Improved Estimation of Bird Abundance and Energy Consumption by Combining Point and Transect Observations

Vinicio Santillán[1], Marta Quitián[3], Eike Lena Neuschulz[4], Matthias Schleuning[4], Katrin Böhning-Gaese[4], Aaron Brothers[2], Maria Jose Ayala Bolagay[1], Orlando Alvarez[1], and Christopher Thron*[2]

[1]Universidad Católica de Cuenca, Ecuador
[2]Texas A&M University-Central Texas, Killeen TX USA
[3]Instituto Mediterráneo de Estudios Avanzados (CSIC-UIB)
[4]Senckenberg Biodiversity and Climate Research Centre, Frankfurt Germany

April 16, 2023

Abstract

Point counts and transect surveys are both common methods for estimating bird abundance (by species, genus, or family) in a given environment. Sometimes both methods are used in the same study. Point counts reflect abundance directly, while transect observations of bird feeding instances reflect both the bird abundance and the mean consumption rate per bird.

In this paper we exploit the relationship between these two types of observations to draw inferences about bioenergy consumption, to develop an improved estimator of bird abundance, and to derive error bars for abundance estimates. These results are implemented practically on a dataset consisting of point counts and transect surveys of Ecuadorian frugivore bird species in six different types of environment (three elevations and two degrees of forestation).

First, by comparing bioenergy consumption models based on the two observation types, we clarify the dependence of bioenergy consumption on mean bird weight and mean fruit weight consumed by bird species. Second, we derive a procedure for forming an optimal linear combination of two relative abundance estimators based on point counts and transect observations respectively. The improvement in accuracy achieved by

*Corresponding author: thron@tamuct.edu

this combined estimator is verified by simulations applied to bird family data from the above-mentioned survey. Finally, we extend the procedure in order to obtain 90% confidence intervals for the bird family relative abundance estimates. These confidence intervals were also verified using simulations.

Keywords: Bioenergy, Abundance, Estimation, Survey, Point counts, Transect sampling, Bird, Ecosystem, Frugivore, Species, Family, Simulation

1 Introduction

Two common methods for estimating abundance of bird types (whether species, genuses, or families) in a given area are point counts and transect surveys. Point counts are obtained when the observer chooses a set of fixed observation points within the area of interest, and counts bird sightings during multiple sessions over an extended period of several days or months. Transect surveys are conducted by the observer specifying a transect (i.e. narrow section) within the area and repeatedly walking through and recording bird observations at the same time of day over an extended period. Transect measurements can target bird feedings, so that the frequency of feeding on each different type of plant is recorded.

These census methods has been extensively discussed in the literature, from several points of view. (See [5] and [6]). Using these two types of measurements, the relative bioenergy consumption of different bird types can be computed in two different ways. The total bioenergy consumption of a bird type can be expressed as the number of individuals of that type times the mean consumption per individual. On the other hand, the same quantity may be expressed as the total number of feedings by birds of that type times the amount consumed per feeding. Point counts can be used to estimate bird abundance required for the first expression, while transect surveys can provide feeding data for the second expression.

We may make use of the two alternative estimates of bioenergy consumption in several different ways. First, by comparing the two expressions, we may gain information about bird consumption rates. Second, by combining the two estimates, we can obtain a more accurate estimate of bioenergy consumption, and hence of bird relative abundance. Finally, given that count measurements are Poisson distributed (which is commonly recognized as valid for count measurements), we can use the results obtained and simulations to obtain confidence intervals for the actual bird relative abundance.

The remainder of this paper is organized as follows. In Section 2 we present the theoretical derivations for the statistical methods used for the data analysis. These methods include the construction of point count-based and transect observation-based estimators for bioenergy consumption (Section 2.1) and for bird abundance (Section 2.2); the combination of bird abundance estimators into a more accurate estimator (Section 2.3); and the description of the dataset used to test the methods derived in the preceding sections. In Section 3 the

methods described in the previous section are verified using simulations based on bird family field data. In Section 4 we summarize our main conclusions.

In Sections 2.1,2.2,and 2.3 we will use "bird type" as a generic term to denote bird species, genuses, or families. In Section 3 we will specialize to the case of bird species (for the determination of model parameters) and bird families (for performance evaluation of different estimators).

2 Methodology

2.1 Construction and comparison of bioenergy consumption measures

As mentioned in the introduction, the bioenergy consumption per bird type may be estimated as the total number of feedings times the mean bioenergy transfer per feeding. Initially we assume the transect observations are proportional to the actual number of feedings (plus some measurement error), so that observations reflect (up to measurement error) the relative frequencies of feedings across all bird types. We also assume initially that bioenergy transfer per feeding depends on both the fruit weight and the bird weight: this assumption presumes that larger fruit may require less foraging, and larger birds may be able to consume more per feeding.

Based on these assumptions, we obtain the following model relating bird bioenergy consumption to transect observations:

$$\mathcal{E}_n = \kappa_1 \sum_{m=1}^{M} (X_{nm} + u_{nm}) f_m^{p_f} w_n^{p_{b1}} \qquad (n = 1 \dots N), \tag{1}$$

where

- $\mathcal{E}_n$: Bioenergy consumption of bird type n.
- κ_1: Constant of proportionality, assumed to be independent of bird type.
- X_{nm}: Number of transect observations of bird type n consuming fruit type m;
- u_{nm}: Observational uncertainty for bird type n consuming fruit type m.
- M, N: Number of fruit types and number of bird types, respectively;
- f_m: mean fruit weight for type m.
- w_n: Mean weight of bird type n;
- p_f, p_{b1}: Exponents of fruit weight and bird weight respectively, in power model.

In the case there is not too much variation in the weight of fruits consumed by individual bird types, we may replace the model with a simplified model:

$$\mathcal{E}_n = \kappa_1 X_n (1 + u_n)(\bar{f}_n)^{p_f} (w_n)^{p_{b1}}, \tag{2}$$

where

- κ_2: Constant of proportionality, initially assumed to be independent of n.
- X_n: Sum of transect observations for bird type n. The formula for X_n is $X_n = \sum_{m=1}^{M} X_{nm}$.
- u_n: Combined observational uncertainty for bird type n. We model u_n as a random variable with mean 0 and variance $\sigma^2_{X,n}$;
- $\bar{f}_n$: Average weight of fruit consumed by type n, given by the formula:

$$\bar{f}_n := \frac{\sum_{m=1}^{M} X_{nm} f_m}{\sum_{m=1}^{M} X_{nm}}. \tag{3}$$

Besides combining all fruit types into a single averaged fruit weight, another difference between (1) and (2) is the change from an additive to a multiplicative form of the error term. If we allow the possibility of u_n depending on X_n then this is just a change of notation: the new u_n is equal to the previous u_n divided by X_n.

A second method for estimating a bird type's bioenergy consumption is to take the bird point count data to estimate the type's abundance, multiplied by the types' mean bioenergy requirement per individual, based on the mean bird weight for that particular type. A common model for bird bioenergy consumption is [2]:

$$\mathcal{E}_n = \kappa_2 Z_n (1 + v_n)(w_n)^{p_{b2}}, \tag{4}$$

where

- Z_n: Point count for bird type n;
- v_n: Uncertainty in point count sightings for bird type n. We model v_n as a random variable with mean 0 and variance $\sigma^2_{Z,n}$;
- p_{b2}: Exponent in power model.

As in (2), the model (4) assumes a multiplicative form of the error term.

Setting the two expressions for $\mathcal{E}_n$ equal to each other, we obtain:

$$\kappa_1 X_n (1 + u_n)(\bar{f}_n)^{p_f} (w_n)^{p_{b1}} = \kappa_2 Z_n (1 + v_n)(w_n)^{p_{b2}}. \tag{5}$$

Taking logarithms and rearranging, we find:

$$\log(X_n) = \log(Z_n) + C_0 + p_b \log(w_n) - p_f \log(\bar{f}_n) + e_n, \tag{6}$$

where:

$$
\begin{aligned}
C_0 &:= \log\left(\frac{\kappa_2}{\kappa_1}\right); \\
p_b &:= p_{b2} - p_{b1} \\
e_n &:= \log\left(\frac{1+u_n}{1+v_n}\right) \text{(combined error term)}
\end{aligned}
\tag{7}
$$

Note that if $u_n, v_n \ll 1$ then $e_n \approx u_n - v_n$. We will assume that $\{e_n, n = 1, ...N\}$ are independent, identically distributed random errors. Then based on on relation (6), we may perform ordinary unweighted least squares regressions to estimate parameters C_0, p_b, and p_f that give the best fit to the data.

2.2 Combined estimates of relative abundance

Derivation and equations

In Section 2.1, we designed models that relate bioenergy consumption to transect sampling and point counts, respectively. In Section 3.1 by comparing these models, we will find that one parameter (fruit weight exponent p_f) is not statistically significantly different from 0. We will also obtain an estimate of the bird weight parameter p_b. In this section we make use of these results to develop two alternative estimates for relative type abundance, where relative type abundance is defined as type abundance divided by total bird abundance. One of these estimates is based on point counts, and the other is based on transect observations.

First we develop the estimator based on point counts. We denote the relative abundance of type n as a_n. We hypothesize that point counts are proportional to relative abundance (apart from measurement error).

$$
a_n = \frac{Z_n}{\sum_{k=1}^{N} Z_k}(1+v_n). \tag{8}
$$

where Z_n and v_n are the point count and measurement error respectively for type n. The division by $\sum_{k=1}^{N} Z_k$ corresponds to the normalization condition that the sum over n of $\{a_n\}$ is equal to 1.

Next, an alternative expression for a_n based on transect observations is obtained by using (24) to substitute in for Z_n:

$$
a_n = \frac{(w_n)^{-p_b} X_n}{\sum_{k=1}^{N} (b_k)^{-p_b} X_k}(1+u_n), \tag{9}
$$

where X_n and u_n are the total transect count and the measurement error respectively for type n.

Equations (8) and (9) give two different estimates of a_n, which we denote as $(\widehat{a}_n)_Z$ and $(\widehat{a}_n)_X$:

$$
(\widehat{a}_n)_Z = \frac{Z_n}{\sum_k Z_k}; \qquad (\widehat{a}_n)_X = \frac{(w_n)^{-\widehat{p}_{bn}} X_n}{\sum_k (b_k)^{-\widehat{p}_{bn}} X_k}, \tag{10}
$$

where $\widehat{p}_{bn}$ is the estimate for the bird weight coefficient p_{bn}. Alternatively, we may define estimators for the logarithm of relative abundance, which we denote as $(\widehat{\ell}_n)_Z$ and $(\widehat{\ell}_n)_X$:

$$\begin{aligned}
(\widehat{\ell}_n)_Z &:= \log\left((\widehat{a}_n)_Z\right) = \log Z_n - \log\left(\sum_k Z_k\right); \\
(\widehat{\ell}_n)_X &:= \log\left((\widehat{a}_n)_X\right) = \log X_n - \widehat{p}_{bn}\log w_n - \log\left(\sum_k (b_k)^{-\widehat{p}_{bn}} X_k\right).
\end{aligned} \tag{11}$$

Since we have estimators based on different, independent measurements (Z_n and X_n), we may take a linear combination of the two estimators to obtain a combined estimator which may have a smaller variance, and thus greater accuracy:

$$\begin{aligned}
(\widehat{a}_n)_{\beta_n,\gamma_n} &:= \beta_n(\widehat{a}_n)_Z + \gamma_n(\widehat{a}_n)_X; \\
(\widehat{\ell}_n)_{\beta'_n,\gamma'_n} &:= \beta'_n(\widehat{\ell}_n)_Z + \gamma'_n(\widehat{\ell}_n)_X.
\end{aligned} \tag{12}$$

It follows by the linearity of distribution mean that:

$$\begin{aligned}
\text{mean}\left[(\widehat{a}_n)_{\beta_n,\gamma_n}\right] &:= \beta_n \text{mean}\left[(\widehat{a}_n)_Z\right] + \gamma_n \text{mean}\left[(\widehat{a}_n)_X\right]; \\
\text{mean}\left[(\widehat{\ell}_n)_{\beta'_n,\gamma'_n}\right] &:= \beta'_n \text{mean}\left[(\widehat{\ell}_n)_Z\right] + \gamma'_n \text{mean}\left[(\widehat{\ell}_n)_X\right].
\end{aligned} \tag{13}$$

Through simulations, it was verified that the estimators $(\widehat{a}_n)_Z$, $(\widehat{a}_n)_X$, $(\widehat{\ell}_n)_Z$, $(\widehat{\ell}_n)_X$ are unbiased, so that:

$$\begin{aligned}
\text{mean}\left[(\widehat{a}_n)_Z\right] &= \text{mean}\left[(\widehat{a}_n)_X\right] = a_n; \\
\text{mean}\left[(\widehat{\ell}_n)_Z\right] &= \text{mean}\left[(\widehat{\ell}_n)_X\right] = \ell_n,
\end{aligned} \tag{14}$$

The requirement that $(\widehat{a}_n)_{\beta_n,\gamma_n}$ and $(\widehat{\ell}_n)_{\beta'_n,\gamma'_n}$ be unbiased estimators of a_n and ℓ_n respectively is expressed mathematically as:

$$\text{mean}\left[(\widehat{a}_n)_{\beta_n,\gamma_n}\right] = a_n; \qquad \text{mean}\left[(\widehat{\ell}_n)_{\beta'_n,\gamma'_n}\right] = \ell_n. \tag{15}$$

Combining (13), (14), and (15) gives the requirements:

$$\gamma_n = 1 - \beta_n; \qquad \gamma'_n = 1 - \beta'_n. \tag{16}$$

Given that $(\widehat{a}_n)_Z$ and $(\widehat{a}_n)_X$ are based on independent observations and are thus statistically independent, it follows:

$$\begin{aligned}
\text{Var}\left[(\widehat{a}_n)_{\beta_n,\gamma_n}\right] &= \beta_n^2 \text{Var}\left[(\widehat{a}_n)_Z\right] + \gamma_n^2 \text{Var}\left[(\widehat{a}_n)_X\right]; \\
\text{Var}\left[(\widehat{\ell}_n)_{\beta_n,\gamma_n}\right] &= (\beta'_n)^2 \text{Var}\left[(\widehat{\ell}_n)_Z\right] + (\gamma'_n)^2 \text{Var}\left[(\widehat{\ell}_n)_X\right],
\end{aligned} \tag{17}$$

and a standard calculation shows that the unbiased minimum-variance estimators for a_n and ℓ_n are obtained by choosing:

$$\begin{aligned} \gamma_n &:= (1-\beta_n); \quad \beta_n := \frac{\mathrm{Var}[(\widehat{a}_n)_X]}{\mathrm{Var}[\widehat{a}_n)_Z] + \mathrm{Var}[(\widehat{a}_n)_X]}; \\ \gamma'_n &:= (1-\beta'_n); \quad \beta'_n := \frac{\mathrm{Var}[(\widehat{\ell}_n)_X]}{\mathrm{Var}[(\widehat{\ell}_n)_Z] + \mathrm{Var}[(\widehat{\ell}_n)_X]} \end{aligned} \tag{18}$$

2.3 Methodology for calculating and verifying near-optimal abundance estimator

The simulation results presented in Section 3.2 show that it is possible to combine point count-based and transect observation-based estimates to obtain more accurate estimators. Of course, in practice one will never see 10,000 ecologists taking measurements on the same site! However, the results show that an ecologist with both point count and transect survey data may be able to combine these data to obtain more accurate abundance estimates. However, there is a difficulty with this in that the ecologist does not know the optimal values of the coefficients β_n (which may differ between bird types) to use to combine the two estimates. We do not have an analytical expression for calculating these values. However, the ecologist can herself run simulations with the data she has obtained to estimate the optimal values of β_n In this section, we will describe a practical procedure by which ecologists can determine the optimum values of the coefficients in (12) via computer simulation. Furthermore, we will verify the effectiveness of this procedure using simulations.

