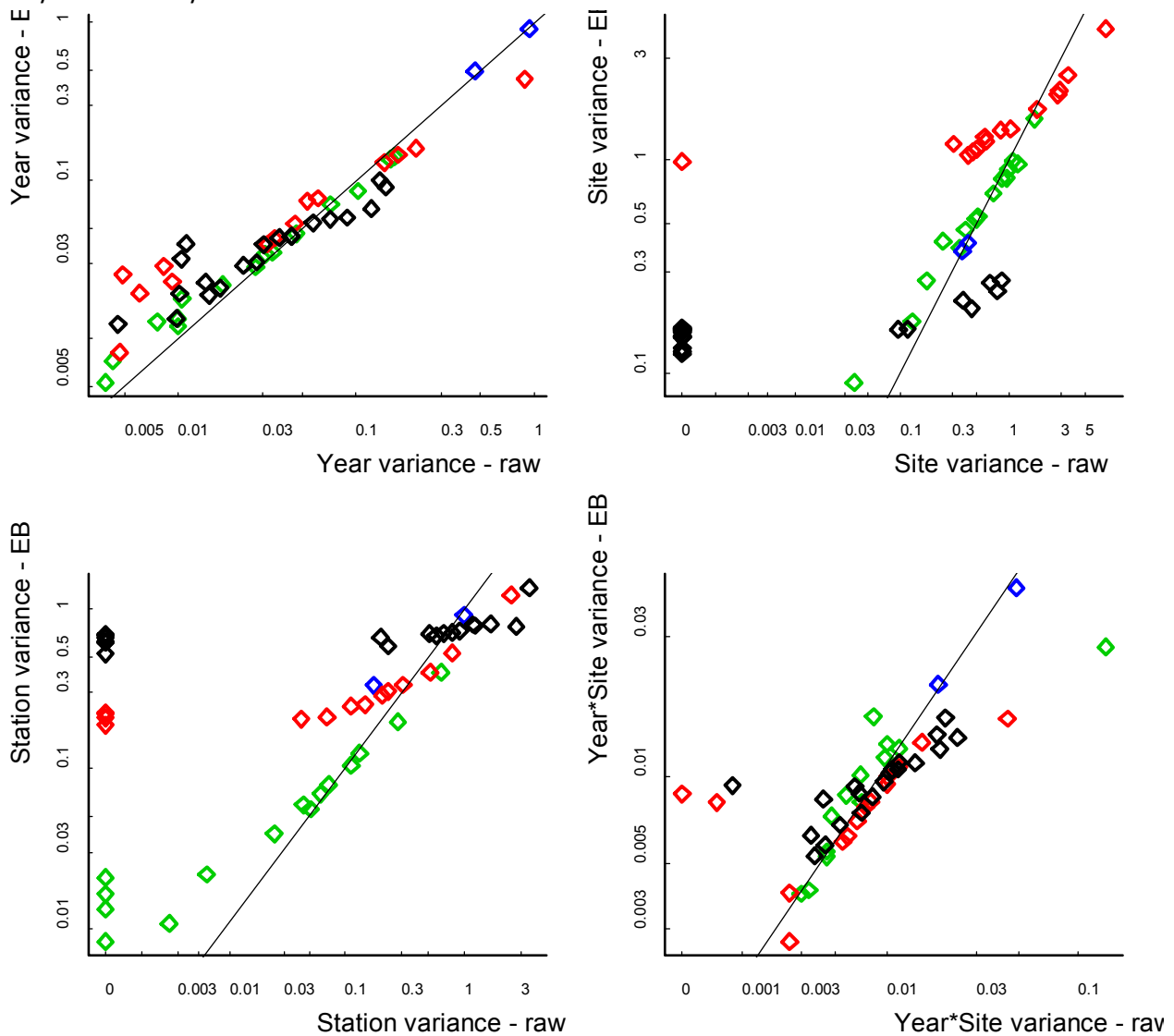
Figure A1.3. Yearly variance estimates from the Fort Liard data versus total detections/year of the species. The spread in the estimate of variance components is higher when total detections are low, because Poisson measurement error becomes dominant, and the GLMM can either “find” high level of extra variation or none.

A1.8 Empirical Bayes smoothing of species’ estimates in groups

An additional step to reduce stochasticity in the species’ estimates was to use an empirical Bayes average, in which species were assigned a weighted mean of their specific value and the mean of values for a group of similar species. The weighting would ideally be based on the uncertainty of the specific estimates, which are not known here, but was again approximated with the total detections. An example of what this would look like is given in Figure 4. I used 4 groups: 1) “Rare” = species with <20 detections/yr in the Fort Liard data, 2) “Moderate” = species with 20-50 detections, 3) “Common” = species with >50 detections, 4) 2 irruptive species (pine siskin and white-winged crossbill). Bohemian waxwing might be included in this group, but it had so few detections that I left it in group 1, the rare

species). I assumed that a species with 50 detections/year⁶ would have the same uncertainty in its estimate as does the mean for the species group. That is, with 50 detections/yr, the species' specific estimate and its group mean have the same weight in the empirical Bayes estimate; with fewer detections, the group mean has more weight, and with more detections the group mean has less weight.

