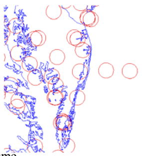
Hierarchical Models for the Analysis of Christmas Bird Count Data

William A. Link, John R. Sauer, and Daniel Niven
USGS Patuxent Wildlife Research Center, and National Audubon Society
12100 American Holly Drive, Laurel, MD 20708

Christmas Bird Count

- Started in 1900
- Birding activity
- Organized in "Circles"
  - 15 mi diameter areas
  - Surveyed by many people
- Not really a survey
  - Sample frame is limited
    - » Locations are not randomly selected
  - Detectability varies over space and time
    - » Counts often associated with effort
    - » Actual populations are never counted!

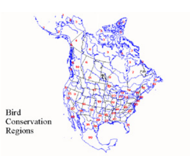


Conservation Interest in Using CBC to Estimate Population Change

- Some species can be counted only in winter
- Alternative data source for many species
  - "Check" other surveys
- Large amounts of data
  - Long time series
  - Geographically extensive
    - » Spatially explicit

Geographic Structuring: Bird Conservation Regions

- Common unit for conservation planning
  - Developed in conjunction with North American Bird Conservation Initiative
  - Ecoregion-based
  - Multistate/province



Stratum	Name	Stratum	Name
1	Aleutian/bering Sea Islands	20	Edwards Plateau
2	Western Alaska	21	Oaks and Prairies
3	Arctic Plains and Mountains	22	Eastern Tallgrass Prairie
4	Northwestern Interior Forest	23	Prairie Hardwood Transition
5	Northern Pacific Rainforest	24	Central Hardwoods
6	Boreal Taiga Plains	25	West Gulf Coastal Plain/ouachita
7	Taiga Shield and Hudson Plains	26	Mississippi Alluvial Valley
8	Boreal Softwood Shield	27	Southeastern Coastal Plain
9	Great Basin	28	Appalachian Mountains
10	Northern Rockies	29	Piedmont
11	Prairie Potholes	30	New England/mid-atlantic Coast
12	Boreal Hardwood Transition	31	Peninsular Florida
13	Lower Great Lakes/ St. Lawrence Plain	32	Coastal California
14	Atlantic Northern Forest	33	Sonoran and Mojave Deserts
15	Sierra Nevada	34	Sierra Madre Occidental
16	Southern Rockies/colorado Plateau	35	Chihuahuan Desert
17	Badlands and Prairies	36	Tamaulipan Brushlands
18	Shortgrass Prairie	37	Gulf Coastal Prairie
19	Central Mixed Grass Prairie		

Plan for Analysis

- Reclassify CBC circles into BCRs
  - GIS overlaying of circles on strata
  - Use them as strata for analysis
- Fit log-linear hierarchical model
  - » Flexible effort adjustment allows for different shapes of relationship
  - » Estimate trend, annual indices by BCR
  - » Additional effects
    - Circle
    - Stratum
    - Overdispersion

Problem is Hierarchical

- Want to summarize collections of estimates
  - Stratum (BCR) and circle effects
- Need model-based adjustments of counts for effort
  - Additional estimation
- Clumsy approaches
  - Route Regression
  - Generalized Linear Model + empirical Bayes (Link and Sauer 1999)
- Better approach:
  - Hierarchical models

New Method: Hierarchical Models

- Data and parameters both random variables
  - Data are observed
  - Distributions described conditionally on realized values of parameters that are also random variables
  - Need to include additional level(s) of distributional assumptions
- Analysis of hierarchical models
  - Bayesian methods are generally required for complicated models
  - Model-based

Defining a Hierarchical Model for CBC Analyses

- Model structure
  - Similar to log-linear model of Link and Sauer (1999) for CBC data
  - Requires additional distributional assumptions
    - » "Random Effects"
    - » Strata, year, and circle effects, overdispersion
    - » Effort Effects
  - Derived variables
    - » Annual Indexes
    - » Trend