The proposed estimation procedure for β_n proceeds as follows. First, using his obtained point counts $\{PC_n\}_{n=1\ldots N}$ and transect survey observations $\{TO_n\}_{n=1\ldots N}$, the ecologist generates a series of K simulated point count-transect observation vectors $\{z_n, x_n, n = 1\ldots N\}_k, k = 1\ldots K$, where K is a large number. Using the results, the ecologist may estimate the statistics of $(\widehat{a}_n)_\beta, (\widehat{a}_n)_\beta, (\widehat{\ell}_n)_\beta$, and $(\widehat{\ell}_n)_\beta$ calculated from $\{z_n, x_n\}$ using (9) and(12). From these results the β_n that minimizes RMSE for each n may be estimated, and used for final abundance estimates. Simulation statistics can also be used to give error bars for these estimates. Note that the relative abundance estimates may not sum exactly to 1. However, we are not attempting to find exact values, just best estimates.

The relative performance of these estimators was established using simulations as follows.

- First, based on each of the six different regions observed in the study, 2500 scenarios were generated with randomly-Poisson-distributed relative abundances $a_n = z_n / \sum_{j=1}^{N} z_j$, where z_n is a Poisson random variable according to (26).

- Next, for each scenario, point counts and transect observations (simulating the dataset collected a single hypothetical "ecologist" studying the

scenario) were randomly generated as z'_n and x'_n respectively, where (compare (26)–(28))

$$z'_n \sim Pois(z_n); \qquad p'_{bn} = Gauss(-0.35, 0.125);$$
$$x'_n \sim Pois\left(z_n w_n^{p'_{bn}} \frac{\sum_{j=1}^N TO_j}{\sum_{j=1}^N z_j w_j^{p'_{bj}}}\right). \tag{19}$$

Each simulated ecologist then uses her data to estimate relative abundances a_n with estimators based on z'_n and x'_n:

$$(\widehat{a}_n)_{z'} := \frac{z'_n}{\sum_k z'_k}; \qquad (\widehat{a}_n)_{x'} := \frac{(w_n)^{-\overline{p}_b} x'_n}{\sum_k (b_k)^{-\overline{p}_b} x'_k}; \tag{20}$$

- Next, each ecologist generates 2500 simulated sets of point counts and transect observations (z''_n and x''_n, respectively) based on her own observations:

$$z''_n \sim Pois(z'_n); \qquad x''_n \sim Pois\left(z'_n w_n^{\overline{p}_b} \frac{\sum_{j=1}^N TO_j}{\sum_{j=1}^N z'_j w_j^{\overline{p}_b}}\right). \tag{21}$$

Using these 500 sets of point counts, the ecologist calculates β''_n for each bird type n, such that $\beta = \beta''_n$ minimizes the RMSE of the estimator

$$(\widehat{a'}_n)_\beta := \beta(\widehat{a'}_n)_{z''} + (1-\beta)(\widehat{a'}_n)_{x''}, \tag{22}$$

- Finally, using data from the 2500 simulated ecologists we calculate the 95% confidence error bars (±2 standard deviations) for the estimators $(\widehat{a}_n)_{z'}$, $(\widehat{a}_n)_{x'}$, and

$$(\widehat{a}_n)_{\beta''} := (\widehat{a'}_n)_\beta := \beta''_n(\widehat{a}_n)_{z'} + (1-\beta''_n)(\widehat{a}_n)_{x'}. \tag{23}$$

2.4 Dataset used in methodology evaluation

Data was collected from 18 1-hectare plots spread across three elevations (1000 m above sea level (asl) (970–1092 ± 47.2 m), 4° 6′ S, 78° 58′ W; 2000 m asl (2025–2119 ± 32.4 m), 3° 58′ S, 79° 4′ W; and 3000 m asl (2874–2966 ± 32.9 m], 4° 6′ S, 79° 10′ W. These plots represented two types of habitat (natural forest and fragmented forest), within and around Podocarpus National Park (PNP) and San Francisco reserve (BRSF) on the southeastern slope of the Andes in Ecuador. The six sites are characterized in Table 1. Study plots were selected jointly by the German Platform for Biodiversity and Ecosystem Monitoring and Research in South Ecuador, to ensure typical representation of each forest type. The area exhibits a humid tropical montane climate [4] with a bimodal rainfall pattern (highest precipitation season from May through June, lowest precipitation season from October through November) [3].

Location	Approx. altitude (m)	Environment type
Bombuscaro	1000	Natural
ECSF	1000	Fragmented
Cajanuma	2000	Natural
Copalinga	2000	Fragmented
Finca	3000	Natural
Bellavista	3000	Fragmented

Table 1: Summary of characteristics for six locations studied

To assess bird diversity and abundance, nine observation points were established within each 1-hectare plot. For a duration of 10 minutes at each point, we recorded and identified all bird species heard or seen within a 20-meter radius [1]. This sampling was repeated eight times per plot over a span of two years (1/6 hours $\times$ 9 observation points $\times$ 8 replicates $\times$ 18 plots = 216 total sampling hours). The total number of bird species (species richness) and the overall number of individual birds (abundance) were quantified by aggregating the records from all observation points and replicates per plot. Bird species were classified as frugivores if their diet consisted of 40% or more fruit [7].

3 Results

3.1 Application of bioenergy estimation methodology to dataset

The models were applied to per-species data from each of the six locations separately; to data grouped by altitude; to data grouped by environment type (natural versus fragmented); and to the entire dataset. In each case, the model parameters were determined by unweighted linear least squares regression.

The six panels in Figure 1 contain scatter plots showing actual transect observations of bird species in log scale (y axis) versus transect observations predicted by the model (6) in log scale, for each of the six locations studied. The x and y axes in these plots represent the right and left sides of (6) respectively. Thus an exact fit to the theoretical model corresponds to a line of slope 1. We may visually appraise goodness of fit by seeing how close data points lie to this theoretical line. The dots' colors represent the mean weight of the corresponding bird species.

Figures 2, 3, and 4 similarly show predicted versus actual transect observations grouped by elevation, fragmentation, and all data respectively. Results of the best fit for the two parameter model (6) for the six separate regions are summarized in Table 2. The table includes the model parameter estimates and 95% confidence intervals, as well as the R^2 values and Akaike information criterion (AIC) values for the different regions, altitudes, environment types, and

Location	C_0 est.	C_0 CI	p_b est.	p_b CI	p_f est.	p_f CI	R^2	AIC
By region								
Bombuscaro	3.6	[2.4,4.9]	-0.6	[-0.9,-0.3]	0.2	[-0.2,0.5]	0.30	92
ECSF	3.7	[0.4,7.1]	-0.6	[-1.4,0.2]	0.2	[-0.8,1.0]	0.19	64
Cajanuma	3.6	[-7.2,14.5]	-0.7	[-4.0,2.1]	0.3	[-1.0,1.6]	0.12	24
Copalinga	2.0	[0.3,3.7]	-0.2	[-0.5,0.2]	0.2	[-0.2,0.5]	0.03	141
Finca	3.0	[0.6,5.3]	-0.7	[-1.3,-0.04]	0.02	[-0.4,0.5]	0.25	60
Bellavista	3.7	[-4.0,11.4]	-0.7	[-2.4,0.9]	0.7	[-0.3,1.8]	0.40	23
By altitude								
Altitude 1000	3.7	[-4.0,11.4]	-0.7	[-2.4,0.9]	0.7	[-0.3,1.8]	0.40	23
Altitude 2000	3.2	[1.4,5.1]	-0.7	[-1.2,0.9]	-0.2	[-0.6,0.2]	0.3	86
Altitude 3000	4.2	[-0.8,9.1]	-0.8	[-2.0,0.3]	-0.7	[-1.4,-0.0]	0.40	33
By environment type								
Natural environment	3.4	[2.2,4.5]	-0.6	[-0.8,-0.3]	-0.1	[-0.4,0.1]	0.40	166
Fragmented environment	1.9	[0.6,3.1]	-0.3	[-0.6,-0.0]	0.0	[-0.2,0.2]	0.08	208
All birds	2.2	[1.2,3.1]	-0.4	[-0.6,-0.1]	-0.1	[-0.3,0.1]	0.11	245

Table 2: Summary of parameter estimations from regression for the two-coefficient model (bird weight and fruit weight) (6)

total for all families for all regions. R^2 values range from 0.03 to 0.40, indicating relatively poor fits.

Figure 5 shows 95% confidence intervals (± 2 standard deviations) for the model coefficients p_b (*left*) and p_f (*right*) as estimated from different data. Confidence intervals are shown for the coefficient estimates using per-species data from six separate locations (yellow bars); using data grouped by altitude (purple bars); using data grouped by fragmentation (blue bars); and using all data combined (green bar). In 11 of the 12 cases (including the overall data model), the confidence intervals for p_f contain 0, indicating a lack of statistical evidence for a nonzero value of p_f. On the other hand, in 6 of the 12 cases (including the model for overall data), the confidence interval for p_b lies entirely below 0, indicating that p_b is negative at 5% significance.

The fact that p_f is not significantly different from 0 suggests we may modify our model by replacing $p_f = 0$ in 6. The model then has only the coefficient p_b plus the constant C_0 that are fitted to the transect and point count data:

$$\log(X_n) = \log(Z_n) + C_0 + p_b \log(w_n) + e_n. \tag{24}$$

Figures 6–9 show the transect observation prediction accuracy for the one-coefficient model (24) by region, elevation, environment type and combined,

Location	C_0 est.	C_0 CI	p_b est.	p_b CI	R^2	AIC
By region						
Bombuscaro	3.4	[2.2,4.5]	-0.6	[-0.9,-0.2]	0.27	91
ECSF	3.4	[0.4,7.1]	-0.6	[-1.3,0.2]	0.18	63
Cajanuma	1.7	[-3.9,7.3]	-0.3	[-1.9,1.4]	0.03	22
Copalinga	1.5	[0.28,2.7]	-0.14	[-0.5,0.2]	0.01	140
Finca	2.9	[0.8,5.0]	-0.7	[-1.3,-0.0]	0.25	58
Bellavista	-0.9	[-5.4,3.6]	0.09	[-1.15,1.33]	0.00	25
By altitude						
Altitude 1000	1.9	[0.9,2.9]	-0.3	[-0.5,0.0]	0.06	156
Altitude 2000	2.9	[1.2,4.6]	-0.6	[-1.1,-0.2]	0.27	85
Altitude 3000	0.2	[-3.7,4.1]	-0.1	[-1.1,1.0]	0.00	37
By environment type						
Natural	3.1	[2.1,4.0]	-0.5	[-0.8,-0.25]	0.22	166
Fragmented	1.9	[0.9,2.9]	-0.3	[-0.6,0.0]	0.08	206
All	2.0	[1.2,2.8]	-0.3	[-0.6,-0.1]	0.1	244

Table 3: Summary of parameter estimations from regression for the single coefficient model (bird weight only) (6)

respectively (compare Figures 1–4, which show corresponding accuracies for the two-coefficient model). Visually, the the one-coefficient model appear to have accuracies comparable to the two-coefficient model.

Table 3 summarizes the fit parameters for the single-coefficient model that includes only bird weight and not fruit weight. The AIC values for all fits are reduced, thus justifying the exclusion of fruit weight as a significant parameter.

Figure 10 summarizes error bars the bird weight coefficient in the single coefficient model of the form (24). Error bars are consistent across all regions, and overlap in the range of -0.3 to -0.5 As with the two-coefficient model, the bird weight coefficients tend to be negative, and some are negative with 95% significance (including the overall estimate). Error bars for high altitudes are large, reflecting the fact that bird weights are less variable at high altitudes. Taken together, the data gives a 95% confidence interval for p_b as $[-0.6, -0.1]$. Recall that p_b is defined in (7) as the difference of the two exponents p_{b1} and p_{b2}, where p_{b1} is related to bird weight's influence on bioenergy consumption per bird, while p_{b2} is related to bird weight's influence on bioenergy consumption per observed feeding. Estimates of p_{b1} in [2] for most families are close to 0.7 (hummingbirds are an exception, with $p_{b1} > 1$). Our finding of $p_b \approx -0.4$ is consistent with a value of $p_{b1} \approx 1$. This would imply that bird energy consumption

per feeding is roughly proportional to bird weight. This is not unreasonable, because heavier birds proportionately have larger anatomical features, including larger stomachs.

Figure 11 shows a scatter plot of the ratio of feeding observation/transect observations versus transect observations frequency. The figure indicates a negative correlation between the feeding/observation ratio and the number of observations. This is an indication of regression to the mean: overpredictions (or underpredictions) based on transect observations will tend to be revised downwards (or upwards) based on feeding observations—and vice versa. The presence of regression to the mean indicates that point counts and transect observations tend to counterbalance each other, and estimators based on both factors may achieve greater accuracy. We investigate this further in the next section.

3.2 Verification of combined estimator using simulations

In order to test the conclusions in Section 2.2 under a variety of conditions, we conducted six series of simulations based on the six different regions studied. Six hypothetical regions (corresponding to the six actual regions in the study) were used, in which the true relative abundances in each hypothetical region were set according to the per-family point count data obtained for the corresponding actual region. We base simulations on per-family data instead of per-species data, this is because per-species point counts are too small to establish differences in accuracies between estimators. It should be noted that the purpose of basing the simulations on measured data is simply to provide a variety of scenarios that may reasonably encountered in practice. The simulations simply take the measured data as ground truth, and evaluate estimators' accuracy in recovering this ground truth from noisy data.

