Modelling Birds Population

Ignacio Alvarez

Fall 2013

1 Introduction

According to (REF, Etterson) one of the basic research goals in Ecology consist on understand the distribution and abundance of the animal population. In this work, the particular goal will be to explore alternatives to model the time trends in Western Great Lakes Birds population over 1994 to 2011.

1.1 Data description

Next table shows total bird count on year 2007 for the most abbundant species, just to compare it with the online annual report. After the table the overall trend for the raw counts over the year are ploted.

	abbrev	count.9020	count.9030	count.9090
361	OVEN	1003	835	1168
362	REVI	823	997	771
363	BTNW	254		
364	NAWA	240	348	867
365	BLJA	222		230
366	CSWA	211	330	375
367	RBGR	208		
368	WTSP	180	387	940
369	HETH	175	249	265
370	AMRO	156		
373	VEER		402	264
375	LEFL		368	
378	AMRE		260	
380	PIWA		215	
386	MAWA			282
389	MOWA			257

Forest 9090 (NAMES) is consistently higer than the other two in terms of overal abundance, also it seems to be an increment of the total bird population over time regardless specie.

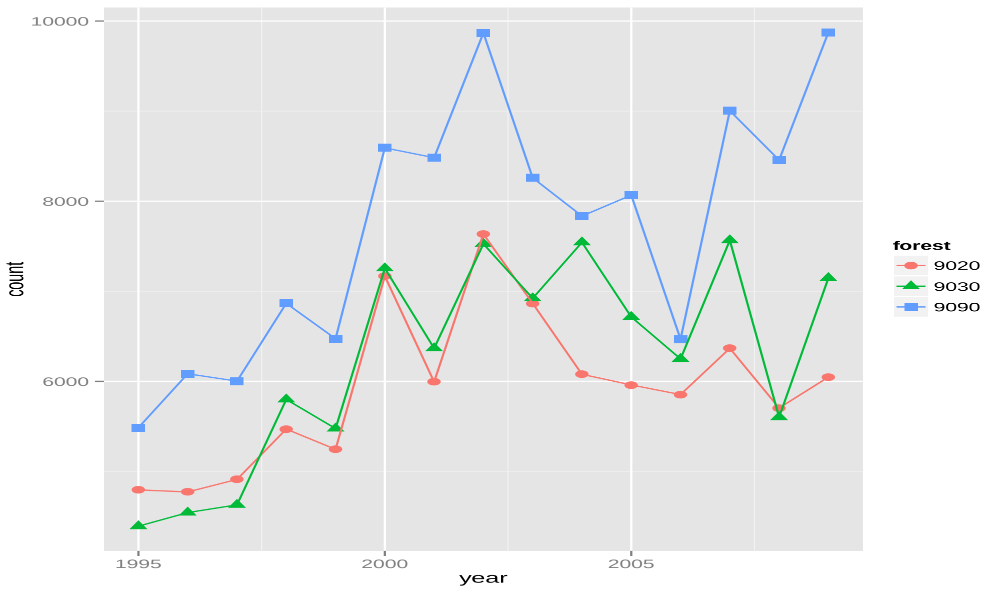

Figure 1: Raw trend in the data

1.2 Initial models

Perhaps the simplest model we could be a quadratic regression separatedly for each specie and forest.

$$\begin{aligned} Y_{tfs} &= \beta_{0fs} + \beta_{1fs}t + \beta_{2fs}t^2 + \epsilon_{tfs} \\ \epsilon_{tfs} &\sim N(0, \sigma^2) \end{aligned} \tag{1}$$

```
# ave : average count per year.
lm(ave ~ year + I(year^2), data = d)
```

where Y_{tfs} represent the average bird count on year t in the forest f for the specie s. There are 73 species and 3 forest in the data set so there are 219 models in total.

Table 1 shows the summary statistics for each coefficient and figure **??** presents histograms for each one of the model coefficients.

A second model considered is a regression using data from all 73 species in each forest and including random terms for the species coefficients.

$$\begin{aligned} log(Y_{tfs}) &= \beta_{0fs} + \beta_{1fs}t + \beta_{2fs}t^2 + \epsilon_{tfs} \\ \beta_{0fs} &\sim N(\beta_{0f}, \sigma^2_{0f}) \\ \beta_{1fs} &\sim N(\beta_{1f}, \sigma^2_{1f}) \\ \beta_{2fs} &\sim N(\beta_{2f}, \sigma^2_{2f}) \\ \epsilon_{tfs} &\sim N(0, \sigma^2) \end{aligned} \tag{2}$$

