APPENDIX 1. Detailed description of the maximum likelihood (ML) model used to generate Canadian Breeding Bird Survey trends before 2010.

Note: We have supplied this document as an appendix because a full description of the ML model was never published. This description of the ML model is modified from an unpublished description of the custom computer program that calculates the ML model (Collins, B. T. unpublished, "Estimating an annual index for the breeding bird survey using a maximum likelihood model").

Model of Annual Index

Let c_{ij} denote the observed count on route i in year j Since not all routes are run in all years, u_{ij} is an indicator variable which shows whether the observations were taken, i.e.,

$u_{ij} = 1$ if an observation was made on route i in year j, otherwise

$u_{ij} = 0$

Observations made under some conditions can not be included in the analysis. The data from each BBS route is partitioned into blocks of observations taken under comparable conditions, meaning by the same observer and at the same time of year (within a 19-day window of observations in other years). There must be at least one sighting of the species in a block in order to include the data from that block. In addition, if a species is never sighted on any route in a given year, that year cannot be included in the model and the annual index for that year is set to zero. To account for the combined effects of these factors that may exclude some observations, v_{ij} is an indicator variable which shows whether the observation can be used in the analysis, i.e.,

$v_{ij} = 1$if an observation made on route *i* in year *j* can be used for the analysis, otherwise $v_{ij} = 0$

The observed counts are assumed to follow a Poisson distribution with mean λ_{ij}, which is modeled as: $\lambda_{ij} = \exp(\mu + \alpha_i + \beta_j + \gamma_{i(k)})$,

where μ denotes the overall average, α_i denotes the effect of route *i*, β_j denotes the effect of year *j*, and $\gamma_{i(k)}$ denotes the effect of the observation conditions in year i, within observation conditions block k. This effect of observation conditions includes the effects of observers and date, which are the factors used to partition the data into comparable blocks. Sampling is selected based on. The likelihood function is weighted by the area degree blocks stratum area:

$$L = \prod_h \left\{ \prod_{i \in h} \prod_j \left[\frac{\lambda_{ij}^k \exp(-\lambda_{ij})}{c_{ij}} \right]^{u_{ij}} \right\}^{w_h}$$

where

$w_{i \in h} = w_h = \frac{A_h}{n_h}$ denotes the weighting term for all routes in degree block *h*, A_h denotes the area of degree block *h*, and n_h denotes the number of routes in degree block *h*.

Multiplying the probability terms together and taking logarithms gives the log-likelihood equation:

$$L = \sum_i^n w_i \sum_j^m v_{ij} \{c_{ij}(\mu + \alpha_i + \beta_j + \gamma_{i(k)}) - \exp(\mu + \alpha_i + \beta_j + \gamma_{i(k)}) - \ln(c_{ij})\} \quad (1)$$

The maximum likelihood estimates of the parameters are the solution to the following set of equations:

$$L = \sum_i^n w_i \sum_j^m v_{ij} \{c_{ij} - \exp(\mu + \alpha_i + \beta_j + \gamma_{i(k)})\} = 0 \quad (2)$$

$$L = \sum_j^m v_{ij} \{c_{ij} - \exp(\mu + \alpha_i + \beta_j + \gamma_{i(k)})\} = 0 \tag{3}$$

$$L = \sum_i^n w_i v_{ij} \{c_{ij} - \exp(\mu + \alpha_i + \beta_j + \gamma_{i(k)})\} = 0 \tag{4}$$

$$L = \sum_{j \in k}^m v_{ij} \{c_{ij} - \exp(\mu + \alpha_i + \beta_j + \gamma_{i(k)})\} = 0 \tag{5}$$

where the summation in equation (5) extends over the years with observations in conditions block *k*.

Estimating the Model Parameters

The maximum likelihood estimators are created through an iterative scheme, because they cannot be written in closed form.