First we describe the simulations' mathematical structure. If $PC_1, PC_2, \ldots PC_N$ are the observed point counts in a given actual region, then the true relative abundance of family n in the corresponding hypothetical region is given by:

$$a_n = \frac{PC_n}{\sum_{j=1}^{N} PC_j}. \tag{25}$$

Next, we simulate ecologist point counts in the hypothetical region. For a given simulation, the point count for each bird family n is generated by a Poisson random variable whose mean is the point count from the observed data. We denote these Poisson random variables as $z_n, n = 1, \ldots N$, where

$$z_n \sim Pois(PC_n). \tag{26}$$

We also generate transect observations randomly as follows. First, for each bird family n we generate a weight exponent p_{bn} as a Gaussian random variable with 95% confidence interval [-0.6,-0.1] (which is the 95% confidence interval for p_b given in Table 3 for all birds):

$$p_{bn} = Gauss(-0.35, 0.125). \tag{27}$$

We then simulate the ecologist's transect observations for birds in family n (denoted as x_n) as follows:

$$x_n \sim Pois\left(PC_n w_n^{p_{bn}} \frac{\sum_{j=1}^{N} TO_j}{\sum_{j=1}^{N} PC_j w_j^{p_{bj}}}\right), \tag{28}$$

where TO_j denotes the number of transect observations of bird family j in the original data ($j = 1 \ldots N$). The choice of distribution mean in (28) was made so that the expected sum of simulated observations over bird families at a given region is equal to the sum of transect observations in the original data.

Given these simulated point counts $\{z_n\}$ and transect observations $\{x_n\}$, We then compute estimates $\widehat{a}_n)_z$ and $\widehat{a}_n)_x$ for a_n based on $\{z_n\}$ and $\{x_n\}$ respectively (compare (9) and (12)):

$$(\widehat{a}_n)_z := \frac{z_n}{\sum_k z_k}; \quad (\widehat{a}_n)_x := \frac{(w_n)^{-\overline{p_b}} x_n}{\sum_k (b_k)^{-\overline{p_b}} x_k}, \tag{29}$$

where $\overline{p_b} = -0.35$ is the mean value of the distribution in (27).

We may also compute the combined estimate $(\widehat{a}_n)_\beta$ (compare (12)):

$$(\widehat{a}_n)_\beta = \beta(\widehat{a}_n)_z + (1-\beta)(\widehat{a}_n)_x, \tag{30}$$

Similarly, we compute the following estimates for $\{\ell_n\}$:

$$(\widehat{\ell}_n)_z := \log(\widehat{a}_n)_z; \qquad (\widehat{\ell}_n)_x := \log(\widehat{a}_n)_x; (\widehat{\ell}_n)_\beta := \beta(\widehat{\ell}_n)_z + (1-\beta)(\widehat{\ell}_n)_x. \tag{31}$$

We simulate 10,000 ecologist estimates according to the above prescription, and calculate the means and variances of all estimates. Results are displayed in Figures 12–19. Figures are described in more detail as follows.

Figure 12 shows the mean over the 10,000 simulations for the estimators $(\widehat{a}_n)_z$ and $(\widehat{a}_n)_x$ for different bird families for different regions, compared to the true value. Results show that there is very little bias in most estimators, where the estimators for birds with lower relative abundance tending have the most bias. Figure 13 shows similar results for the estimators $(\widehat{\ell}_n)_z$ and $(\widehat{\ell}_n)_x$.

Figure 14 shows the relative standard deviations for $(\widehat{a}_n)_\beta$ for values of β between 0 and 1, for all different bird families in the six different regions. The standard deviations are based on estimates of 10,000 simulated ecologists. The triangles on the horizontal axis mark the optimal values of β as calculated by (18). The figure shows that (18) gives a very accurate estimate of the β that minimizes standard deviation. All of the optimal β values are between 0.5 and 1, reflecting the fact that $(\widehat{a}_n)_z$ (corresponding to $\beta = 1$) is consistently smaller than $(\widehat{a}_n)_x$ (corresponding to $\beta = 0$). Figure 15 shows similar results for the standard deviations for $(\widehat{\ell}_n)_\beta$ for the different bird families in different regions. Figure 16 shows the relative root mean squared error (RMSE) of $(\widehat{a}_n)_\beta$ as a function of β for different bird families per region. The graph is almost identical to Figure 14, showing that bias error is negligible compared to standard

deviation. Figure 17 gives similar results for $(\widehat{\ell}_n)_\beta$ as similar results for the standard deviations for $(\widehat{\ell}_n)_\beta$ for the different bird families in different regions.

Figure 18 shows RMSE-minimizing β values for $(\widehat{a}_n)_\beta$ versus relative abundance of bird family for the six regions. It is clear that lower β values only occur for birds with relative abundance less than about 0.1, with only two exceptions. Figure 19 shows that similar results hold for RMSE-minimizing β values for $(\widehat{\ell}_n)_\beta$.

Figures 12–19 presume that the ecologists do not have accurate information about p_{bn}, so the use the mean value $\bar{p}_{bn} = 0.35$ to calculate the abundance estimates in (29). If the ecologists do have accurate information about p_{bn}, then the situation changes considerably. Figures 20–27 correspond to Figures 12–19, except they show results obtained when ecologists have accurate information about p_{bn}. It is evident that the estimates $(\widehat{a}_n)_X$ and $(\widehat{\ell}_n)_X$ (corresponding to $\beta = 0$ in Figures 22–25) are more accurate than the corresponding estimates when p_{bn} is not known exactly (as depicted in Figures 22–25). As a result the optimal values of β tend to be lower when p_{bn} is known exactly, as shown in Figures 26 and 27 (compare Figures 18 and 19 for the inexact case).

3.3 Verification of accuracy of near-optimal estimator

Results from the simulations described in Section 2.3 below are shown graphically in Figures 28–31 and summarized numerically in Tables 4 and 5. The figures show that relative abundance or log relative abundance estimates using the combined estimator $(\widehat{a}_n)_{\beta''}$ are always at least as accurate as the best of the two estimators $(\widehat{a}_n)_{z'}$ and $(\widehat{a}_n)_{x'}$.

Figures 28–29 include results for transect-based estimators that were calculated using the average value $\overline{p_b} = -0.35$. The transect observation based estimators are the most inaccurate, especially for bird families with smaller relative abundance. The combined estimates in some cases are slightly better than the point count based estimators, especially for families with smaller abundances. Figure 29 shows the same conclusions hold for estimates of log relative abundance.

The relatively large errors transect based estimators is the fact that $\overline{p}_b$ is used in the estimate rather than the true value of p_{bn} which may vary from family to family. If the ecologist has more accurate information about p_{bn}, then the situation changes considerably. Figure 30 shows relative accuracies of relative abundance estimators that make use of exact values for p_{bn} for each family, rather than the overall average $\overline{p_b}$. The figure shows that in this case, the transect observation based abundance estimators are much more accurate. Furthermore, the combined estimators achieve an appreciable gain over the two individual estimators for many bird families.

	$\widehat{a}_Z$ rel. err.		$\widehat{a}_X$ rel. err. (mean $\overline{p}_b$ / exact p_b)		$\widehat{a}_\beta$ rel. err. (mean $\overline{p}_b$ / exact p_b)	
	Mean	Std	Mean	Std	Mean	Std
Bombuscaro	0.28	0.21	0.55 / 0.16	0.18 / 0.11	0.22 / 0.14	0.12 / 0.10
Copalinga	0.30	0.19	0.64 / 0.23	0.27 / 0.17	0.25 / 0.18	0.15/ 0.13
ECSF	0.23	0.1	0.66 / 0.21	0.29 / 0.12	0.21 / 0.15	0.15 / 0.08
Finca	0.16	0.10	0.54 / 0.16	0.28 / 0.14	0.15 / 0.11	0.09 / 0.08
Cajanuma	0.18	0.11	0.33 / 0.14	0.21 / 0.09	0.16 / 0.10	0.1 / 0.07
Bellavista	0.20	0.16	0.6 / 0.49	0.34 / 0.50	0.2 / 0.18	0.16 / 0.15
Altitude 1	0.29	0.20	0.61 / 0.20	0.24 / 0.15	0.24 / 0.16	0.14 / 0.12
Altitude 2	0.20	0.10	0.61 / 0.19	0.29 / 0.13	0.18 / 0.13	0.09 / 0.08
Altitude 3	0.19	0.15	0.51 /0.37	0.33 / 0.44	0.18 / 0.16	0.14 / 0.13
Natural	0.25	0.18	0.57 / 0.26	0.26 / 0.27	0.21 / 0.16	0.11 / 0.13
Fragmented	0.25	0.18	0.61 / 0.26	0.29 / 0.27	0.22 / 0.16	0.15 / 0.13
TOTAL	0.25	0.18	0.59 / 0.22	0.27 / 0.22	0.21 / 0.15	0.13 / 0.11

Table 4: Means and standard deviations of RMSE for relative error of Z-based, X-based, and combined relative abundance estimators, for individual regions and grouped by altitude, environment type, and total.

	$\widehat{l}_Z$ error		$\widehat{l}_X$ error (mean $\overline{p}_b$ / exact p_b)		$\widehat{l}_\beta$ error (mean $\overline{p}_b$ / exact p_b)	
	Mean	Std	Mean	Std	Mean	Std
Bombuscaro	0.26	0.15	0.53/0.18	0.15/0.13	0.22/0.14	0.11/0.091
Copalinga	0.29	0.17	0.58/0.23	0.18/0.16	0.25/0.18	0.13/0.11
ECSF	0.24	0.12	0.63/0.22	0.2/0.13	0.22/0.16	0.1/0.086
Finca	0.17	0.11	0.52/0.16	0.18/0.14	0.16/0.12	0.097/0.087
Cajanuma	0.2	0.14	0.33/0.15	0.2/0.098	0.17/0.12	0.11/0.077
Bellavista	0.21	0.18	0.53/0.34	0.11/0.2	0.18/0.17	0.12/0.12
Altitude 1	0.28	0.16	0.56/0.21	0.17/0.15	0.24/0.16	0.12/0.11
Altitude 2	0.21	0.12	0.58/0.19	0.2/0.13	0.2/0.14	0.1/0.089
Altitude 3	0.21	0.16	0.46/0.27	0.17/0.19	0.18/0.15	0.12/0.11
Natural	0.25	0.14	0.55/0.19	0.2/0.13	0.22/0.15	0.11/0.089
Fragmented	0.25	0.17	0.56/0.23	0.17/0.17	0.22/0.16	0.13/0.11
Total	0.25	0.15	0.57/0.23	0.22/0.22	0.22/0.15	0.12/0.1

Table 5: Means and standard deviations of RMSE for error of Z-based, X-based, and combined log relative abundance estimators, for individual regions and grouped by altitude, environment type, and total.

3.4 Estimation of confidence intervals for relative abundance and log relative abundance

As with all scientific measurements, abundance estimates should be accompanied by confidence intervals that indicate the accuracy of these estimates. The

methods described above for combining estimators can also be exploited to obtain confidence intervals for the combined estimators. In order to estimate the parameter β for each bird family, an ecologist must run simulations based on her own measurements. These simulation produce a distribution of relative abundance estimates, with an associated confidence interval. Since the ecologist's measurements approximate the true relative abundances, it follows that the confidence intervals that she obtains from simulations can also approximate the confidence intervals for her own estimates (this is an example of "bootstrapping", which is a widely-used technique in computational statistics).

The technique described in the previous paragraph was applied to obtain confidence intervals for relative abundance and log relative abundance for the different families observed in the six different regions. The simulations used to create the estimates in Tables 4 and 5 were also used to generate 90% confidence intervals for each family. To evaluate the reliability of these error bars, for each family we also computed the proportion of time the "real" relative abundance and log relative abundance lay within the computed 90% confidence interval. Results are shown in Table 6.

	$\widehat{a}_\beta$ 90% C.I. proportions (mean $\overline{p}_b$ est. / exact p_b est.)		$\widehat{\ell}_\beta$ 90% C.I. proportions (mean $\overline{p}_b$ est. / exact p_b est.)	
	Mean proportion	Std. of proportion	Mean proportion	Std. of proportion
Bombuscaro	0.26	0.15	0.53/0.18	0.15/0.13
Copalinga	0.29	0.17	0.58/0.23	0.18/0.16
ECSF	0.24	0.12	0.63/0.22	0.2/0.13
Finca	0.17	0.11	0.52/0.16	0.18/0.14
Cajanuma	0.2	0.14	0.33/0.15	0.2/0.098
Bellavista	0.21	0.18	0.53/0.34	0.11/0.2

Table 6: Means and standard deviations of proportion of instances where the true relative abundance (respectively log relative abundance) lay within the confidence intervals computed by the technique described in this section.

4 Conclusions

The principal conclusions of this study are as follows:

- There is no statistical evidence that families' observed consumption events via transect observations depends on the weight of fruit consumed.
- There is statistical evidence supporting existing literature that bioenergy consumption of bird families depends on bird weight.
- There is also statistical evidence that bioenergy consumption per transect observation is larger for heavier birds than for lighter birds.

	$\widehat{a}_\beta$ 90% C.I. proportions (mean $\overline{p}_b$ est. / exact p_b est.)		$\widehat{\ell}_\beta$ 90% C.I. proportions (mean $\overline{p}_b$ est. / exact p_b est.)	
	Mean proportion	Std. of proportion	Mean proportion	Std. of proportion
Bombuscaro	0.9061/0.9004	0.0074/0.0035	0.9065/0.9008	0.0074/0.0035
Copalinga	0.9053/0.9008	0.0078/0.0041	0.9057/0.9012	0.0078/0.0041
ECSF	0.9046/0.9033	0.0071/0.0019	0.905/0.9037	0.0071/0.0019
Finca	0.9027/0.9053	0.0053/0.0078	0.9031/0.9057	0.0053/0.0078
Cajanuma	0.902/0.9076	0/0	0.9024/0.908	0/0
Bellavista	0.9035/0.9011	0.0053/0.0088	0.9039/0.9015	0.0053/0.0088
At1	0.9056/0.9006	0.0077/0.0039	0.906/0.901	0.0077/0.0039
At2	0.9037/0.9042	0.0064/0.0054	0.9041/0.9046	0.0064/0.0054
At3	0.903/0.9033	0.0044/0.0078	0.9034/0.9037	0.0044/0.0078
Natural	0.905/0.9024	0.007/0.0035	0.9054/0.9028	0.007/0.0035
Fragmented	0.9044/0.9019	0.007/0.0064	0.9048/0.9023	0.007/0.0064
TOTAL	0.9047/0.9021	0.007/0.0053	0.9051/0.9025	0.007/0.0053

- Point counts and transect observations can be used to obtain independent estimates of relative abundance and log of relative abundance. Typically, the point count-based estimates are more accurate.
- The two estimates mentioned in the previous item may be linearly combined to form another estimate which has lower mean squared error.
- Ecologists can make use of simulations to estimate the parameter β needed to form the optimal linear combination of point count based and transect observation based estimators.
- With more accurate information concerning the dependence of bird type bioenergy consumption as a function of weight, a better estimate of β can be made that leads to combined estimators of abundance with lower error.

References

[1] Colin J Bibby, Neil D Burgess, David M Hillis, David A Hill, and Simon Mustoe. *Bird census techniques*. Elsevier, 2000.

[2] Gurney J. Crocker D., Hart A. and McCoy C. Project pn0908: Methods for estimating daily food intake of wild birds and mammals. Technical report, Central Science Laboratory, 2002.

[3] Paul Emck. *A climatology of south Ecuador-with special focus on the major Andean ridge as Atlantic-Pacific climate divide*. PhD thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), 2007.

