Modeling Detection Bias of Mobile Species

Adi Chowers and Micha Mandel
The Hebrew University of Jerusalem

1.Introduction

Species distribution models (SMDs) estimate species habitats using environmental data. A key challenge is imperfect detection, where a species may be present but undetected. False absences introduce bias into model predictions. Accounting for detectability is essential for reliable SDMs inference.

2. Zero Inflated Model

Hierarchical zero-inflated binomial (ZIB) models are commonly used to incorporate imperfect detection in presence-absence studies. These models assume two processes (over N sites):

- Y_i - Presence of an individual at site i
- D_i- Detection of an individual at site i

$$Y_i \sim \text{Bernoulli}(\psi_i)$$
$$D_i | Y_i \sim \text{Bernoulli}(Y_i p_i)$$

With the likelihood:

$$L(\psi_i, p_i) = \Pi_i^N \left[\psi_i p_i^{D_i}(1-p_i)^{1-D_i} + I(D_i = 0)(1-\psi_i)\right]$$

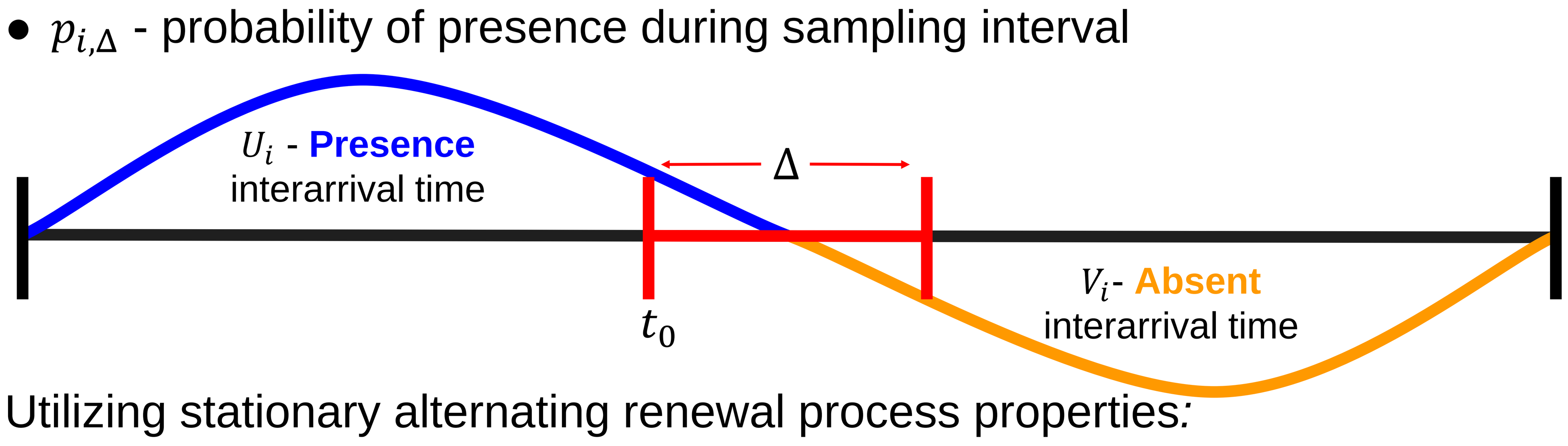
Where ψ_i and p_i can be modeled as functions of environmental variables

- Repeated surveys increase detection probability, however they are logistically complex.
- One visit is problematic for mobile species, since presence is often dynamic, leading to temporal absences rather than true site-level absence.

3. Alternating Renewal Process (ARP)

We model species presence at site i using an alternating renewal process:

- $U_i \sim F_{U_i}(u)$ - Presence interarrival time, with $E[U_i] = \mu_{U_i}$
- $V_i \sim F_{V_i}(v)$ - Absence interarrival time, with $E[V_i] = \mu_{V_i}$
- Sampling interval between t_0 and $t_0 + \Delta$
- $p_{i,\Delta}$ - probability of presence during sampling interval



Utilizing stationary alternating renewal process properties:

$$D_i \sim \text{Bernoulli}\left(p_{i,\Delta}\right)$$
$$L(p_i) = \Pi_i^N \left[p_{i,\Delta}^{D_i}\left(1-p_{i,\Delta}\right)^{1-D_i}\right]$$
$$p_{i,\Delta} = \frac{\mu_{U_i}}{\mu_{U_i}+\mu_{V_i}} + \frac{\int_0^\Delta 1 - F_{V_i}(v)dv}{\mu_{U_i}+\mu_{V_i}}$$

Estimation using MLE with regularization. For regression purposes we model $\phi_i = \frac{\mu_{U_i}}{\mu_{U_i}+\mu_{V_i}}$, and F_{V_i} as a function of environmental covariates.

