Species Conservation in North American Forests, 1982-2020

Danielle LaVine, Mathematics
Mentor: Dr. Adam King
Kellogg Honors College Capstone Project

Introduction

The United Nations estimates that over one million species are threatened with extinction worldwide and notes that the abundance of species in most terrestrial habitats has decreased by at least 20% since 1900. To help understand the nature of species decline within the United States, we will study data from the International Union for Conservation of Nature (IUCN)'s Red List of Threatened Species for species inhabiting North America's forests. The IUCN Red List recognizes seven different categories which describe a species' risk of extinction: least concern (7), near threatened (6), vulnerable (5), endangered (4), critically endangered (3), extinct in the wild (2), and extinct (1). A species is considered *threatened* if it is assigned a Red List Category of 'vulnerable', 'endangered', or 'critically endangered'. We will study the Red List status of 773 species in North American forests from 1982 to 2020. We will estimate the trend of conservation status over time by fitting several mixed effects models that show how Red List Status has changed over time, considering the influence of age, period, and cohort on conservation status.

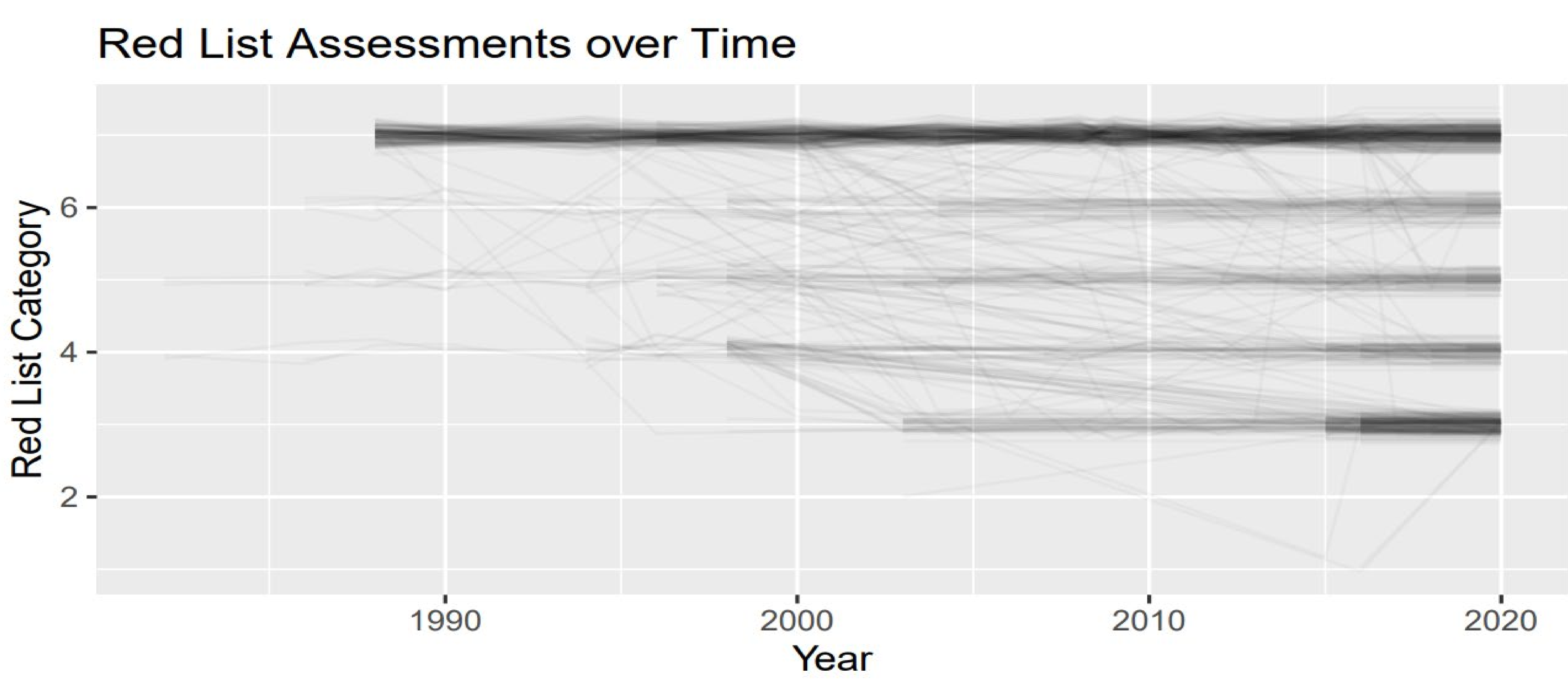


Figure 1: Trajectories of all 773 species. A slight jitter was added to this visual in order to avoid line overlap.

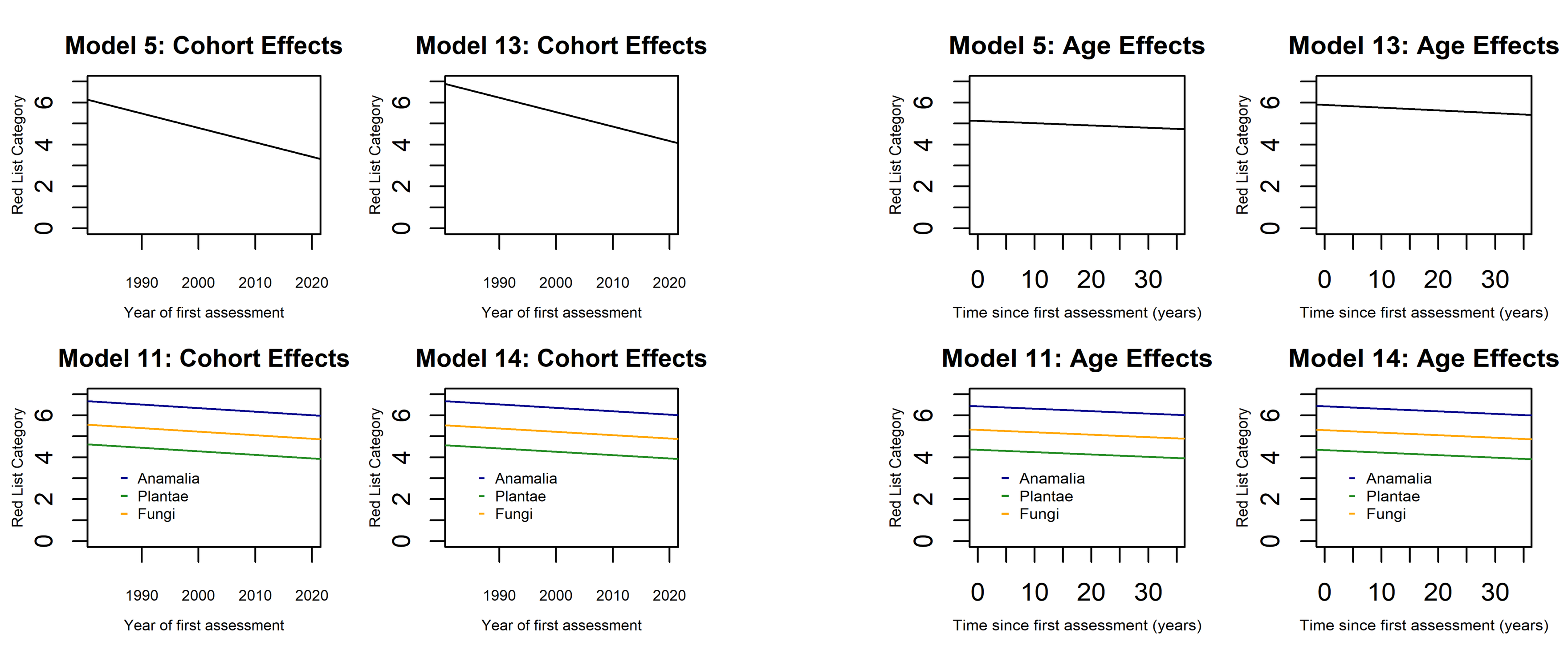
Results

We fit 14 different models which consider different combinations of age, period, cohort, taxonomy information, and random effects.