[4] Markus Kottek, Jürgen Grieser, Christoph Beck, Bruno Rudolf, and Franz Rubel. World map of the köppen-geiger climate classification updated. *Meteorologische Zeitschrift*, 15(3):259–263, 07 2006.

[5] C. John Ralph, John R. Sauer, and Sam Droege. Monitoring bird populations by point counts. Technical report, 1995.

[6] William J Sutherland. *Ecological census techniques: a handbook*. Cambridge university press, 2006.

[7] Hamish Wilman, Jonathan Belmaker, Jennifer Simpson, Carolina de la Rosa, Marcelo M Rivadeneira, and Walter Jetz. Eltontraits 1.0: Species-level foraging attributes of the world's birds and mammals: Ecological archives e095-178. *Ecology*, 95(7):2027–2027, 2014.

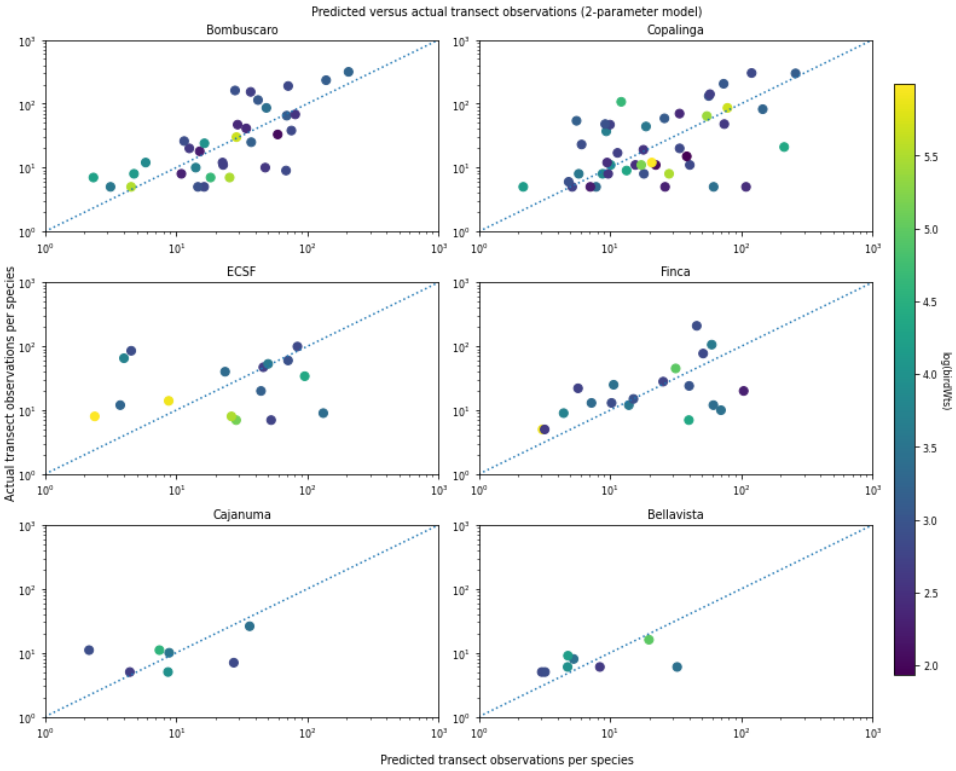

Figure 1: Transect observation prediction accuracy by location for two-coefficient model that includes both bird and fruit weights. The 45 degree line corresponds to zero prediction error.

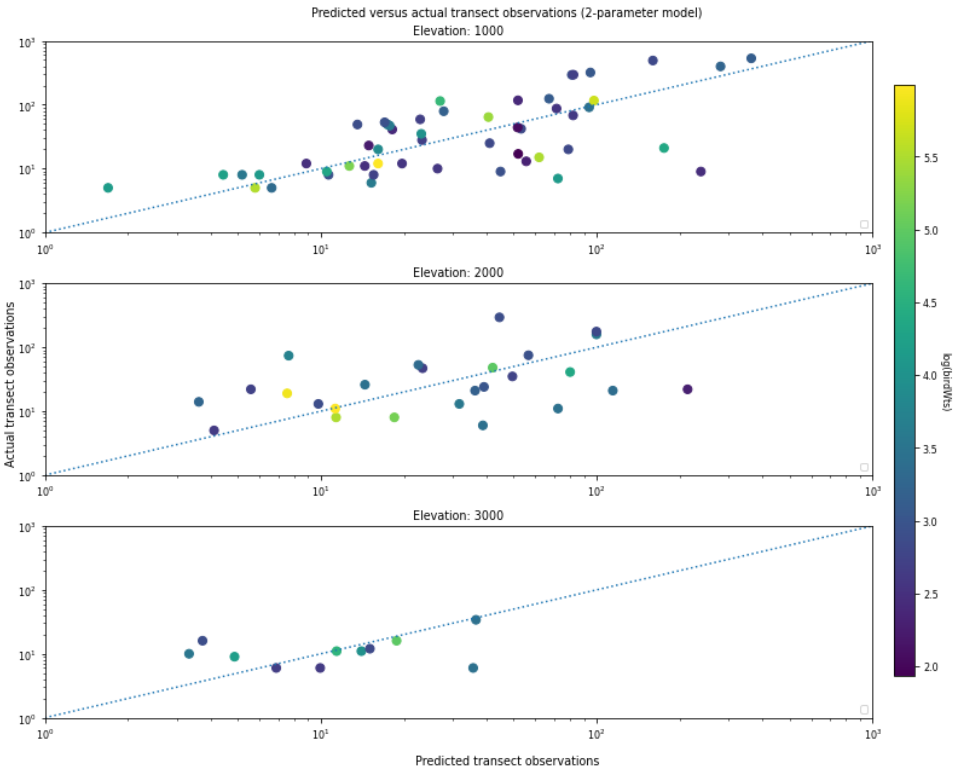

Figure 2: Transect observation prediction accuracy of two-coefficient model, by elevation.

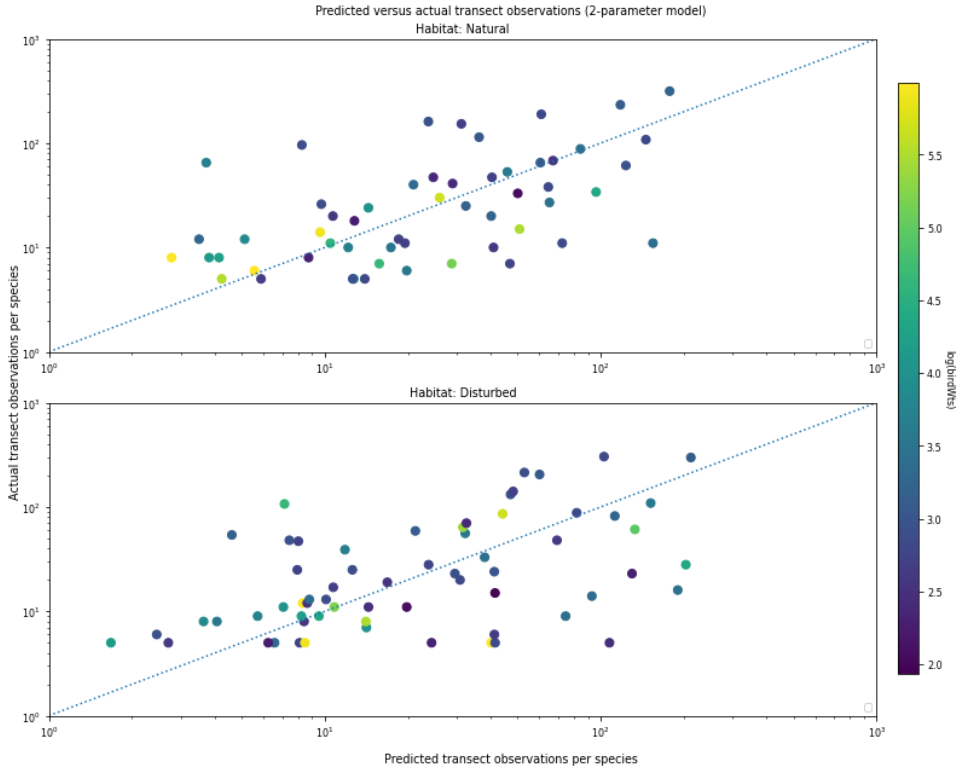

Figure 3: Transect observation prediction accuracy of two-coefficient model, by fragmentation.

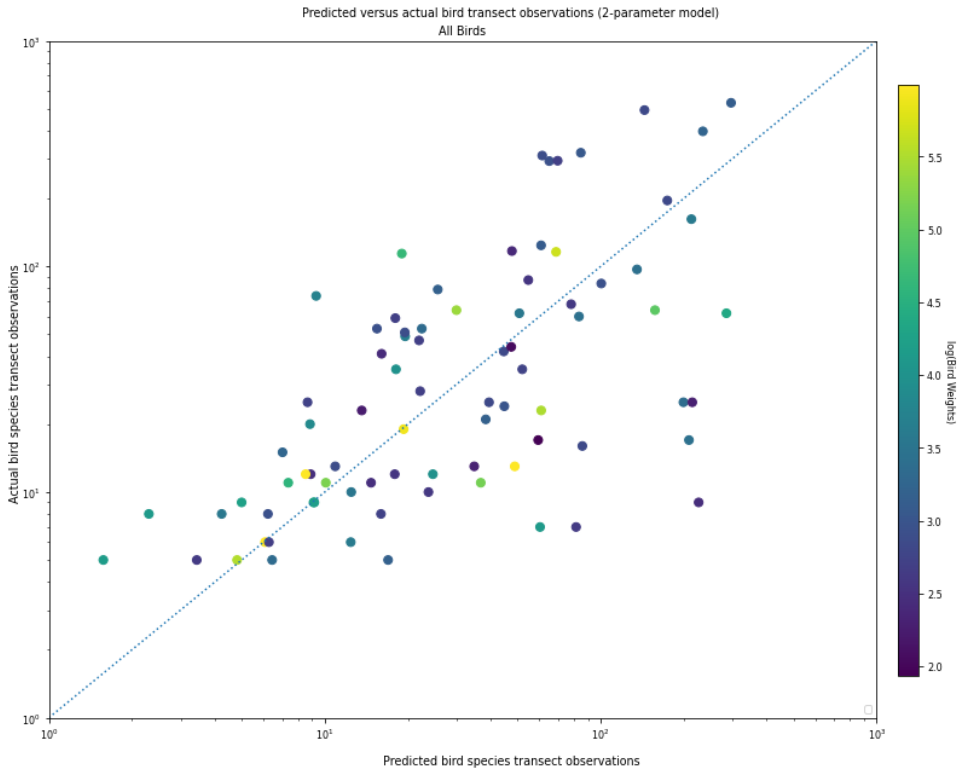

Figure 4: Transect observation prediction accuracy for two-coefficient model, all observations.

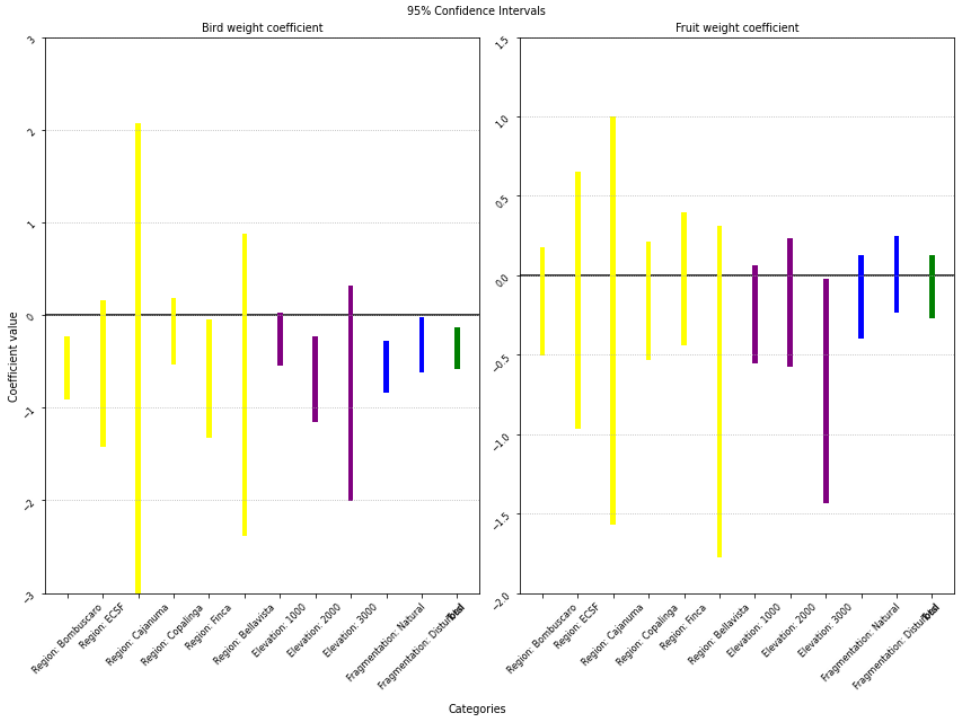

Figure 5: Error bars for bird weight coefficient (*left*) and fruit weight coefficient (*right*) in two-coefficient model (bird_weight +fruit_weight) by location (yellow), by elevation (purple), by fragmentation (blue) and overall(green)

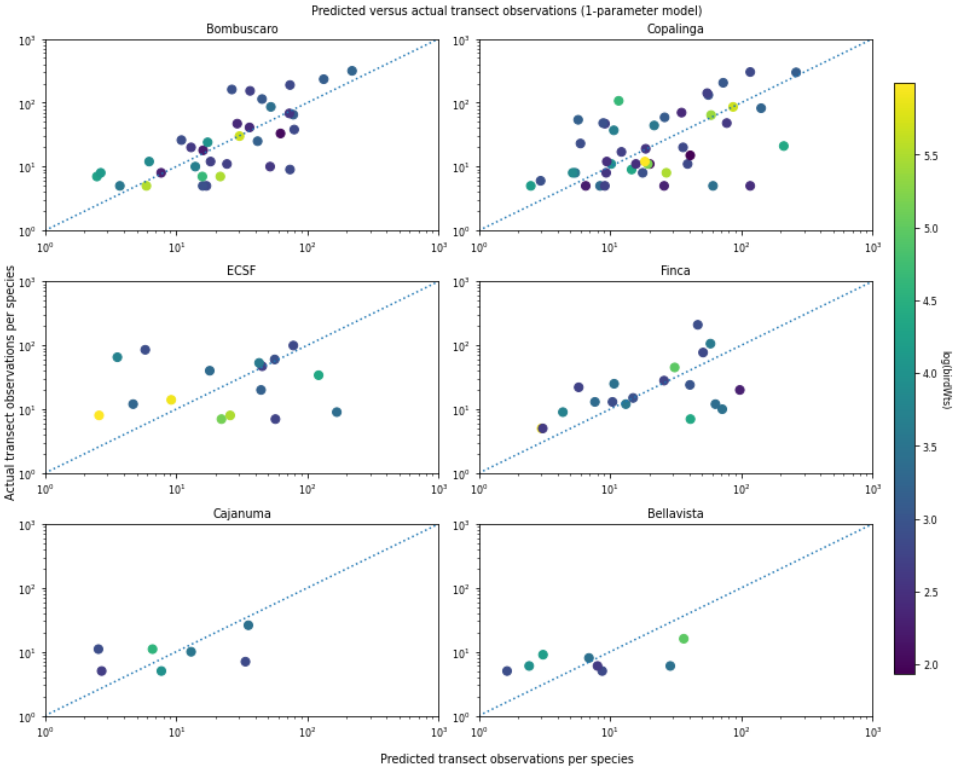

Figure 6: Transect observation prediction accuracy for one-coefficient model, by region.