Initial estimates are set to $\hat{\alpha}_i^{(0)} = 0$, $\hat{\beta}_j^{(0)} = 0$, $\hat{\gamma}_{i(k)}^{(0)} = 0$ and

$$\hat{\mu}^{(0)} = \frac{\sum_i^n w_i \sum_j^m v_{ij} c_{ij}}{\sum_i^n w_i \sum_j^m v_{ij}} \text{ (i.e., the area weighted average of all observations)} \tag{6}$$

Given the estimates at step *g* ($\alpha_i^{(g)}, \beta_j^{(g)}, \gamma_{i(k)}^{(g)}$), the estimates for the next step are calculated as:

$$\sum_j^m v_{ij} \left\{c_{ij} - \exp\left(\hat{\mu}^{(0)} + \hat{\alpha}_i^{(g+1)} + \hat{\beta}_j^{(g)} + \hat{\gamma}_{i(k)}^{(g)}\right)\right\} = 0 \tag{7}$$

$$\sum_j^m w_i v_{ij} \left\{c_{ij} - \exp\left(\hat{\mu}^{(0)} + \hat{\alpha}_i^{(g+1)} + \hat{\beta}_j^{(g+1)} + \hat{\gamma}_{i(k)}^{(g)}\right)\right\} = 0 \tag{8}$$

$$\sum_j^m v_{ij} \left\{c_{ij} - \exp\left(\hat{\mu}^{(0)} + \hat{\alpha}_i^{(g+1)} + \hat{\beta}_j^{(g+1)} + \hat{\gamma}_{i(k)}^{(g+1)}\right)\right\} = 0 \tag{9}$$

The left hand sides of equations (7)-(9) each have only one unknown variable and are monotone decreasing functions. The solution is derived by initially calculating the equation at the current level of the parameter and then stepping up or down by one until the solution is bounded. The bounded solution is then refined using a binary search.

When the iteration is completed, the change in the parameter estimates is calculated as:

$$C^{(g)} = \sum_{i}^{n} (\alpha_i^{(g)} - \alpha_i^{(g+1)})^2 + \sum_{j}^{m} (\beta_j^{(g)} - \beta_j^{(g+1)})^2 + \sum_{i}^{n} \sum_{k} (\gamma_{i(k)}^{(g)} - \gamma_{i(k)}^{(g+1)})^2$$

The iterations continue until $C^{(g)} < 0.00001$.

In the above algorithm, the estimate $\hat{\mu}^{(0)}$ is never updated by the iteration steps, because it is a redundant parameter that can be set arbitrarily. When the iterations have converged, the estimates are adjusted to fit the constraints: $\sum_{k} \hat{\gamma}_{i(k)} = 0$ for all *i*, $\sum_{i}^{n} \hat{\alpha}_i = 0$ for all i, and $\sum_{j}^{m} \beta_j = 0$ as follows. If *G* denotes the last iteration, then the conditions block effects are adjusted to sum to zero within each route as follows:

$$\bar{\gamma}_i^{(G)} = \sum_{k} \hat{\gamma}_{i(k)}^{(G)} / K_i \qquad \hat{\gamma}_{i(k)} = \hat{\gamma}_{i(k)}^{G} - \bar{\gamma}_i^{(G)} \qquad \hat{\alpha}_i^{(G+1)} = \hat{\alpha}_i^{(G)} + \bar{\gamma}_i^{(G)}$$

The block effect are then adjusted to that they each sum to zero through the following calculations:

$$\bar{\beta}_j^{(G)} = \sum_{j}^{m} \hat{\beta}_j^{(G)} / m \qquad \hat{\beta}_j = \hat{\beta}_j^{(G)} - \bar{\beta}_j^{(G)} \qquad \hat{\mu} = \hat{\mu}^{(0)} + \bar{\alpha}^{(G+1)} + \bar{\beta}^{(G)}$$

The adjustments retain the same predicted count for every observation.

Annual Indices and Trend

The annual population index (I_j) is an estimate of the count that would have been observed in year *j* if all routes had been run under average conditions. The index is calculated as:

$$I_j = \left(\frac{\sum_i w_i v_{ij}}{\sum_i w_i u_{ij}}\right) \exp\,(\hat{\mu} + \hat{\beta}_j) \qquad (10)$$

The first term in brackets in equation (10) scales the index downward to reflect the fact that the model can only be applied to a portion of the data set. For years in which the species is never sighted, I_j is set to zero, which is the maximum likelihood estimate in this situation.

An estimate of trend in the ln scale is calculated as the slope of a linear regression of the $\ln(I_j + 0.001)$ on year.

Estimation of Standard Error

The standard errors of the annual index and trend are estimated through a jack-knife procedure. In this algorithm, pseudo-replicates are created through discarding one route at a time and recalculating the annual index and trend. The variability among the pseudo-replicates is used to estimate the variance of the annual index and trend. In the calculation implemented in the program, all routes (including those in which the species was never seen) are used in the calculation of the variance. Because the number of routes visited varies among years, the number of pseudo-replicates varies among the annual indices.