- Models 1-6 consider a random intercept with different combinations of our time-varying components.
- Models 7-12 are identical to models 1-6, with taxonomy added as a categorical predictor.
- Models 13 and 14 are extensions of two well performing models and consider a random slope for age.

We are interested in examining the age-cohort models (5 and 11) further because of the highly significant slope estimates in both versions of this model. We will examine the parameter estimates for the age-cohort model considering a random slope for age in Table 3.

- Models 13 and 14 show highly significant age effects when other factors are held fixed, and significant cohort effects when other factors are fixed.
- Both models appear to perform better than the age-cohort models that do not include a random intercept.
- The first models suggests that for every 15-year increase in the year that a species is first assessed, the Red List category decreases by 1, on average when age is fixed. Our second model shows that for every 15-year increase in the year that a species is first assessed, the Red List category decreases by 0.2 on average for fixed age and kingdom.



Methods

- **Mixed Effects Models**

Mixed effects models allow us to consider the variation between subjects, as well as identify average trends for all 773 species. Mixed effects models allow us to include a random intercept and/or random slopes. Let X_{ijk} denote the value of the k^{th} predictor at the j^{th} observation of the i^{th} species. Then the random intercept model is of the form

$$Y_{ij} = (\alpha + u_i) + \beta_1 X_{ij1} + \beta_2 X_{ij2} + \cdots + \beta_K X_{ijK} + \epsilon_{ij},$$

where α, β_K, and ε_{ij} are the usual intercept, slopes, and error constant of a standard linear regression model, and represent the estimated population average of these parameters. The value u_i represents a random intercept, which we assume has a normal distribution: $u_i \sim N(0,\sigma_u^2)$, where σ_u^2 is the individual species variation (van Belle et al., 2004).

- **Age-period-cohort analysis**

In longitudinal studies, age, period, and cohort effects can be useful in understanding time-varying components, such as how species conservation changes over time. For our data, we define age, period, and cohort as follows:

  - Age: time that has passed since the species' first assessment
  - Period: assessment year
  - Cohort: the year that a species is first assessed

Ideally, we would be able to include all three effects into a single model, however, this is impossible due to the exact linear dependence between age, period, and cohort, i.e., Period = Age + Cohort (Bell and Jones, 2015).

- **Assessing Goodness of Fit**

The goodness of fit for mixed effects models are described by a conditional R^2 and a marginal R^2. Marginal R^2 (R_m^2) measures the proportion of the variance explained by the fixed effects only, whereas conditional R^2 (R_c^2) measures the proportion of the variance that is explained by both fixed and random effects and is interpreted as the variance explained by the full model. Nakagawa and Schielzeth (2013) developed the following formulas for conditional and marginal R^2:

$$R_m^2 = \frac{\sigma_f^2}{\sigma_f^2 + \sigma_u^2 + \sigma_e^2}$$

$$R_c^2 = \frac{\sigma_f^2 + \sigma_u^2}{\sigma_f^2 + \sigma_u^2 + \sigma_e^2},$$

where σ_f^2 is the variance of the fixed effects, σ_u^2 is the variance of the random effects, and σ_e^2 is the variance of the residuals. We will use R_c^2 to assess the fit of our models.