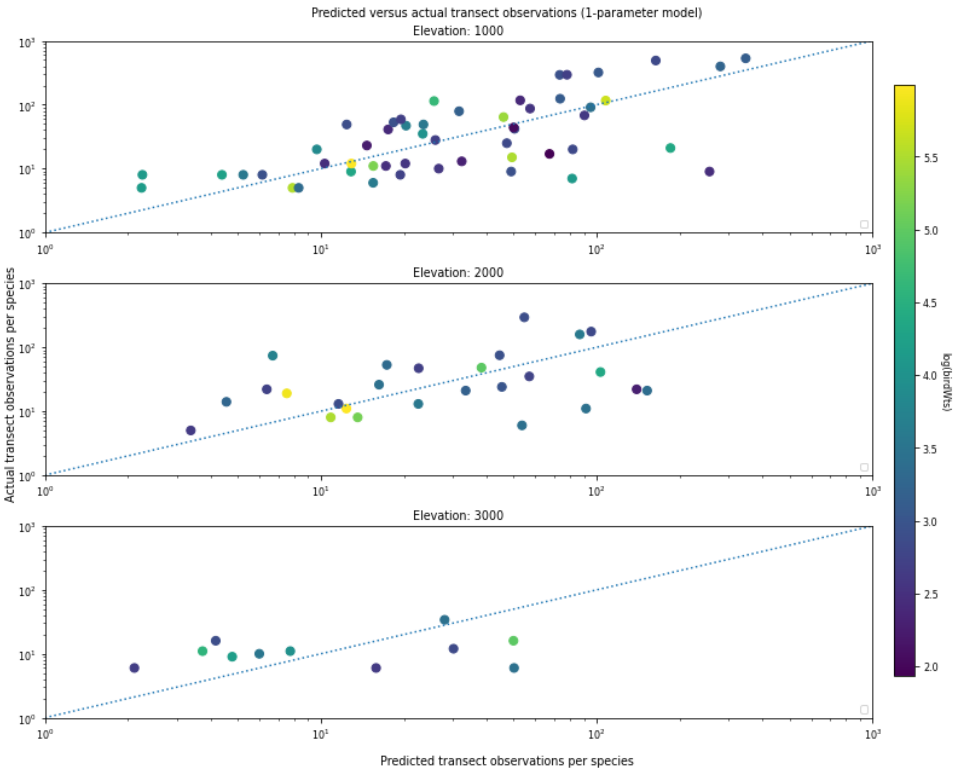

Figure 7: Transect observation prediction accuracy for one-coefficient model, by elevation.

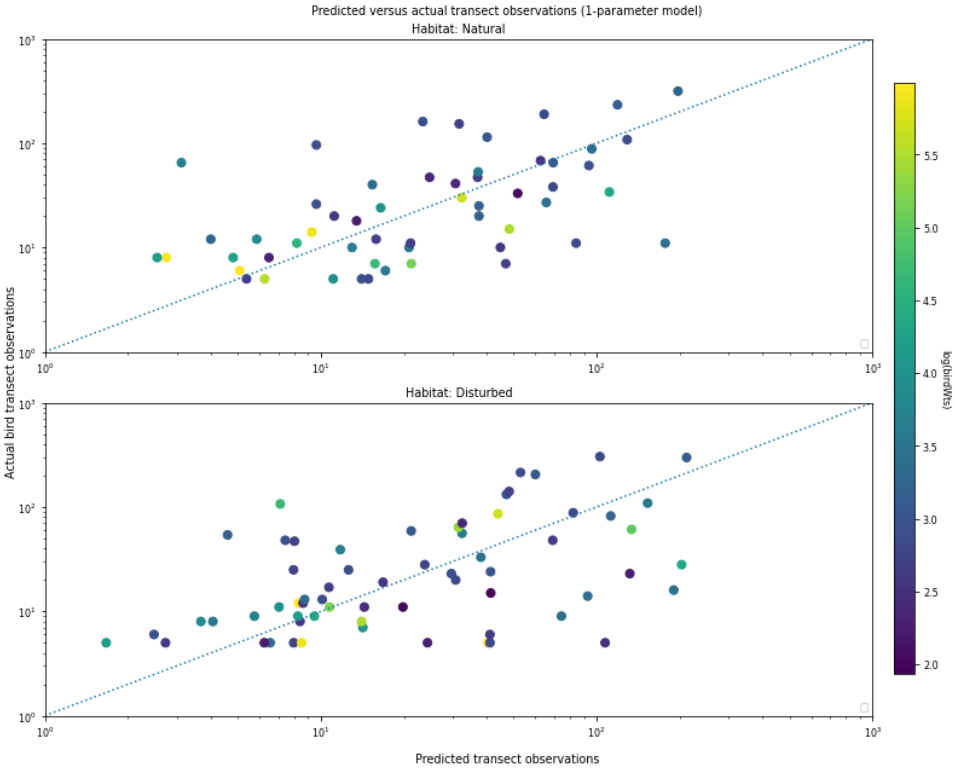

Figure 8: Transect observation prediction accuracy for one-coefficient model, by fragmentation.

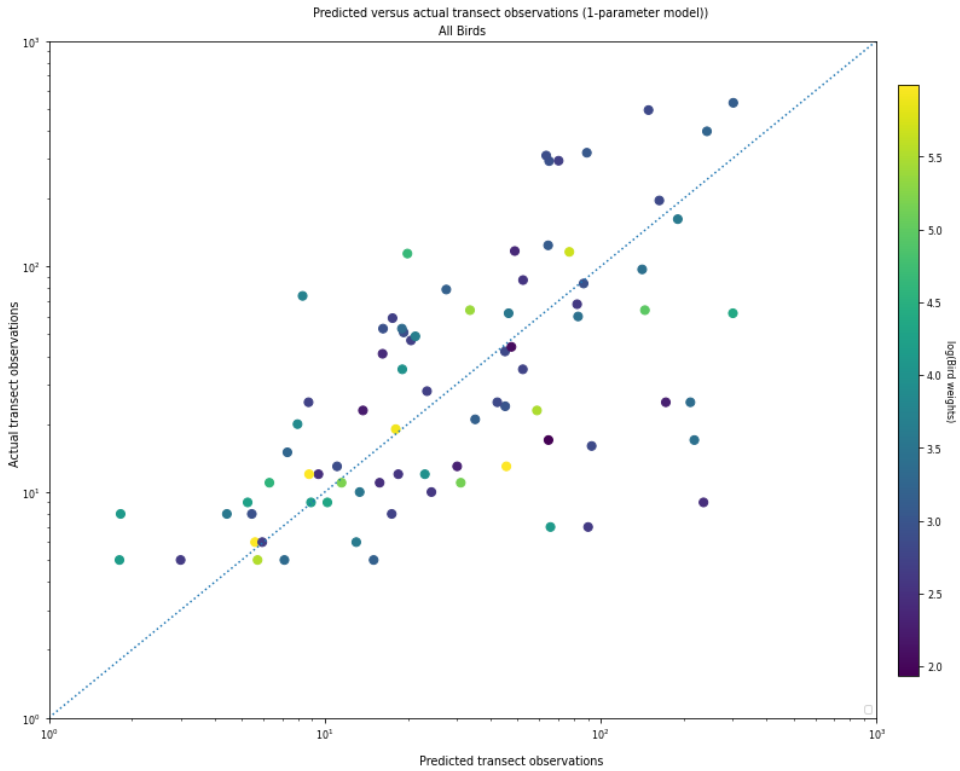

Figure 9: Transect observation prediction accuracy for one-coefficient model, all observations.

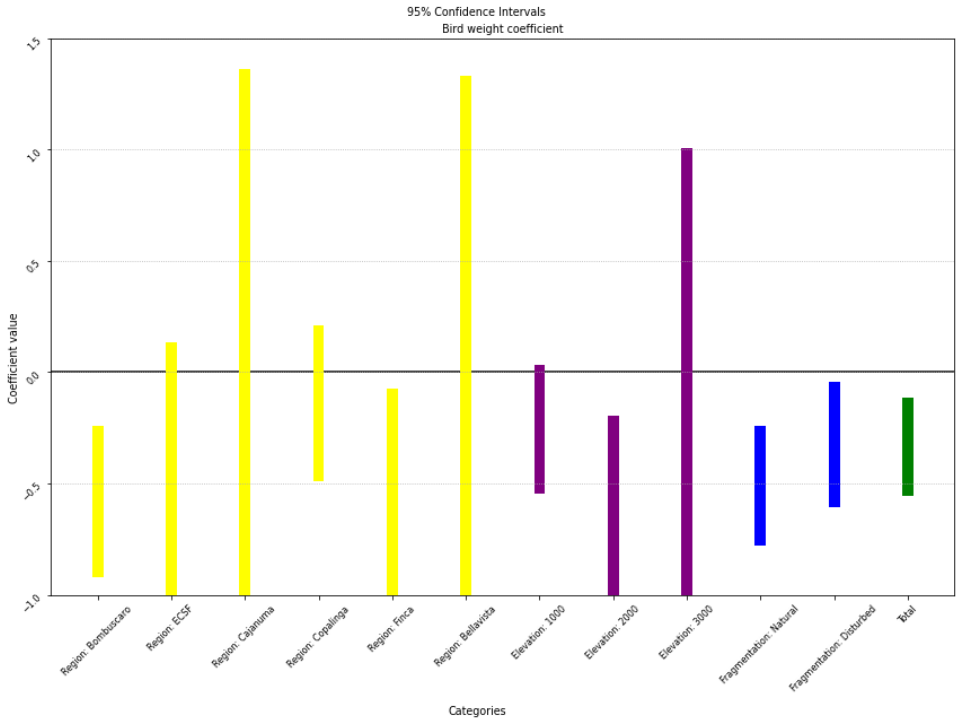

Figure 10: Error bars for one-coefficient model (bird weight only) model coefficients by location, by elevation, by fragmentation and overall

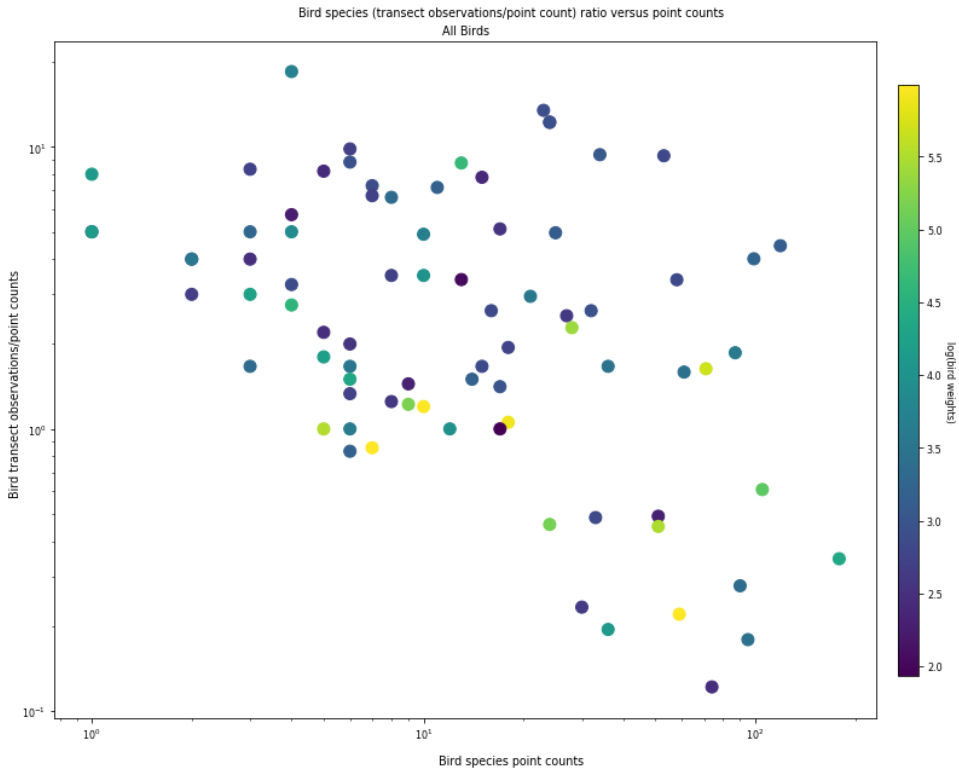

Figure 11: Per-family (transect observation/point count) ratio versus point count for combined data.

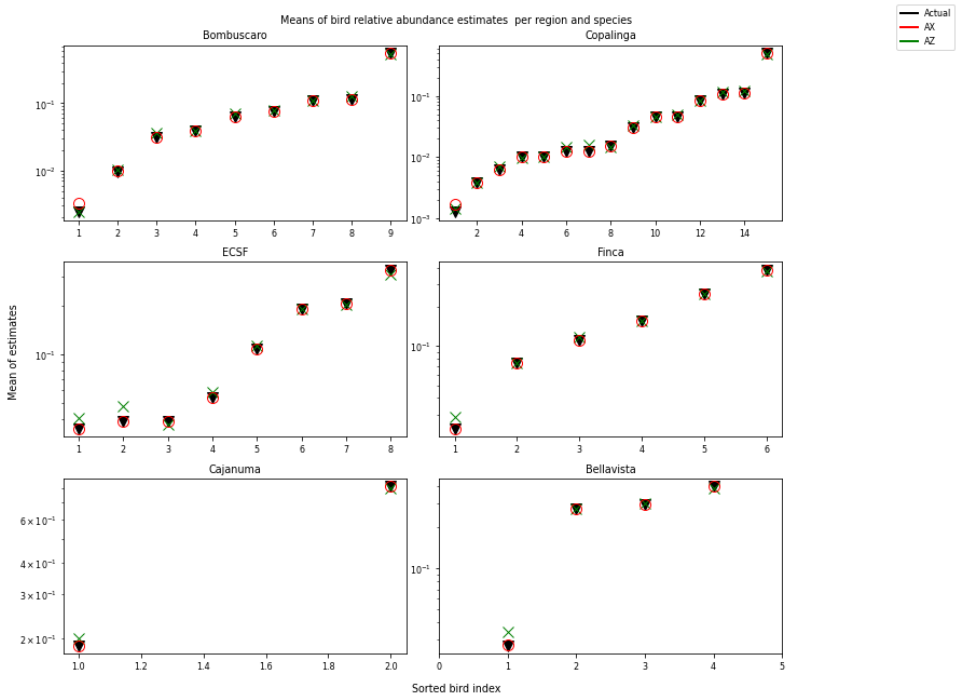

Figure 12: Mean $(\widehat{a}_n)_z$ and $(\widehat{a}_n)_x$ compared to true relative abundance for different bird families in different regions. Bird families in each region are sorted in increasing order of relative abundance.