Population Change and the CBC: Stratum-specific Effort Adjustments

Effort (Party Hours) is scaled relative to an overall mean

$$\xi_* = \frac{\xi_i}{\overline{\xi}}$$

p_i determines the shape of the effort adjustment in stratum i

$$B_i \frac{(\xi_*^{\,p_i} - 1)}{p_i}$$

$p_i < 0$: finite asymptote

$p_i < 1$: change in concavity over range of ξ

Population Change and the CBC: Regression Model

"i" = stratum, "j" = circle, "t" = year

$$\log(\lambda_{i,j,t}) = S_i + \beta_i (t - t^*) + \gamma_{i,t} + \omega_j + B_i \frac{(\xi_*^{p_i} - 1)}{p_i} + \varepsilon_{i,j,t}$$

- logarithm of expected count: $\log(\lambda_{i,j,t})$
- stratum-specific slope and intercept: $S_i + \beta_i (t - t^*)$
- year effects: $\gamma_{i,t}$
- circle effects: ω_j
- effort effects: $B_i \frac{(\xi_*^{p_i} - 1)}{p_i}$
- overdispersion: $\varepsilon_{i,j,t}$

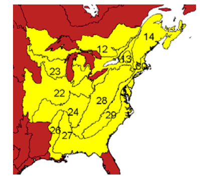
Regional Indices and Trends

$$n_{i,t} = z_i \exp\left(S_i + \beta_i (t - t^*) + \gamma_{i,t} \right)$$

$$T_i = \left\{ \frac{n_{i,t_b}}{n_{i,t_a}} \right\}^{\frac{1}{t_b - t_a}}$$

Note: z_i scales indexes to accommodate proportion of circles in a BCR with 0 counts

Example Analysis: Black Ducks in 11 Bird Conservation Regions, 1966-2003



Gritty Details of Example Analysis

- Set up BUGS code and data
- Compile and run
  - Burn-in time (1,500 iterations to reach stationarity)
  - 15,000 additional iterations
  - From this, get
    - Estimates and distributions (Posterior Distributions)
    - Diagnostics (did it work correctly?)

Estimates of *p* and *B* by BCR

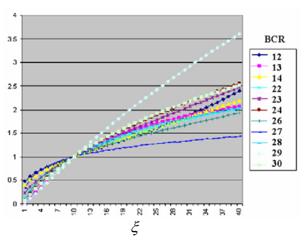
	p			B		
BCR	Median	2.50%	97.50%	Median	2.50%	97.50%
12	0.203	-0.975	1.335	0.401	0.028	0.834
13	-0.567	-0.805	-0.341	0.747	0.602	0.894
14	0.152	-0.187	0.550	0.478	0.285	0.676
22	-0.267	-0.555	-0.003	0.587	0.410	0.768
23	-0.024	-0.464	0.373	0.621	0.378	0.851
24	-0.298	-0.885	0.181	0.754	0.419	1.068
26	-0.735	-2.027	0.108	0.467	0.063	1.193
27	-0.418	-1.144	0.306	0.296	0.088	0.527
28	-0.523	-0.793	-0.272	0.699	0.519	0.879
29	-0.268	-0.579	0.049	1.079	0.850	1.307
30	-0.291	-0.485	-0.086	0.802	0.659	0.941

Relative Effort Adjustment by BCR, Calculated as:

$$e^{B_i(\xi_*^{p_i} - 1)/p_i}$$

where

$$\xi_* = \frac{\xi}{\overline{\xi}}, \overline{\xi} = 10$$



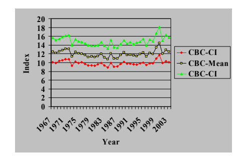
Scaled Effort (ξ_*) by BCR

BCR	Median	2.50%	97.50%	N Circles
12	0.597	0.1137	2.1722	1006
13	0.9523	0.3127	3.1092	3705
14	0.6311	0.1308	1.9331	2306
22	0.6538	0.1279	2.0987	3973
23	0.7036	0.1279	2.3378	2340
24	0.7164	0.1835	1.7566	1312
26	0.6581	0.135	2.1626	512
27	0.8422	0.2365	2.3433	1970
28	0.7889	0.199	2.405	3703
29	1.1371	0.2946	3.2661	1817
30	1.2935	0.253	3.9116	3503

