# Title *Myioborus* … (Aves: Parulidae)

Sebastián Gallan a, Cornelio A. Bota-Sierra a,b, Jorge Lizarazo c,d , and Gustavo A. Londoño e

a Facultad de Ciencias Exactas y Naturales, Universidad de Antioquia, Universidad de Antioquia, Colombia.

b Grupo de Entomología Universidad de Antioquia (GEUA), Universidad de Antioquia, Colombia.

c Departamento de Humanidad, Tecnología, y Ciencia (HTC), Facultad de Ciencias Humanas, Universidad Icesi, Cali, Colombia.

d Department of Biology, Faculty of Science, Western University, London, Ontario, Canada N6A 5B7.

e Facultad Barberi de Ingeniería, Diseño y Ciencias Aplicadas, Universidad Icesi, Cali, Colombia.

Corresponding authors: Sebastian Gallan sebastiangallan31@gmail.com and Gustavo Londoño galondono@icesi.edu.co

# Abstract

### Título:

### Resumen

# Introduction

# Methods

## Study area

## Egg and Nestlings collection data

## Data analysis

All analyses were conducted in Python (v3.10–3.11) using PyCharm (JetBrains; project-based virtual environments) and Jupyter notebooks for exploratory steps. Core libraries included pandas (data wrangling), *numpy* (numeric), *scipy* (statistics), *scikit-learn* (modeling and validation), and *matplotlib* (graphics).

### Incubation behaviors recognition and cauterization

We analyzed incubation behavior in *Myioborus miniatus*, *M. melanocephalus*, and *M. ornatus* using nest temperature time series collected from multiple sites in Colombia and Peru. All files were standardized and processed using the BoutScout pipeline and deep learning model (Lizarazo & Londoño, in print), which performs automated cleaning, 24-hour segmentation (1,440 minutes), and transformation of time into a cyclical feature via sin(2π·minutes/1440). Sequences were passed into its pre-trained Bidirectional Long Short-Term Memory (BiLSTM) model that assigns a behavioral state—on-bout, off-bout, or nocturnal—to each minute. Model outputs were compiled with metadata including the source file, day index, and temperature values.

Post-inference, we filtered out sequences with sensor errors, misaligned timestamps, or implausible behavioral patterns. To quantify daily incubation strategies, we computed a set of behavioral metrics per nest-day. These included the duration of the diurnal period (defined as the start from the first off-bout to the start of the final nocturnal bout), nest attentiveness (proportion of time spent in on-bouts during the diurnal period), total and average on-, off-, and nocturnal bout durations, bout counts. These summaries were merged with daily mean ambient temperature and enriched with morphological attributes, including elevation, egg mass, length, and width, included at the end a full set of 15 variables.

Through exploratory analysis and correlation assessment, retained a subset of nine variables showing the strongest ecological interpretability and non-redundant structure. These final features included, nest attentiveness, total on-bout duration, duration of the active diurnal period, average on-, off-, and nocturnal bout durations, daily ambient temperature, elevation, and egg weight.

This refined dataset was z-score normalized using *StandardScaler* (from the *sklearn.preprocessing* module) and used for Principal Component Analysis (PCA) implemented via the PCA class from *sklearn.decomposition*. The first five components explaining >80% of total variance were retained. K-means clustering was applied to the PCA-transformed space, with the optimal number of clusters (k = 2) selected based on silhouette score and inertia across multiple random initializations. Cluster identity was cross-tabulated with species to evaluate the association between behavioral–morphological groupings and taxonomic identity. A contingency table was constructed using *pd.crosstab* in pandas, and a chi-square test of independence was performed using *chi2\_contingency* from *scipy.stats* to assess whether the distribution of species differed significantly across clusters. To characterize the behavioral and ecological composition of each cluster, we computed the mean and standard deviation of each variable included in the PCA input for all individuals within each group to facilitate ecological interpretation and cross-cluster comparison.