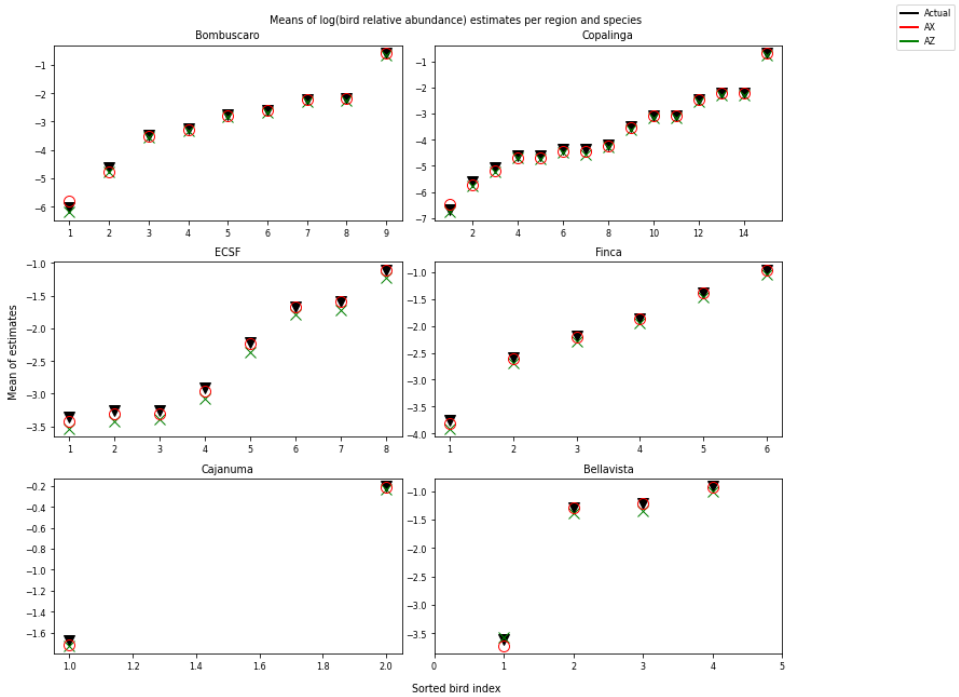

Figure 13: Mean $(\widehat{\ell_n})_z$ and $(\widehat{\ell_n})_x$ compared to true log relative abundance for different bird families in different regions. Bird families in each region are sorted in increasing order of relative abundance.

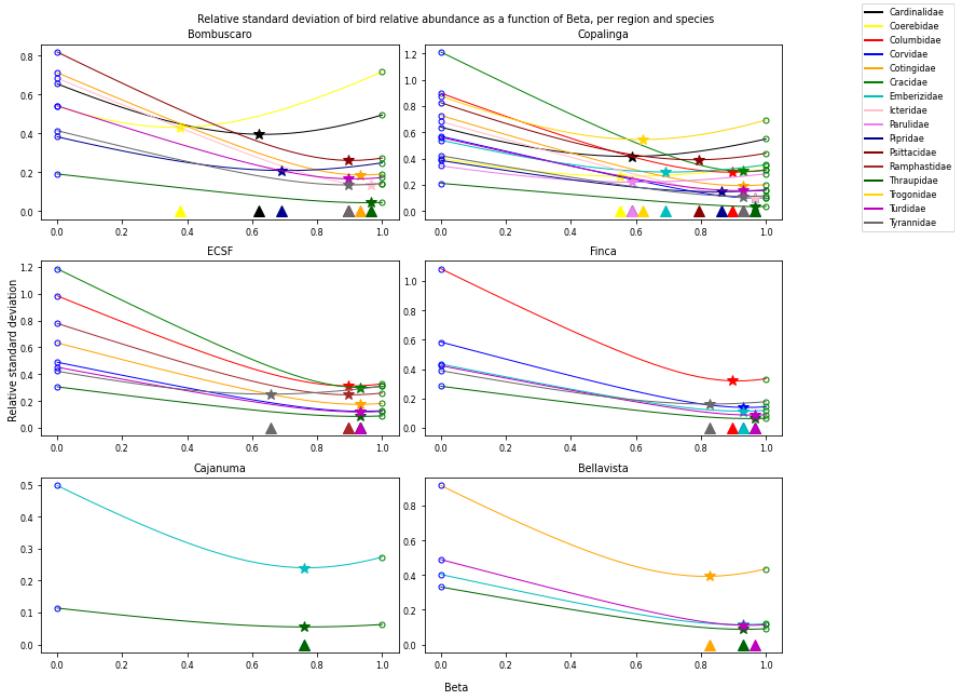

Figure 14: Relative standard deviation (standard deviation divided by true value) of $(\widehat{a}_n)_\beta$ for $0 \leq \beta \leq 1$, for different bird families in different regions. $\beta = 0$ and $\beta = 1$ correspond to $(\widehat{a}_n)_X$ and $(\widehat{a}_n)_Z$, respectively. The stars locate the minima of the different standard deviation curves, while the corresponding triangles on the x axis indicate the value of β corresponding to the minimum value.

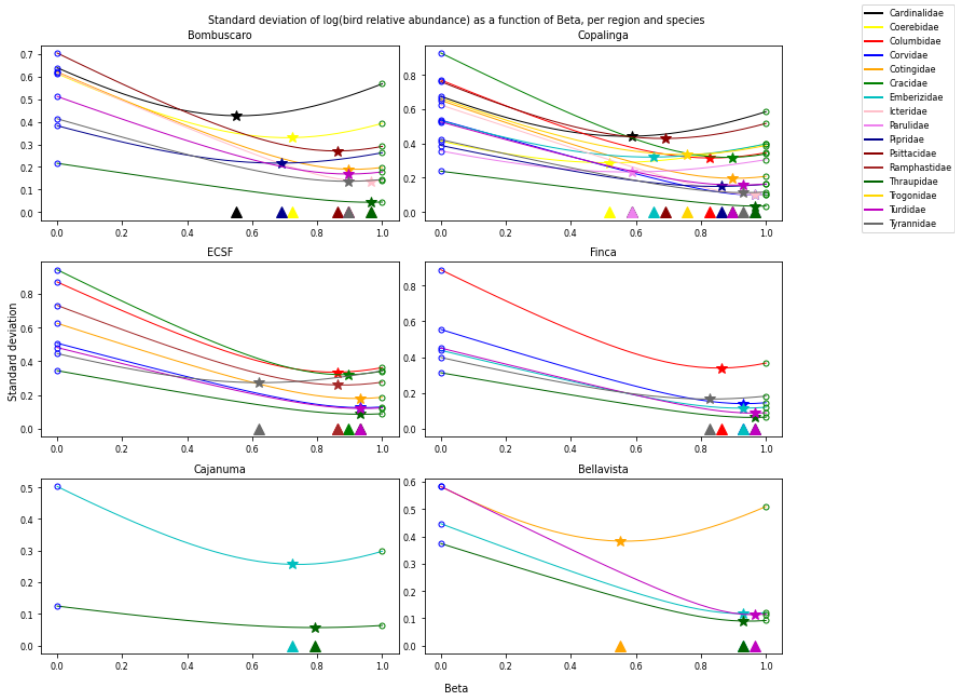

Figure 15: Standard deviation of $(\widehat{\ell}_n)_\beta$ for $0 \leq \beta \leq 1$, for different bird families in different regions. $\beta = 0$ and $\beta = 1$ correspond to $(\widehat{\ell}_n)_X$ and $(\widehat{\ell}_n)_Z$, respectively. The stars locate the minima of the different standard deviation curves, while the corresponding triangles on the x axis indicate the value of β corresponding to the minimum value.

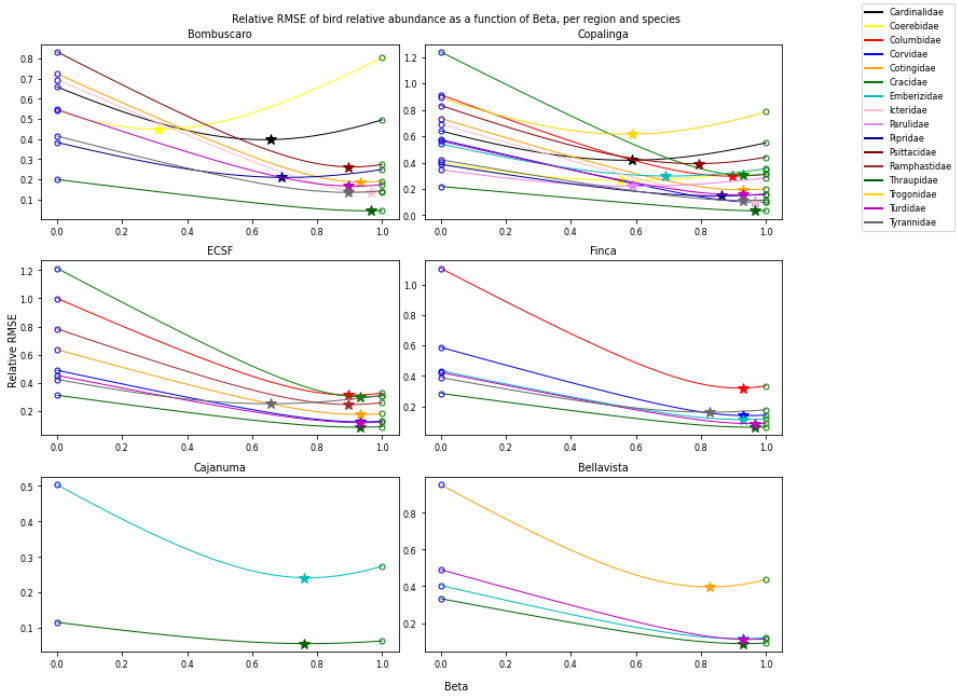

Figure 16: Relative root mean squared error (RMSE) (RMSE / true value) of $(\widehat{a}_n)_\beta$ for $0 \leq \beta \leq 1$, for different bird families in different regions. $\beta = 0$ and $\beta = 1$ correspond to $(\widehat{a}_n)_X$ and $(\widehat{a}_n)_Z$, respectively. The stars locate the minima of the different standard deviation curves, while the corresponding triangles on the x axis indicate the value of β corresponding to the minimum value.

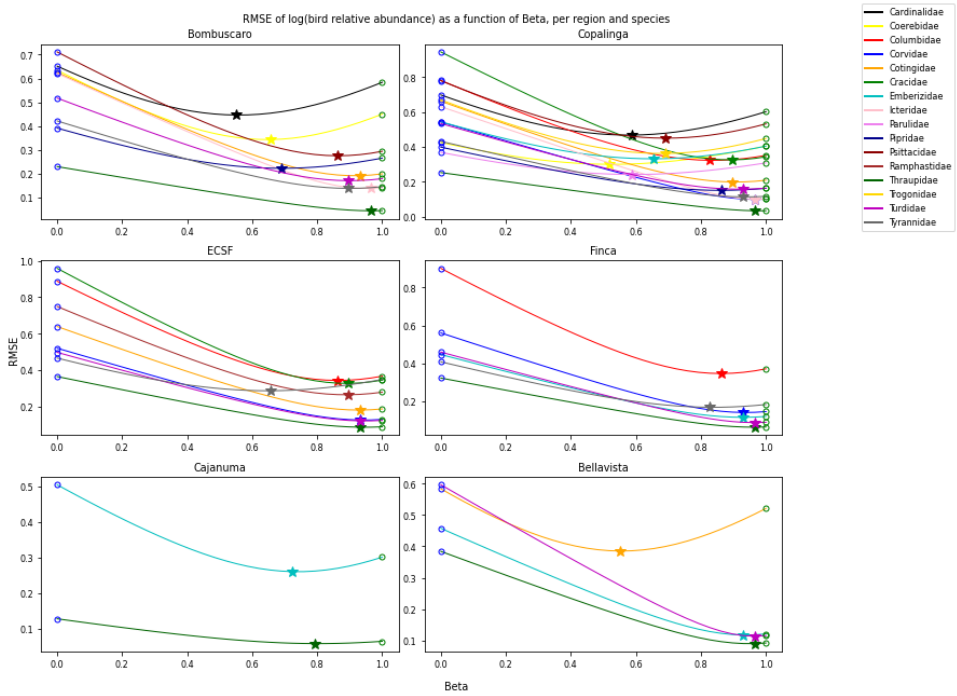

Figure 17: Root mean squared error (RMSE) of $(\widehat{\ell}_n)_\beta$ for $0 \leq \beta \leq 1$, for different bird families in different regions. $\beta = 0$ and $\beta = 1$ correspond to $(\widehat{\ell}_n)_X$ and $(\widehat{\ell}_n)_Z$, respectively. The stars locate the minima of the different standard deviation curves, while the corresponding triangles on the x axis indicate the value of β corresponding to the minimum value.

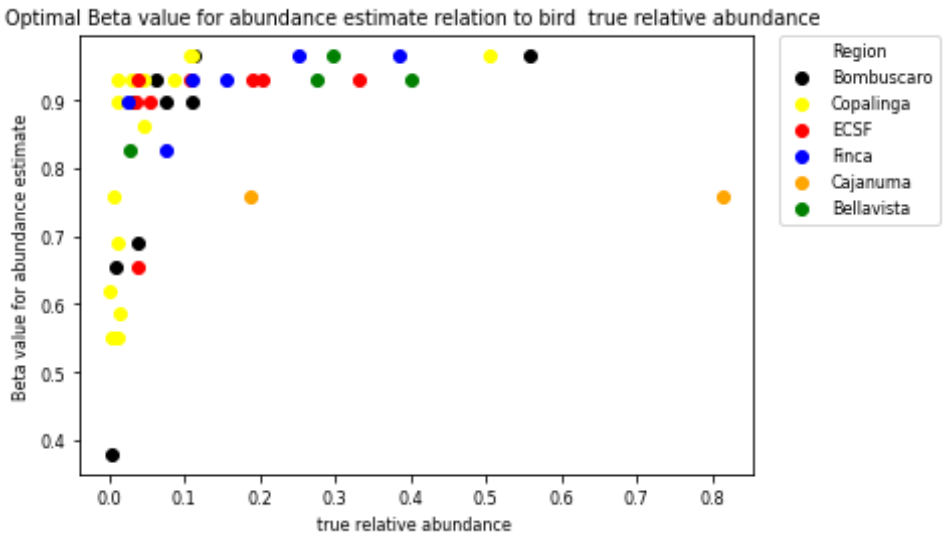

Figure 18: Optimal value of β (i.e. value of β that minimizes RMSE error) for $(\widehat{a}_n)_\beta$ for $0 \leq \beta \leq 1$ as a function of bird family true relative abundance.