Model No.	Model	Intercept	Age	Period	Cohort	σ_u^2	σ_e^2	R_m^2	R_c^2
1	Age- Only	5.238	-0.010***			2.543	0.228	0.003	0.918
2	Period-Only	5.238		-0.013***		1.825	0.240	0.006	0.912
3	Cohort-Only	5.081			-0.063***	1.825	0.240	0.164	0.903
4	Age-Period	5.134	0.058***	-0.069***		1.864	0.228	0.161	0.909
5	Age-Cohort	5.134	-0.011***		-0.069***	1.864	0.228	0.161	0.909
6	Period-Cohort	5.134		-0.011***	-0.058***	1.864	0.228	0.161	0.909

Table 1: Parameter estimates and measures of goodness of fit for initial age-period-cohort analysis. Note that all the two-factor models are mathematically equivalent. Significance levels denoted by* 0.05 ** 0.01 ***0.001

Model No.	Model	Intercept	Age	Period	Cohort	σ_u^2	σ_e^2	R_m^2	R_c^2
7	Age- Only	6.424	-0.011***			1.284	0.228	0.367	0.904
8	Period-Only	6.435		-0.012***		1.266	0.228	0.369	0.904
9	Cohort-Only	6.309			-0.011*	1.241	0.241	0.374	0.899
10	Age-Period	6.436	0.005	-0.017***		1.265	0.228	0.369	0.904
11	Age-Cohort	6.436	-0.012***		-0.017***	1.265	0.228	0.369	0.904
12	Period-Cohort	6.436		-0.017	-0.005***	1.265	0.228	0.369	0.904

Table 2: Parameter estimates and measures of goodness of fit for age-period-cohort analysis with kingdom added as a categorical predictor. Significance levels denoted by * 0.05 ** 0.01 ***0.001.

Model No.	Model	Intercept	Age	Cohort	σ_{u1}^2	σ_{u2}^2	σ_e^2	R_m^2	R_c^2
13	Without taxonomy predictor	5.888	-0.013***	-0.069***	1.848	0.001	0.153	0.151	0.941
14	With taxonomy predictor	6.438	-0.012***	-0.016**	1.298	0.001	0.153	0.366	0.936

Table 3: Age-cohort models with a random slope. Significance levels denoted by * 0.05 ** 0.01***0.001.

Discussion

Throughout the 14 models we produced, the slope estimates for age vary much more than the slope estimates for period and cohort. These effects also vary in significance. Age as a stand-alone predictor or including it with either cohort or period effects changes the trajectories of these estimate somewhat dramatically, even changing the sign of these estimates. The age-only and the age-cohort models suggest that conservation status, on average, deteriorates the longer a species is observed by the IUCN for fixed first assessment year. We also observe that adding information about taxonomy as a predictor results in notable differences in our model estimates, particularly within cohort effects.

The age-cohort models that we explored (models 5, 11, 13 and 14) thoroughly suggest that the Red List status of forest-dwelling species has, on average, been declining at alarming rates. While we are unsure of the exact reasons why we're estimating such declines, the IUCN recognizes a variety of different threats such as residential development, pollution, and climate change in different data sets which may be of interest in future study. We're also left with some ambiguity about the estimates for age, period, and cohort. While we can get a sufficient picture for how each of these effects affect conservation status, we can't know the true extent that each of these time-varying components interact with one another.

Acknowledgements

Special thanks to Dr. King, my family and friends, and the Kellogg Honors College for the guidance and support on this project and more. I would also like to acknowledge the International Union for Conservation of Nature, who provided the data utilized throughout this project.

References

- Bell, A. and Jones, K. (2015). Age, period and cohort processes in longitudinal and life course analysis: A multilevel perspective. In A life course perspective on health trajectories and transitions, pages 197–213. Springer, Cham
- IUCN (2020).The IUCN Red List of Endangered Species. Accessed: 2020-08-21.
- Mace, G. M. (2004). The role of taxonomy in species conservation. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 359(1444):711–719.
- Nakagawa, S. and Schielzeth, H. (2013). A general and simple method for obtainingR2fromgeneralized linear mixed-effects models. Methods in ecology and evolution, 4(2):133–142.
- van Belle, G., Fisher, L. D., Heagerty, P. J., and Lumly, T. (2004).Biostatistics: A Methodology For the Health Sciences. Wiley.