### Growth-curve modeling

We estimated missing post-hatch ages from *Myioborus miniatus* and *M. melanocephalus* using a species-specific Decision Tree Regressor (scikit-learn) that maps body measurements to age. Age in days was the response; predictors were body mass and tarsus length (Supplementary S2). Model fitting was restricted to records with complete age and predictors. For each species, we fit a *DecisionTreeRegressor* (max\_depth = 3, random\_state = 42). Train/test splits were 85/15 for *M. miniatus* (N= 232, RMSE=1.01, MAE= 0.80, R2 =0.84) and 90/10 for *M. melanocephalus* (N= 16, RMSE =1.28, MAE= 1.25, R2 =0.93). After training, we imputed age where age was missing and both predictors were present (*M. miniatus*: imputed ages for 30; *M. melanocephalus*: imputed ages for 4), writing predictions into the age column and applying half-up rounding to integer days for reporting. M. ornatus, no imputation was performed owing to one nest with complete nestling data (remaining records contained egg-only information).

We conducted a growth-curve analyses using the *Nestling Growth App* (cite). For each species × trait (mass, wing, tarsus) the app fit a standard panel of biological functions—Logistic, Gompertz, Richards, von Bertalanffy, and EVF—and exported parameter estimates and information-criterion tables (AIC/BIC) alongside figures and their results. We retained the top-ranked model per trait by AIC/BIC; when the Logistic model was not top-ranked, we additionally report its rate parameters K and T0 to facilitate cross-trait/species comparisons.

# Results

Se encontraron tantos nidos de los cuales tantos con huevos frescos, tantos con huevos en desarrollo, tantos con polluelos. Fueron depredados tantos…. Etc.

## Nest

Nidos domo en tal áreas, con este placement en tal parte a tantos metros sobre el suelo… etc. Medidas internas esto, medidas externas exto…..etc materiales fueron ….. diferencias genrales entre los tres y similitudes.

## Incubation behaviors and its ecology

Clustering analysis based on behavioral and morphological traits revealed two distinct incubation strategies across the evaluated *Myioborus* species, which were not structured by taxonomy *per se*. Instead, the clusters reflected variation in incubation behavior within species. Particularly within *M. miniatus*, the most abundant taxon in both groups (75% in each; Table S3).

The first group (Cluster 1; Figure 3) exhibited longer diurnal periods (650 ± 312 min), lower nest attentiveness (0.36 ± 0.17 %), more frequent and prolonged off-bouts (30.57 ± 18.95 min), and shorter nocturnal incubation (401 ± 155 min). This group was also associated with slightly warmer ambient temperatures (17.13 ± 5.85 °C), lower elevation (1882 ± 457 m), and lighter eggs (1.56 ± 0.25 g). In contrast, the second group (Cluster 2) showed higher attentiveness (0.61 ± 0.11), shorter diurnal periods (334 ± 188 min), brief off-bouts (12.67 ± 4.35 min), and prolonged nocturnal incubation (709 ± 57 min). Nests in this group were generally located at higher elevations (1928 ± 556 m), in slightly cooler environments (16.29 ± 4.43 °C), and had marginally heavier egg masses (1.61 ± 0.23 g).

Although the chi-square test indicated a statistically significant association between species and cluster identity (χ² = 18.51, p < 0.001), this was primarily driven by the skewed distributions of the less frequent species. *M. melanocephalus* was strongly concentrated in Cluster 2 (33 of 36 individuals), while *M. ornatus* occurred predominantly in Cluster 1 (6 of 8 individuals). Suggesting that elevation and its associated environmental constraints may play a key factor shaping incubation strategies, with more continuous and thermally buffered incubation observed at higher elevations.