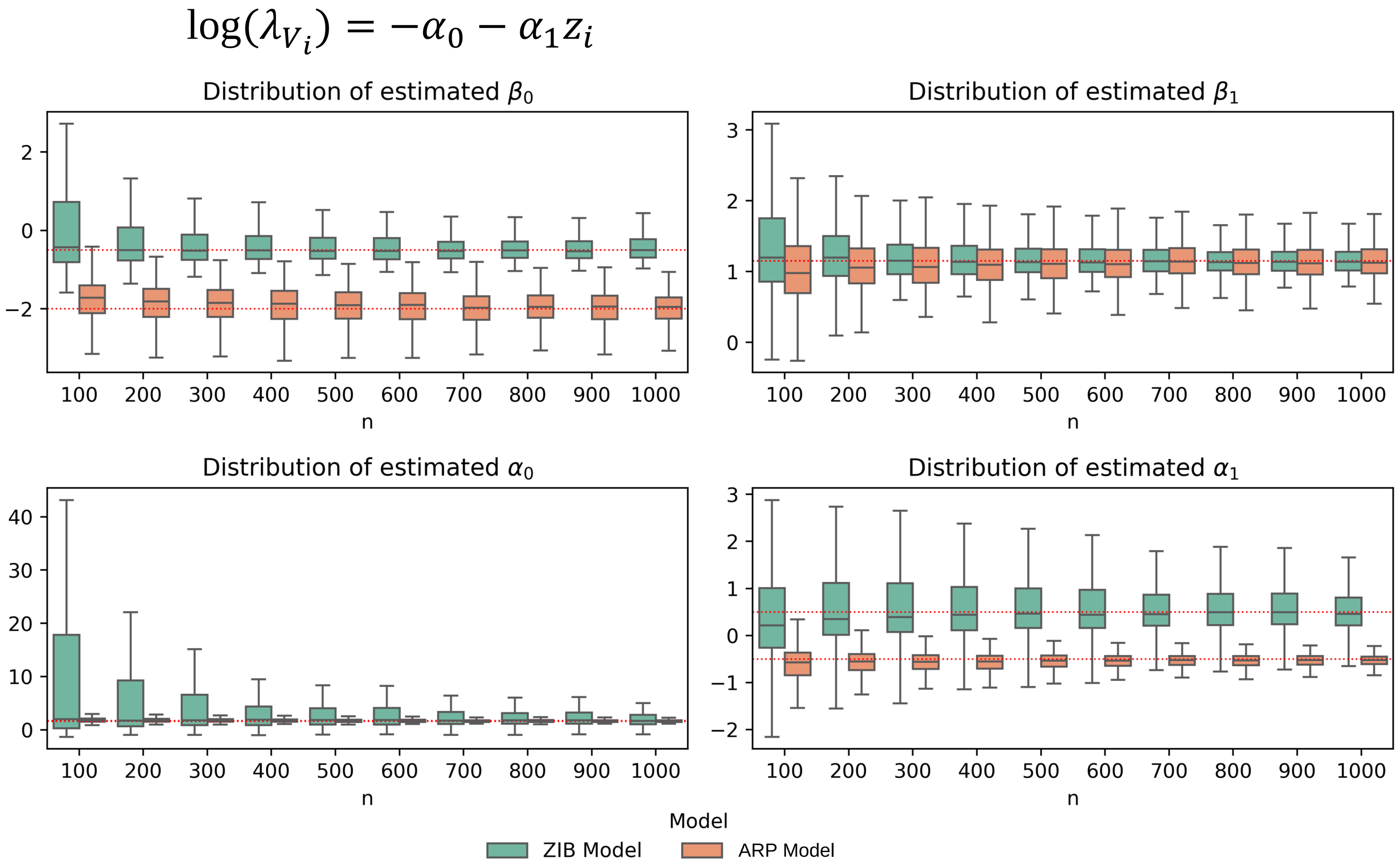
4. Simulation

ARP model

$$U_i \sim \text{Exp}(\lambda_{U_i}), V_i \sim \text{Exp}\left(\lambda_{V_i}\right)$$
$$p_i = 1-(1-\phi_i)(1-F_{V_i})$$
$$\text{logit}(\phi_i) = \beta_0 + \beta_1 x_i$$
$$\log\left(1 - F_{V_i}(\Delta)\right) = -\lambda_{V_i}\Delta$$
$$\log(\lambda_{V_i}) = -\alpha_0 - \alpha_1 z_i$$

ZIB model

$$\text{logit}(\psi_i) = \beta_0 + \beta_1 x_i$$
$$\text{logit}(p_i) = \alpha_0 + \alpha_1 z_i$$