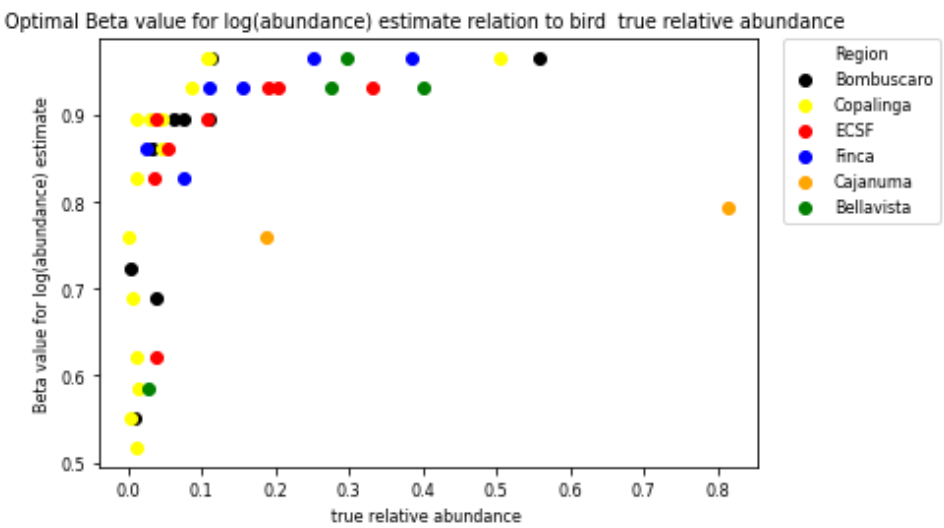

Figure 19: Optimal value of β (i.e. value of β that minimizes RMSE error) for $(\widehat{\ell}_n)_\beta$ for $0 \leq \beta \leq 1$ as a function of bird family true relative abundance.

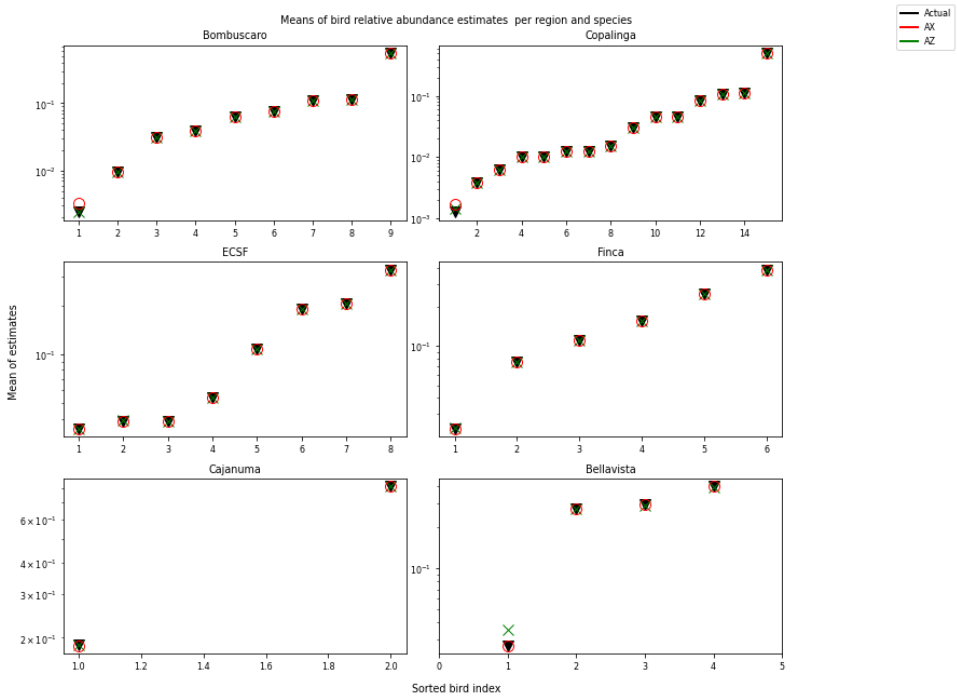

Figure 20: Mean $(\widehat{a}_n)_z$ and $(\widehat{a}_n)_x$ compared to true relative abundance for different bird families in different regions, for ecologists with exact information about p_b. Bird families in each region are sorted in increasing order of relative abundance.

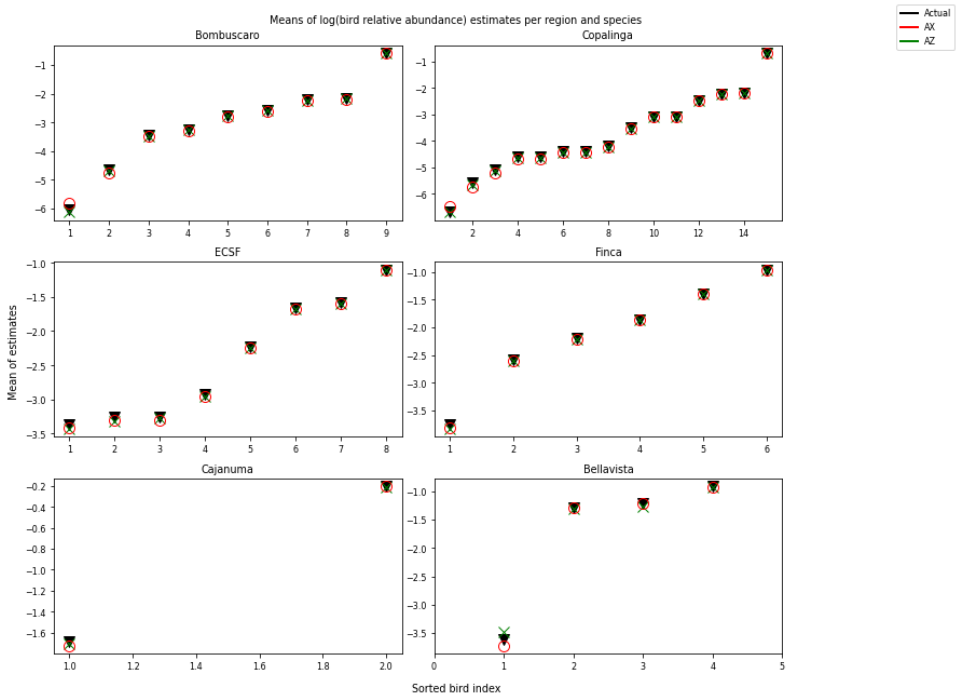

Figure 21: Mean $(\widehat{\ell_n})_z$ and $(\widehat{\ell_n})_x$ compared to true log relative abundance for different bird families in different regions, for ecologists with exact information about p_b. Bird families in each region are sorted in increasing order of relative abundance.

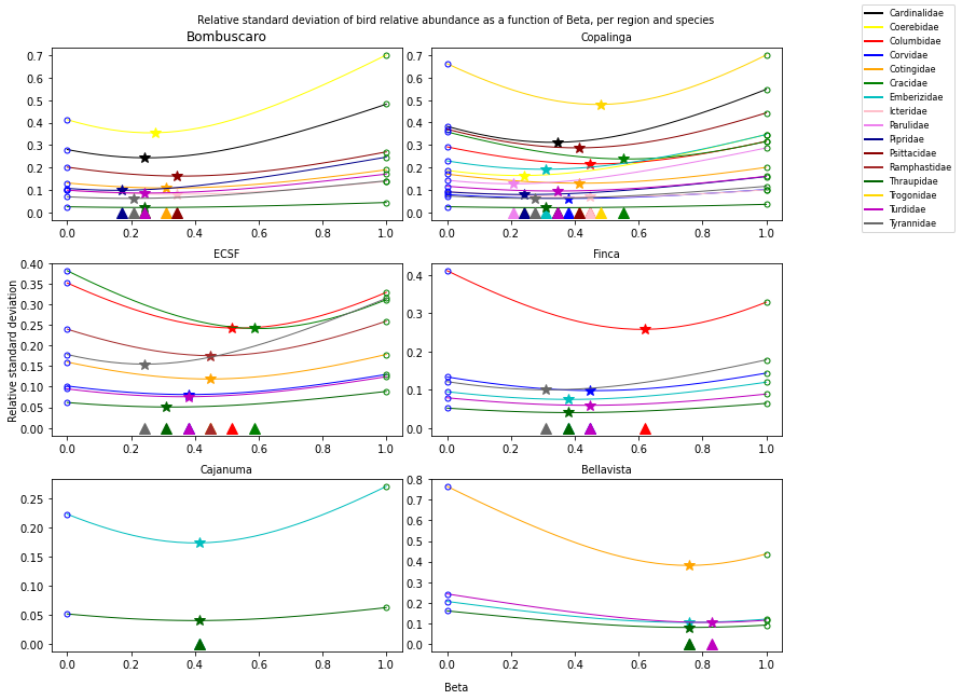

Figure 22: Relative standard deviation (standard deviation divided by true value) of $(\widehat{a}_n)_\beta$ for $0 \leq \beta \leq 1$, for different bird families in different regions. $\beta = 0$ and $\beta = 1$ correspond to $(\widehat{a}_n)_X$ and $(\widehat{a}_n)_Z$, respectively, for ecologists with exact information about p_b.

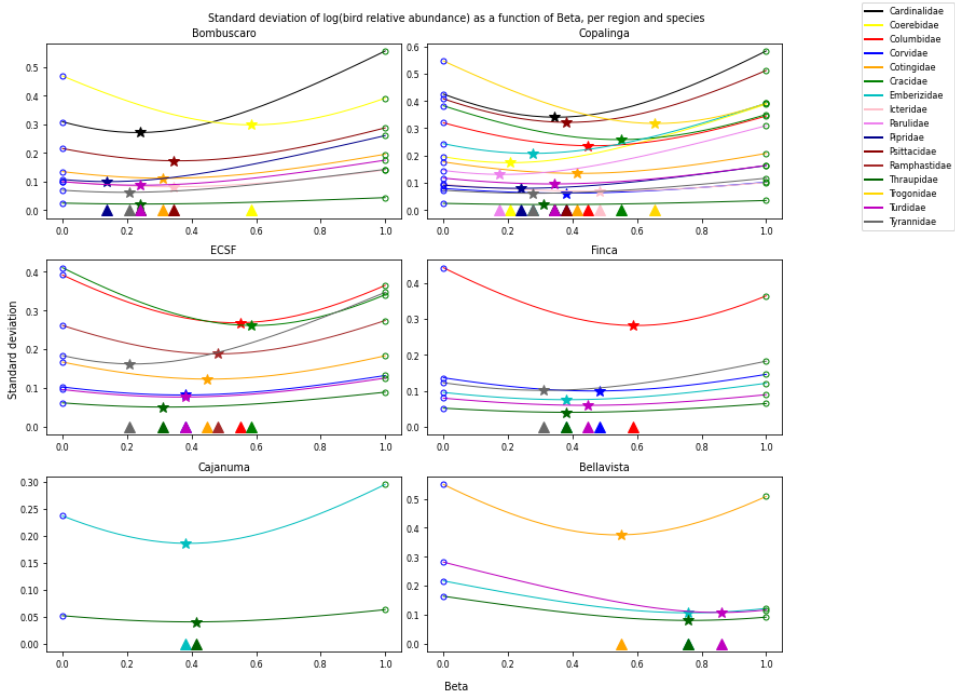

Figure 23: Standard deviation of $(\widehat{\ell}_n)_\beta$ for $0 \leq \beta \leq 1$, for different bird families in different regions. $\beta = 0$ and $\beta = 1$ correspond to $(\widehat{\ell}_n)_X$ and $(\widehat{\ell}_n)_Z$, respectively, for ecologists with exact information about p_b.

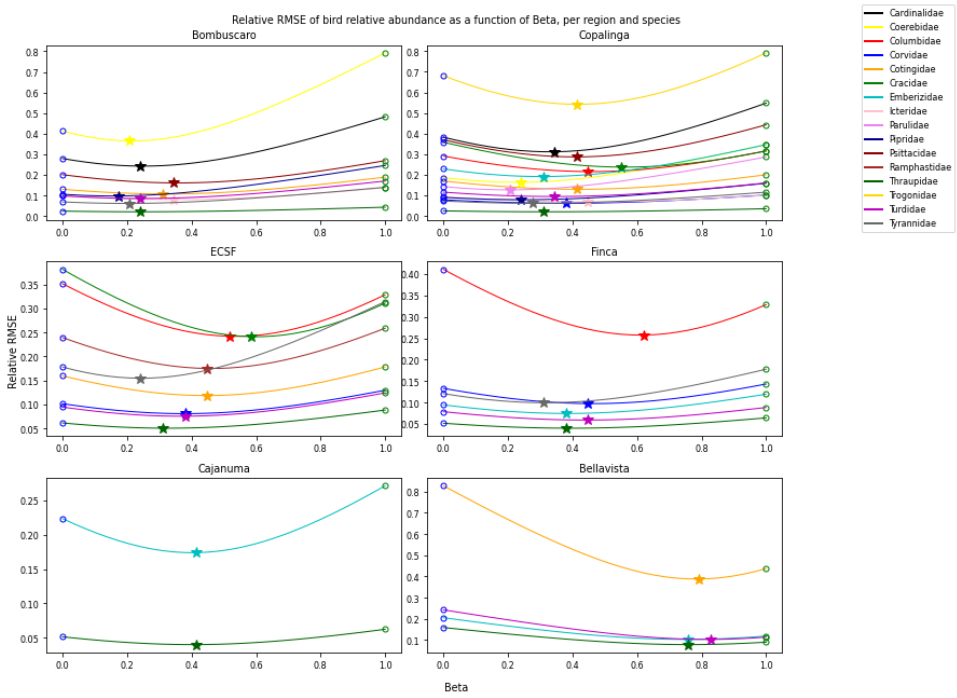

Figure 24: Relative root mean squared error (RMSE) (RMSE / true value) of $(\widehat{a}_n)_\beta$ for $0 \leq \beta \leq 1$, for different bird families in different regions. $\beta = 0$ and $\beta = 1$ correspond to $(\widehat{a}_n)_X$ and $(\widehat{a}_n)_Z$, respectively, for ecologists with exact information about p_b.