## Nestling growth modeling

Informacion sobre los polluelos de forma cualitativa. Al primer dia tal cosa con downy prehatching…. Cuatro dia tal cosa, octavo dia tal cosa. En promedio el periodo de polluelos duro……etc

Nestling mass followed a shared sigmoidal increase across species, with modest differences in rate. The schedule of growth is strikingly similar (Table S2), peak mass accumulation occurs at 4–5 days post-hatching (logistic 𝑇; Figure 4). *M. miniatus* showed the slowest general rise (best fit: Richards; AIC = −22.11, BIC = −5.47; logistic 𝑘 = 0.466, 𝑇 = 4.4 day; Figure 4(a)), whereas *M. melanocephalus* was faster (best fit: logistic; 𝑘 = 0.647, 𝑇 = 3.9 day; Figure 4(b)) and *M. ornatus* was similarly fast (best fit: logistic 𝑘 = 0.609, 𝑇 = 4.7 day; Figure 2(c)).

For wing (Table S3), *M. miniatus* was best described by a logistic curve (𝑘 = 0.330, 𝑇 = 6.8 day; Figure 4(a)). *M. melanocephalus* favored Gompertz (AIC = 62.70, BIC = 66.10; logistic for comparison: 𝑘 = 0.34, 𝑇 = 7.1 day; Figure 4(b)). *M. ornatus* showed a tie among logistic/Gompertz/von Bertalanffy (ΔAIC ≈ 0; AIC = −11.95; logistic: 𝑘 = 0.45, 𝑇 = 6.6 day; Figure 4(d)). Taken together, *ornatus* exhibits the fastest apparent wing elongation (highest logistic 𝑘), with *miniatus* and *melanocephalus* similar and slightly slower. The timing of peak wing growth (logistic 𝑇) clustered at 6–7 days post-hatching.

For tarsus (Table S3), *M. miniatus* was best fit by logistic (𝑘 = 0.303, 𝑇 = 3.548 day; Figure 4(a)). *M. melanocephalus* slightly favored Gompertz (AIC = 25.22, BIC = 28.63; logistic for comparison: 𝑘 = 0.39, 𝑇 = 3.9 day; Figure 4(b)). *M. ornatus* again showed a tie (logistic/Gompertz, ΔAIC ≈ 0; AIC = −6.40; logistic: 𝑘 = 0.16, 𝑇 = 11.9 day; Figure 4(d)). Consistent with these rates, *melanocephalus* displays the fastest tarsus growth, *miniatus* is intermediate, and *ornatus* is slowest and markedly later in timing. Peak tarsus growth (logistic T) occurred early (3.5–4.0 days) in miniatus and melanocephalus, but was delayed to ~12 days in ornatus.

# Discussion

# Funding

Field work was supported by the National Science Foundation [Grant number DEB-1120682​].

# Acknowledgements

# References

# Figure Legend

Figure 3. Behavioral–morphological clustering of incubation strategies in *Myioborus* species. Principal Component Analysis (PCA) based on incubation behavior, ambient temperature, elevation, and egg morphological traits with group partitioning based on K-means clustering applied to the first five principal components (explaining 82.0% of total variance). The optimal number of clusters (k = 2) was selected using silhouette score maximization. Cluster 1 (green) and Cluster 2 (purple) represent two distinct incubation strategies, differing in nest attentiveness, diurnal bout structure, and nocturnal incubation duration (Table S2). Cluster centroids are shown as X symbols. Species identity was significantly associated with cluster membership (χ² = 18.51, p < 0.001) driven by the restricted distribution of M. melanocephalus (mostly in Cluster 2) and M. ornatus (mostly in Cluster 1). Nevertheless, *M. miniatus* is distributed across both clusters, highlighting behavioral divergence within species. Shapes represent species identity, circles (*M. melanocephalus*), squares (*M. miniatus*), and diamonds (*M. ornatus*). Watercolor illustrations by Jorge Lizarazo.

Figure 4. Nestling growth in three *Myioborus* species. Panels show body mass (left column) and appendage growth (right column) as a function of days post-hatching for (a) *M. miniatus*, (b) *M. melanocephalus*, and (c) *M. ornatus*. Points are individual observations and the solid lines are the best-fit trend for that species. Model fits follow the best model per species × trait (AIC/BIC; Table S2, S3, S4).