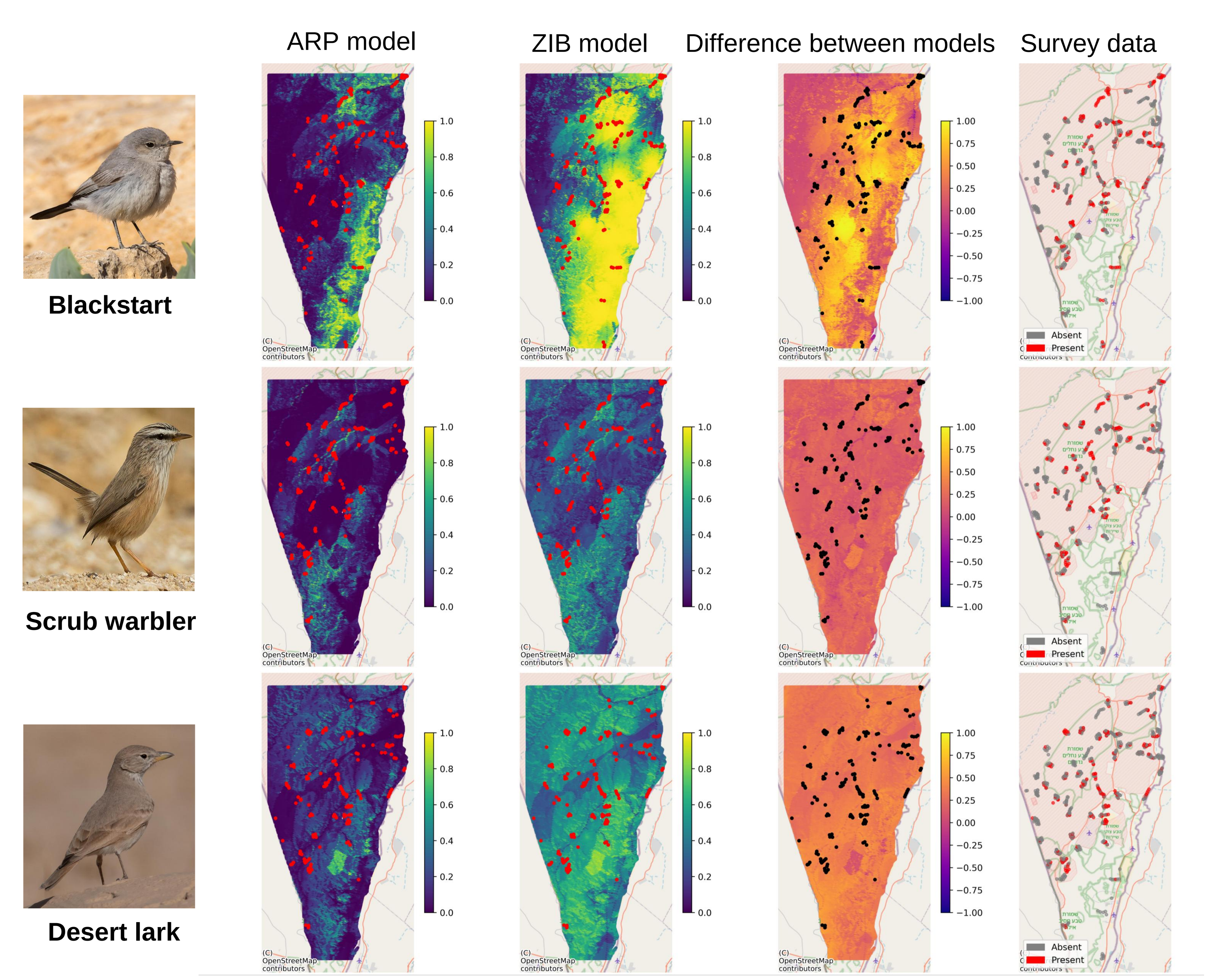


The red dotted lines are the real parameters' value (the values of the parameters were selected to yield approximately the same detection probability in both models).

The simulation results shows that the ARP model performs better than the ZIB model, under the single-visit setting.

5. Birds' distribution in southern Israel

We apply the methods to three bird species based on survey data that were collected by the Eilat Birding Center (2018–2020) in southern Israel. Prediction is based on 11 environmental covariates



The red/black points = presence, grey = absence.

The results are mixed: ZIB model overestimates presence for Blackstart, assigning probability 1 to many areas, while the ARP model underestimates areas where the Scrub warbler is clearly observed.

6. Discussion

- The ARP model outperforms the ZIB model, but for repeated visits (not shown) ZIB model may be better.
- The model enables mean presence time estimation.
- The maps assists in identifing sensitive areas for nature conservation.

Acknowledgments

- GIS Center, the Hebrew University of Jerusalem
- CIDR - HUJI Center for interdisciplinary Data Science Research
- Israel science foundation