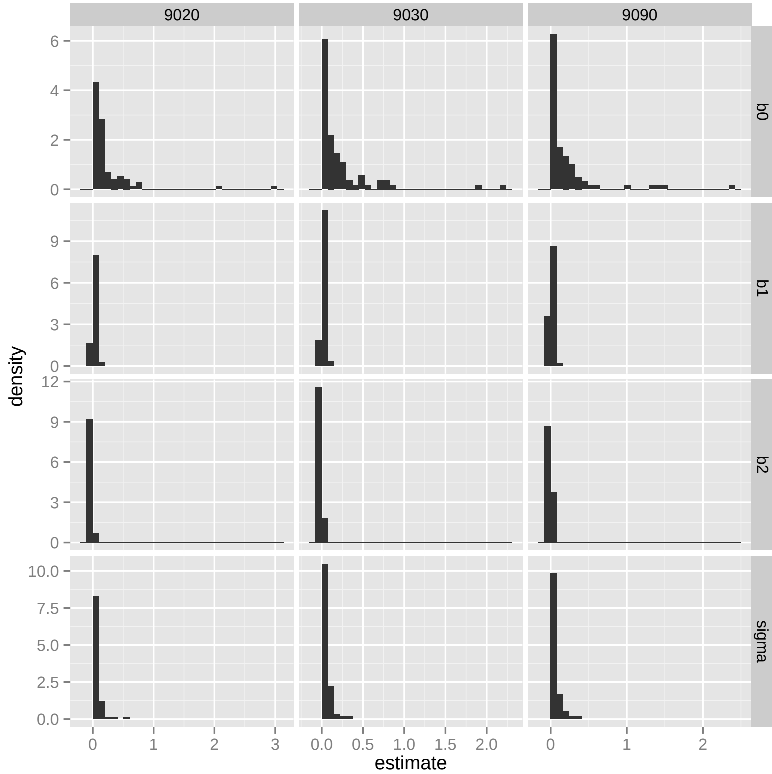

Figure 2: Histograms for coefficients of model 1.

```
# ave : average count per year.
library(lme4)
lmer(ave ~ year + I(year^2) + (year + I(year^2) | abbrev), data = d)
```

Table 1: Summary stats for coefficients of model 1

	forest	parameter	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1	9020	b0	0.003	0.043	0.115	0.238	0.212	3.003
2	9020	b1	-0.012	0.001	0.003	0.010	0.011	0.140
3	9020	b2	-0.029	-0.002	-0.001	-0.002	-0.000	0.000
4	9020	sigma	0.004	0.021	0.041	0.063	0.071	0.528
5	9030	b0	0.001	0.042	0.082	0.224	0.244	2.221
6	9030	b1	-0.007	0.001	0.003	0.010	0.014	0.125
7	9030	b2	-0.015	-0.001	-0.000	-0.001	-0.000	0.001
8	9030	sigma	0.002	0.018	0.032	0.054	0.074	0.370
9	9090	b0	0.000	0.015	0.072	0.217	0.223	2.407
10	9090	b1	-0.004	-0.000	0.003	0.008	0.007	0.110
11	9090	b2	-0.016	-0.001	-0.000	-0.001	0.000	0.001
12	9090	sigma	0.001	0.011	0.028	0.054	0.070	0.349

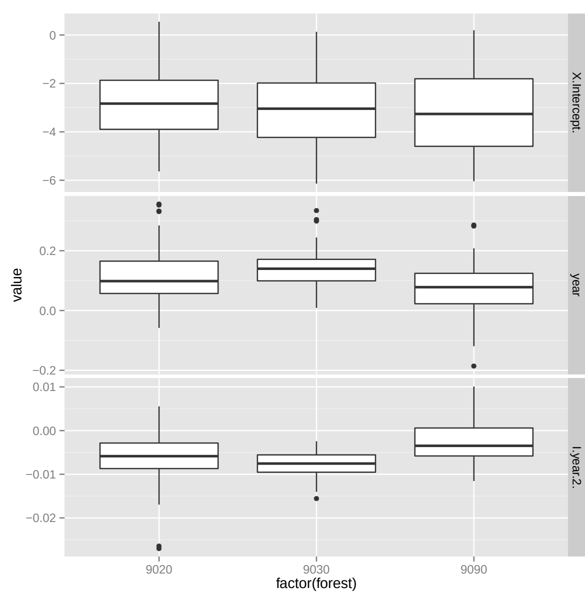

Figure 3: Boxplots for random effects in model 2.

Table 2: Estimated mean and variance in model **??**

	forest	parameter	mean	variance
1	9020	b0	-2.900	1.462
2	9020	b1	0.119	0.126
3	9020	b2	-0.007	0.008
4	9020	residual	0.000	0.487
5	9030	b0	-3.131	1.602
6	9030	b1	0.143	0.090
7	9030	b2	-0.008	0.004
8	9030	residual	0.000	0.496
9	9090	b0	-3.236	1.704
10	9090	b1	0.067	0.121
11	9090	b2	-0.003	0.006
12	9090	residual	0.000	0.518

2 Statistical Model

$$\begin{aligned}
log(Y_{tfs}) &= \beta_{0fs} + \beta_{1fs}t + \beta_{2fs}t^2 + \epsilon_{tfs} \\
\beta_{fs} &= \begin{pmatrix} \beta_{0fs} \\ \beta_{1fs} \\ \beta_{2fs} \end{pmatrix} \sim N(0, \Sigma_{fs}) \\
\Sigma_{fs} &\sim \text{inv-gamma, scaled inv-gamma, ...} \\
\epsilon_{tfs} &\sim N(0, \sigma_\epsilon^2) \\
\sigma_\epsilon^2 &\sim inv - gama(\alpha, \gamma)
\end{aligned} \tag{3}$$