In this example, site and station variance are most effected, because they are based on the Fort Liard data only, and hence the total detections are less – the groups' means are therefore weighted relatively more. The rare species are strongly affected, while the common species are fairly unaffected, which follows directly from how the procedure was set up. This is a typical finding with empirical Bayes estimates, and one of the main reasons for its use. Outliers tend to be moved more towards the mean than less extreme values. Usefully, the 0 estimates for some values – where the GLMM could not “find” any variation beyond the dominant Poisson error – are all moved to non-0 values.



⁶ The detections/yr is BBS+Fort Liard for year variance, Fort Liard only for site and station, and BBS only for site*year, because those were the datasets used to generate those estimates.

Figure 4. An example of empirical Bayes (EB) adjustments of species' parameters, using simple groups of species (black=rare, red=moderate, green=common, blue=irruptive). Points are the EB value plotted against the original value for each species. Note the logarithmic scaling on both axes. Line is 1:1 (no change). The mean for a group is where its points cross the 1:1 line.

Appendix 2. Sensitivity of Trend Precision to Variance Estimates

As an initial step to know how much effort was needed in deriving variance components for the monitoring simulations, I assessed how sensitive precision estimates for trend are to the four variance components – year, site, station and year*site – by running simulations where each of these was set to its mean estimate, 1/3 the mean and 3 times the mean. These were done individually (i.e., there were 9 scenarios, 1 mean + 2 altered levels of each component, not all 3x3x3x3 combinations). For each of the 9 sets of variance levels, I did simulations for the 24 combinations of: 3 species (Swainson's thrush – common; magnolia warbler – moderately common; and red-breasted nuthatch – moderate abundance), 2 durations (11 and 21 years), 2 numbers of sites (45 and 161, representing the 20x20 and 10x10 spacing in the largest NWA), and 2 revisit designs (all sites done every year, and all sites done every 5th year).

Results

Expected SE's of trend are in Table A2.1 for the combinations of simulation factors and variance levels. The results of altering the variance components are most easily seen in the "x3:x1/3" column, which shows the ratio of the expected SE of the trend with 3x the mean variance component to the expected SE with 1/3x the mean variance component, for each combination of simulation factors.

The main finding in terms of sensitivity to variance levels is that the yearly variance is important for trend precision, while the other components are much less so, or not important at all. The 9-fold variation in the level of yearly variance changed the expected SE by 1.6 to 2.8 times. The effect was consistently higher with more sites, and it also tended to be higher with the 5-yr revisit period rather than the yearly revisits. There was no real difference in the effect for the 3 species.

The lack of effect of the site and station variance make sense in retrospect. Because all sites are visited in a sampling year and the Poisson simulation and analysis is done on the log scale, increased site variance varies the intercept for different sites, but it does not change their trend, so the overall trend across all sites is not really affected by the site- or station-level variance. Added site or station variance does affect the Poisson measurement error, but this includes increased relative variability when counts are low and decreased relative variability when counts are high (i.e., a Poisson random number from a low-mean site has relatively more variability than from a high-mean site). These effects seem to more-or-less cancel out across the reasonable number of sites used here. There could be greater effects when there are few sites with rare species, or where simulated strong declines make species rare, when extra site variance might exacerbate the already dominant Poisson error effects. Site variance should also have a bit more effect with a panel design, where only some sites are measured each year. In that case, greater site variability would cause somewhat greater year-to-year variability.

The limited effect of changes in the year*site variance component was a bit surprising. However, it seems that even 45 sites is enough for the upward drift of some sites relative to the overall trend to cancel out the downward drift of others.

The results in Table A2.1 also show the strong effect of duration of sampling, the additional important effect of revisit schedule (the yearly revisits here represent 5 times the effort compared to 5-yr revisits), the moderate effects of number of sites, and the substantial differences in expected precision between the common, moderately common and moderate-abundance species.

The level of accuracy of yearly variance is therefore important to the results, while the precise values of the other variance components are relatively less important (at least at the levels of the species used here). Overall abundance of the species is also an important parameter.

Table A2.1: Expected SE of trends for 3 species under different combinations of monitoring duration, number of sites, revisit schedule and levels of the variance components ($x_{1/3} = 1/3 \times$ the mean, $x_3 = 3 \times$ the mean, $x_3:x_{1/3}$ is the ratio of the SE with 3x the variance to the SE with 1/3 the variance).