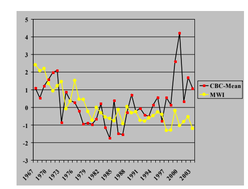
Trend (Proportional) Estimates

Stratum		Posterior Distribution Percentiles		
Num	Name	Median	2.50%	97.50%
12	Boreal Hardwood Transition	1.043	1.02	1.066
13	Lower Great Lakes/ St. Lawrence Plain	1.026	1.015	1.037
14	Atlantic Northern Forest	1.037	1.02	1.054
22	Eastern Tallgrass Prairie	0.9751	0.9605	0.9897
23	Prairie Hardwood Transition	0.9958	0.9872	1.008
24	Central Hardwoods	0.9677	0.9483	0.9854
26	Mississippi Alluvial Valley	0.9435	0.9042	0.9825
27	Southeastern Coastal Plain	0.9426	0.9272	0.9589
28	Appalachian Mountains	0.9875	0.976	0.9989
29	Piedmont	0.9801	0.97	0.9924
30	New England/mid-atlantic Coast	0.9962	0.9904	1.002
	Survey-Wide	0.9999	0.9939	1.006

Annual Indexes from Hierarchical Model



Comparison of CBC Annual Index with FWS Mid-winter Inventory Counts (Indexes Standardized for Comparison)



BUGS Code for Analysis

```
model;
{
 ######   Functions of Data   ######################
  xbar.effort <- mean(effort[1:ncounts])
  for( i in 1 : ncounts ) {
    ScaledEffort[i]<-effort[i]/xbar.effort
  }
 #### counts and overdispersion effects  ######
  for( k in 1 : ncounts ) {
    transeffort[k] <- (pow(ScaledEffort[k] ,p[strat[k]])-1)/p[strat[k]]
    Eloglambda[k] <- beta[strat[k]] * (year[k] - fixedyear) + cir[circle[k]] + strata[strat[k]] + B[strat[k]]*transeffort[k]+ yeareffect[year[k],strat[k]]
    loglambda[k] ~ dnorm(Eloglambda[k],taunoise)
    log(lambda[k]) <- loglambda[k]
    count[k] ~ dpois(lambda[k])
  }
  taunoise ~ dgamma(0.001,0.001)
  sdnoise <- 1 / pow(taunoise, 0.5)
  #### circle effects  ######
  for( i in 1 : ncircles ) {
    cir[i] ~ dnorm( 0.0,taucir)
  }
  taucir ~ dgamma(0.001,0.001)
  sdcir <- 1 / pow(taucir, 0.5)
  #### stratum effects  ######
 for( s in 1 : nstrata ) {
    p[s] ~ dunif(-4,4)
    B[s] ~ dnorm(0.0,1.0E-6)
    beta[s] ~ dnorm( 0.0,1.0E-6)
    strata[s] ~ dnorm( 0.0,1.0E-6)
    tauyear[s] ~ dgamma(0.001,0.001)
    sdyear[s] <- 1 / pow(tauyear[s],0.5)
  #### stratum specific year effects  ######
    for( y in 1 : nyears) {
       yeareffect[y,s] ~ dnorm( 0.0, tauyear[s])
    }
  }
  ###### summary statistics  ######
  totareaweight <- sum(areaweight[1:nstrata])
  for( i in 1 : nstrata ) {
    for( t in 1 : nyears ) {
       n[i,t] <- nonzeroweight[i]*exp(strata[i]+beta[i]*(t-fixedyear)+yeareffect[t,i])
       N[i,t] <- areaweight[i]*n[i,t]/totareaweight
    }
  }
  for( i in 1 : nstrata ) {
                Trend[i] <- pow(n[i,nyears]/n[i,1],1/(nyears-1))
  }
  for( t in 1 : nyears ) {
                CompIndex[t] <- sum(N[1:nstrata,t])
  }
  CompTrend <- pow(CompIndex[nyears]/CompIndex[1],1/(nyears-1))
  ######   Derived Effort Parameters ######################
  for( s in 1 : nstrata ) {
                for( i in 1 : 40 ) {
                EffortEffect[s,i]<- exp(B[s]*(pow(i/10,p[s])-1)/p[s])
                }
  }
  ######   overdispersion by stratum ######################
   for(s in 1 : nstrata ) {
     expstrata[s] <- exp(strata[s])
     expbeta[s] <- exp(beta[s])
     overdisp[s] <- 1 + 1/(expstrata[s]*taunoise)
  }
}
```