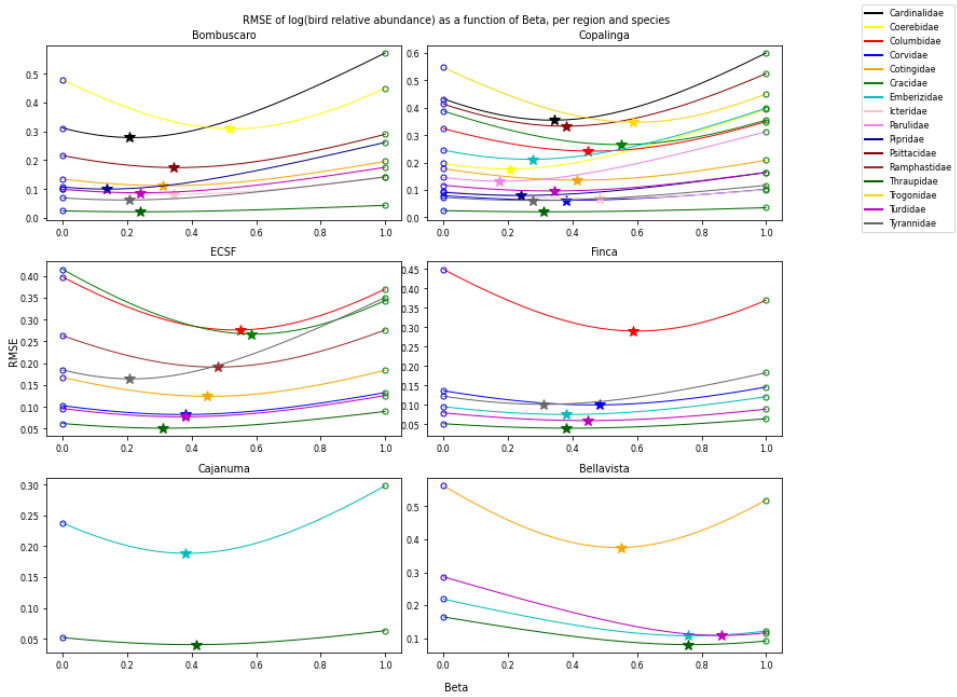

Figure 25: Root mean squared error (RMSE) of $(\widehat{\ell}_n)_\beta$ for $0 \leq \beta \leq 1$, for different bird families in different regions. $\beta = 0$ and $\beta = 1$ correspond to $(\widehat{\ell}_n)_X$ and $(\widehat{\ell}_n)_Z$, respectively, for ecologists with exact information about p_b

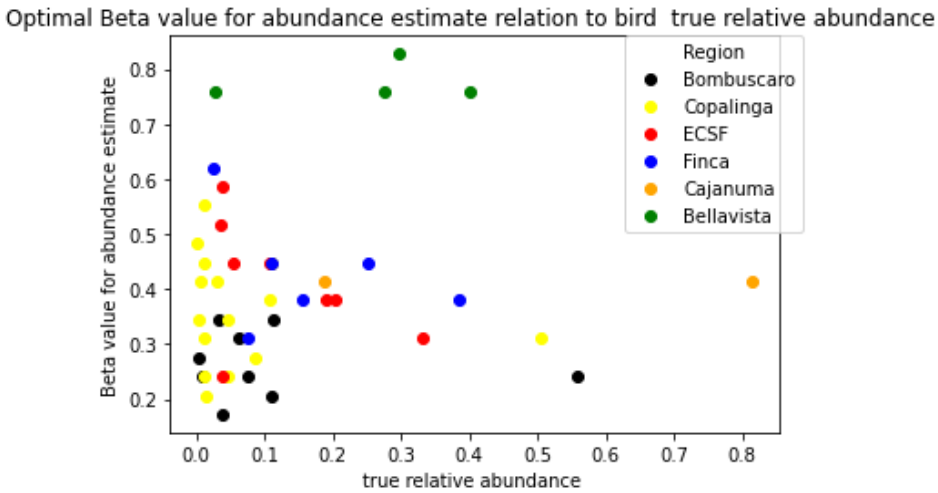

Figure 26: Optimal value of β (i.e. value of β that minimizes RMSE error) for $(\widehat{a}_n)_\beta$ for $0 \leq \beta \leq 1$ as a function of bird family true relative abundance, for ecologists with exact information about p_b.

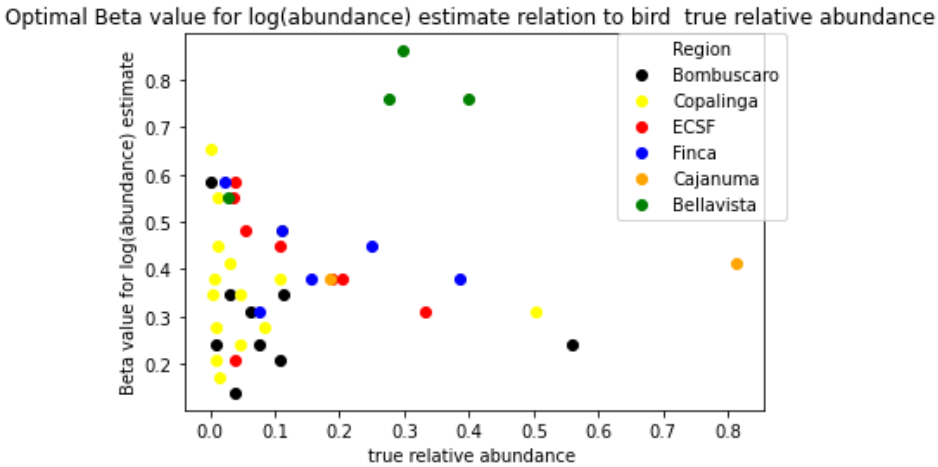

Figure 27: Optimal value of β (i.e. value of β that minimizes RMSE error) for $(\widehat{\ell}_n)_\beta$ for $0 \leq \beta \leq 1$ as a function of bird family true relative abundance, for ecologists with exact information about p_b.

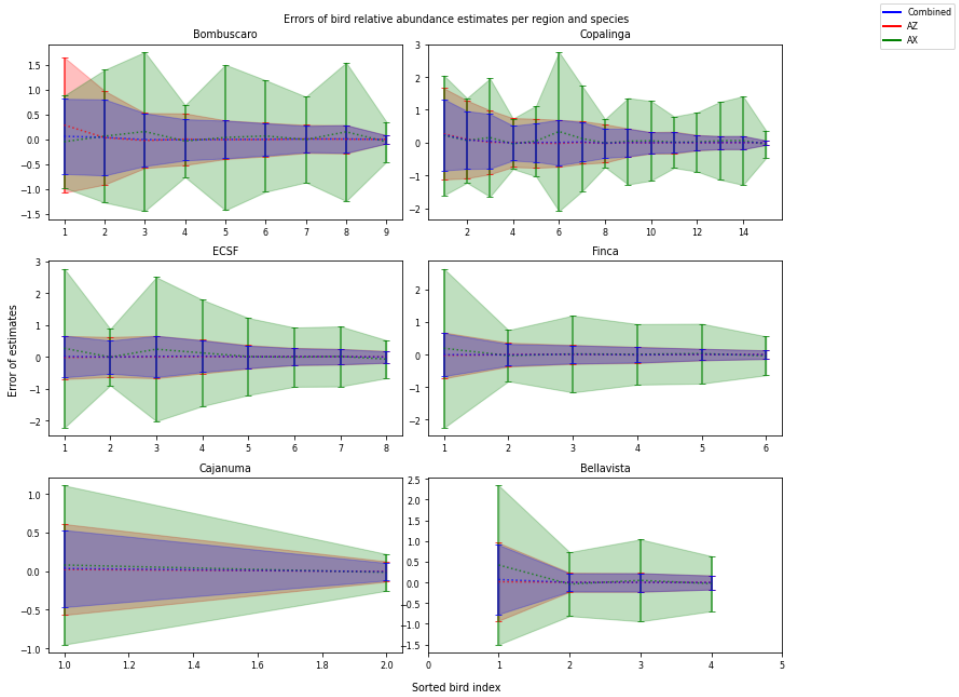

Figure 28: 95% confidence intervals for relative errors (error divided by true value) for three estimators of relative abundance (point-count based, transect count based, and combined) for different bird families for scenarios based on data from six different regions. The transect count based estimator is calculated using $\overline{p}_b$ as in (20). For each region, families are ordered in order of increasing abundance.

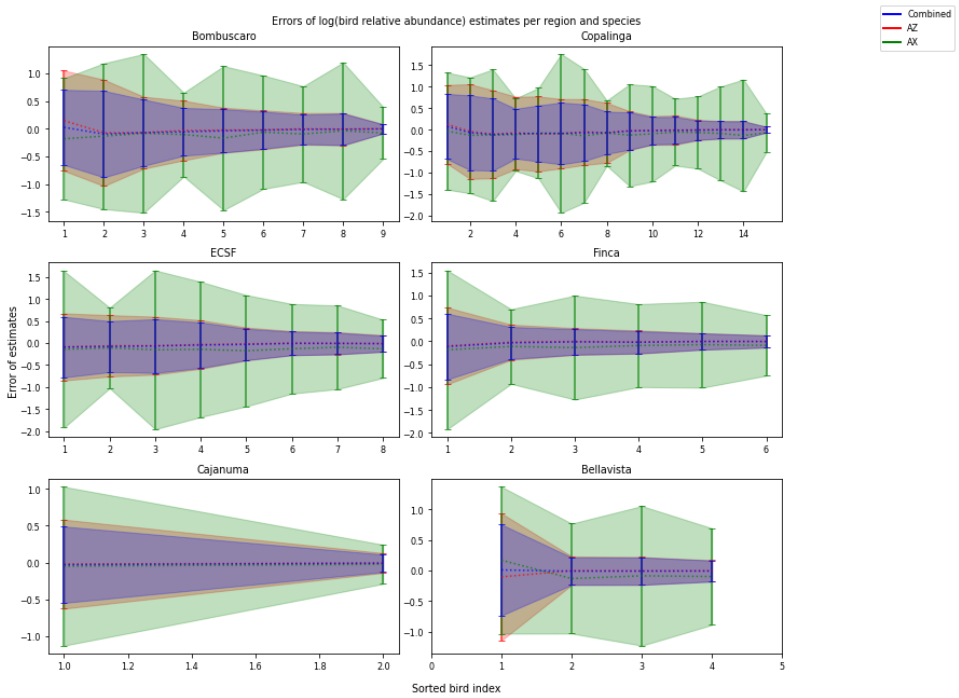

Figure 29: 95% confidence intervals for three estimators of logarithm of relative abundance (point-count based, transect count based, and combined) for different bird families for scenarios based on data from six different regions. The transect count based estimator is calculated using $\overline{p}_b$ as in (20). For each region, families are ordered in order of increasing abundance.

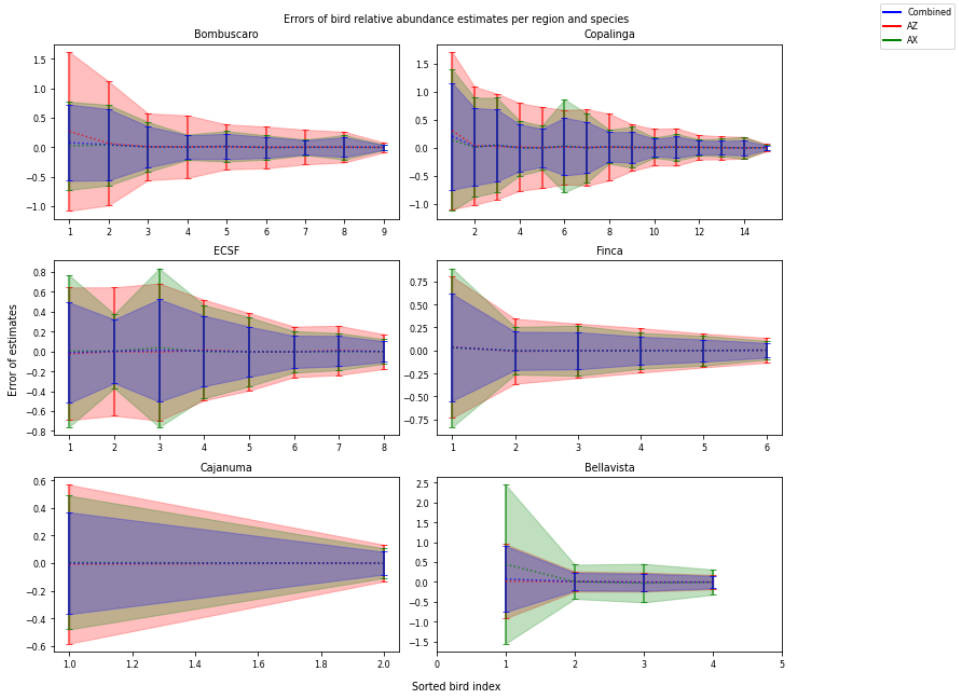

Figure 30: 95% confidence intervals for relative errors (error divided by true value) for three estimators of relative abundance (point-count based, transect count based, and combined) for different bird families for scenarios based on data from six different regions. The transect count based estimator is calculated using the exact value of p'_{bn}. For each region, families are ordered in order of increasing abundance.

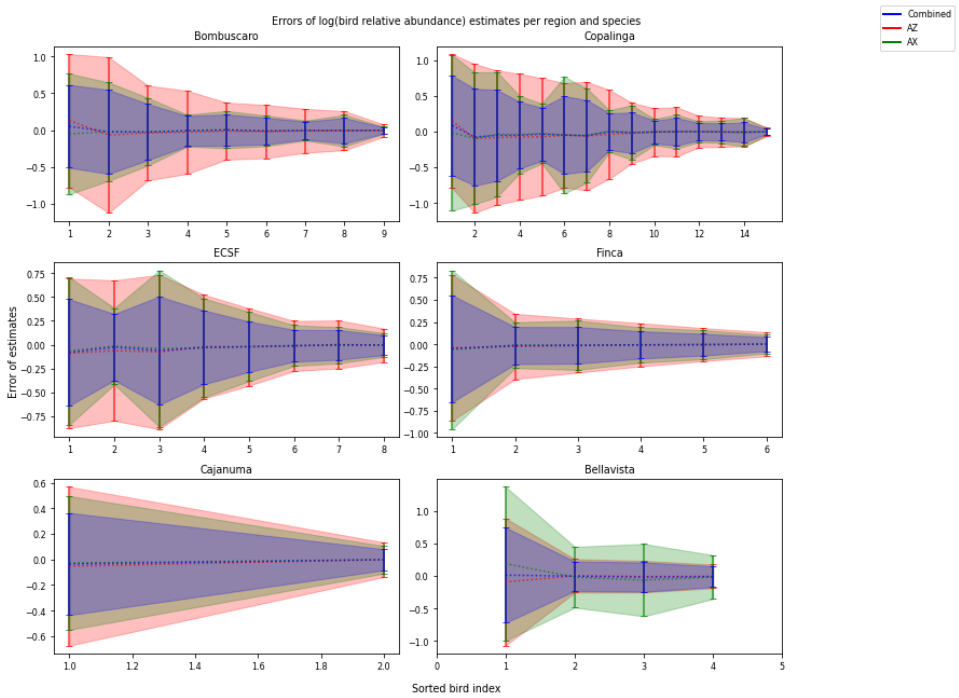

Figure 31: 95% confidence intervals for three estimators of logarithm of relative abundance (point-count based, transect count based, and combined) for different bird families for scenarios based on data from six different regions. The transect count based estimator is calculated using the exact value of p'_{bn} For each region, families are ordered in order of increasing abundance.