Sp	Duration (yr)	Sites	Revisit (yr)	Year variance				Site variance				Station variance				Year*Site Variance			
				$x_{1/3}$	Mean	x_3	$x_3:x_{1/3}$	$x_{1/3}$	Mean	x_3	$x_3:x_{1/3}$	$x_{1/3}$	Mean	x_3	$x_3:x_{1/3}$	$x_{1/3}$	Mean	x_3	$x_3:x_{1/3}$
SWTH	11	45	1	0.0064	0.0072	0.0109	1.7044	0.0070	0.0072	0.0073	1.0498	0.0074	0.0072	0.0074	0.9910	0.0073	0.0072	0.0076	1.0432
			5	0.0167	0.0223	0.0334	1.9970	0.0220	0.0223	0.0218	0.9915	0.0214	0.0223	0.0226	1.0574	0.0208	0.0223	0.0235	1.1335
		161	1	0.0043	0.0056	0.0097	2.2585	0.0058	0.0056	0.0059	1.0185	0.0061	0.0056	0.0065	1.0669	0.0058	0.0056	0.0058	0.9901
			5	0.0118	0.0189	0.0306	2.6061	0.0194	0.0189	0.0191	0.9871	0.0203	0.0189	0.0188	0.9298	0.0166	0.0189	0.0188	1.1316
	21	45	1	0.0025	0.0025	0.0039	1.5809	0.0030	0.0025	0.0026	0.8749	0.0028	0.0025	0.0030	1.0695	0.0027	0.0025	0.0027	1.0245
			5	0.0059	0.0081	0.0128	2.1538	0.0088	0.0081	0.0076	0.8621	0.0081	0.0081	0.0098	1.2086	0.0081	0.0081	0.0083	1.0215
		161	1	0.0015	0.0019	0.0035	2.2934	0.0020	0.0019	0.0020	0.9557	0.0021	0.0019	0.0023	1.1165	0.0022	0.0019	0.0020	0.9081
			5	0.0045	0.0066	0.0114	2.5105	0.0079	0.0066	0.0071	0.9004	0.0063	0.0066	0.0083	1.3062	0.0067	0.0066	0.0066	0.9872
MAWA	11	45	1	0.0145	0.0232	0.0330	2.2743	0.0213	0.0232	0.0203	0.9520	0.0206	0.0232	0.0231	1.1244	0.0208	0.0232	0.0227	1.0910
			5	0.0368	0.0556	0.0761	2.0660	0.0498	0.0556	0.0476	0.9548	0.0458	0.0556	0.0511	1.1156	0.0477	0.0556	0.0539	1.1301
		161	1	0.0114	0.0208	0.0317	2.7795	0.0185	0.0208	0.0184	0.9911	0.0176	0.0208	0.0218	1.2411	0.0180	0.0208	0.0202	1.1239
			5	0.0271	0.0504	0.0726	2.6756	0.0463	0.0504	0.0410	0.8850	0.0414	0.0504	0.0472	1.1389	0.0418	0.0504	0.0451	1.0776
	21	45	1	0.0054	0.0081	0.0120	2.2272	0.0071	0.0081	0.0066	0.9218	0.0080	0.0081	0.0084	1.0428	0.0086	0.0081	0.0099	1.1537
			5	0.0128	0.0186	0.0265	2.0793	0.0186	0.0186	0.0187	1.0083	0.0175	0.0186	0.0203	1.1580	0.0192	0.0186	0.0202	1.0540
		161	1	0.0042	0.0071	0.0117	2.7682	0.0063	0.0071	0.0060	0.9518	0.0073	0.0071	0.0074	1.0076	0.0071	0.0071	0.0089	1.2666
			5	0.0107	0.0161	0.0256	2.3943	0.0161	0.0161	0.0165	1.0231	0.0152	0.0161	0.0183	1.2012	0.0163	0.0161	0.0175	1.0769
RBNU	11	45	1	0.0486	0.0649	0.0884	1.8170	0.0634	0.0649	0.0615	0.9706	0.0602	0.0649	0.0585	0.9708	0.0654	0.0649	0.0581	0.8883
			5	0.1335	0.1412	0.2234	1.6737	0.1338	0.1412	0.1511	1.1292	0.1429	0.1412	0.1906	1.3334	0.1818	0.1412	0.1129	0.6211
		161	1	0.0356	0.0550	0.0859	2.4127	0.0530	0.0550	0.0529	0.9975	0.0517	0.0550	0.0523	1.0130	0.0561	0.0550	0.0501	0.8929
			5	0.0691	0.1011	0.1911	2.7678	0.1043	0.1011	0.1020	0.9777	0.1095	0.1011	0.1139	1.0399	0.1130	0.1011	0.0992	0.8785
	21	45	1	0.0208	0.0235	0.0386	1.8579	0.0248	0.0235	0.0261	1.0524	0.0253	0.0235	0.0238	0.9416	0.0256	0.0235	0.0255	0.9976
			5	0.0439	0.0473	0.0768	1.7487	0.0483	0.0473	0.0475	0.9825	0.0560	0.0473	0.0513	0.9175	0.0493	0.0473	0.0501	1.0159
		161	1	0.0146	0.0205	0.0376	2.5692	0.0226	0.0205	0.0223	0.9854	0.0229	0.0205	0.0239	1.0435	0.0235	0.0205	0.0215	0.9134
			5	0.0263	0.0325	0.0677	2.5711	0.0372	0.0325	0.0392	1.0527	0.0440	0.0325	0.0430	0.9766	0.0397	0.0325	0.0403	1.0153