**Risk-based viable population monitoring from community science data**

**Rapid assessment of local persistence from population trends with community science data**

Orlando Acevedo-Charry1,2,3, \*, José M. Ponciano4, Caroline L. Poli5, Brian M. Jeffrey2, Robert J. Fletcher Jr. 2, María Angela Echeverry-Galvis6, Bette A. Loiselle2,7, Scott K. Robinson3, Miguel A. Acevedo2

1School of Natural Resources and Environmental, University of Florida, Gainesville, FL 32611, USA

2Department of Wildlife Ecology and Conservation, University of Florida, Gainesville, FL 32611, USA

3Florida Museum of Natural History, University of Florida, Gainesville, FL 32611, USA

4Department of Biology, University of Florida, Gainesville, FL 32611, USA

5Florida Fish and Wildlife Conservation Commission (FWC), FWC Fish and Wildlife Research Institute, Saint Petersburg, FL 33701, USA

6Departamento de Ecología y Territorio, Pontificia Universidad Javeriana, Bogotá D.C., Colombia

7Center for Latin American Studies and Tropical Conservation Development Program, University of Florida, Gainesville, FL 32611, USA

**\*Corresponding author.** [**acevedocharry@gmail.com**](mailto:acevedocharry@gmail.com)

**1 | INTRODUCTION**

“*The science of ecology is the study of distribution and* ***abundance***.” Nicholas J. Gotelli (2008)

As ecologists, a pivotal part of our interest relies on quantifying changes in species abundances, a quintessential ecological process to inform conservation actions (Callaghan et al., 2024; Dennis et al., 1991; Staples et al., 2005). Different methods have been used to estimate abundance (Dénes et al., 2015), abundance rate of change (Bart et al., 2003; Fink, Johnston, et al., 2023), and thus inferring population trends for conservation (Staples et al., 2005). For example, a hierarchical Bayesian model was used to report the net loss of ~3 billion birds during the last five decades in the continental United States (U.S.) and Canada, combining monitoring data from long-term breeding counts and whether radars for resident and migrant birds (Rosenberg et al., 2019). However, such long-term and large-scale structured monitoring efforts are not available elsewhere (Robinson et al., 2021), leading to uncertainty of population dynamics for most biodiversity on the planet (Lees et al., 2022; Neate-Clegg et al., 2020).

Keep conducting standardized long-term monitoring is essential to understand population dynamics (Blake & Loiselle, 2024; Neate‐Clegg et al., 2021). In fact, long-term monitoring is the basic ingredient to robust methods in conservation, like population viability analysis that aims to understand what has happened with the monitored populations. It has been suggested, however, that conservation ecologists could ask rather how the population’s status is changing during the monitoring program, for instance assessing the risk of decline below a prespecified threshold (Staples et al., 2005). Within short times, the dynamic of the population monitored might be assessed from the complement of the risk of decline (), which we define here as local persistence probability (). An important conundrum is that two simultaneous but independent sources of variation occur during population dynamic assessment: the ecological and the observational processes (Dénes et al., 2015; Dennis et al., 2006; Humbert et al., 2009).

Estimating change in abundance in time or space is complex, requiring the inclusion of mathematical models in the ecological (state) and observational (sampling) processes. The ecological process can follow deterministic models like textbook examples of exponential or logistic growth (Gotelli, 2008). Unfortunately, most populations do not respond to simple deterministic models, requiring modeling the ecological underlying process with the inclusion of demographic and/or environmental stochasticity (Evans et al., 2023; Fink, Johnston, et al., 2023; Humbert et al., 2009). This “process noise” works independently of the observation process, also named the “observation error” when detection is imperfect (Dénes et al., 2015; Dennis et al., 2006). Different hierarchical models could integrate these population processes. For example, by using “state-space” models on time-series observational data, researchers can estimate the latent values of the population abundance time-series (Auger‐Méthé et al., 2021; Dennis et al., 2006; Evans et al., 2023). True abundance estimations under state-space models can assess independent population parameters like the growth rate, density dependence (when present), as well as stochastic “process noise” and “observation error” (Dennis et al., 2006; Dennis & Ponciano, 2014; Humbert et al., 2009; Ponciano et al., 2009). These estimated parameters provide robust assessment of population dynamics, which in turns serve for understanding changes of the populations status through time under risk-based viable population monitoring or population viability monitoring (PVM) (see Section 2.2) (Staples et al., 2005).

When abundance data follows discrete and equal sampling, time-series abundance can be fitted with the Gompertz State-Space (GSS) population model (Dennis et al., 2006; Dennis & Ponciano, 2014). This model is a practical way to include density dependence, stochastic process variation, and stochastic observational variation with realistic dynamics of animal populations (Dennis et al., 2006; Evans et al., 2023). Although incorporating replicated sampling into the GSS model improves parameter estimation for population dynamics (Dennis et al., 2010), this improvement in repeated sampling methods still requires equal sampling in discrete time-series data. Special cases of GSS for unequal sampling include the density-independent Exponential Growth (or decay) State Space (EGSS) model and the continuous-time Ornstein-Uhlenbeck State-Space (OUSS) model (see Section 2.1). The EGSS model includes approaches of traditional log-linear regression of counts against time and Brownian motion diffusion processes (Humbert et al., 2009); the first approach assumes that variability in the data occurs only by the observation error noise, while the second approach assumes that populations are censured (without observation error) and variability in the abundance data occurs only by growth rate fluctuations from environmental variability, or ecological process noise (Dennis et al., 1991; Humbert et al., 2009). Similarly, the OUSS allows ecologists to estimate strength of density dependence, process noise, and observation error in data that includes, for example, gaps in sampling observations or missing data (Dennis & Ponciano, 2014). This kind of unequal sampling with gaps and observation error is usual in biodiversity monitoring initiatives that involve volunteer participation during data collection, or community science data.

Community science initiatives (also known as citizen or participative science) involve monitoring by volunteer participation mainly for collecting or reviewing data, becoming an increasing source of biodiversity information worldwide (Amano et al., 2016). From the spectrum of community science initiatives, there is a trade-off between broad participation and having clear objectives, research design, and rigorous protocols during sampling efforts; this is evident between the unstructured and structured community science initiatives (Kelling et al., 2019). For example, iNaturalist ([www.inaturalist.org](http://www.inaturalist.org)) is a flexible global platform that stimulates massive participation of the public to submit incidence occurrence data (Di Cecco et al., 2021) or identify records (Campbell et al., 2023); this initiative has not taxonomic or geographic constraints, albeit absence of sampling event effort information. In contrast, the North American Breeding Bird Survey is an example of a structured community science data, following a rigorous protocol by only highly skilled volunteers for monitoring bird populations in continental North America (Ziolkowski et al., 2023). Semi-structured approaches, somewhere in the middle of the spectrum, provide robust information for biodiversity monitoring while including essential details of the observation process (Kelling et al., 2019). From the semi-structured approaches, eBird is perhaps the most popular for analyzing bird records (Johnston et al., 2021; Kelling et al., 2019; Sullivan et al., 2014). Current estimates of species status and trends from eBird data strictly control for the observation process variability and other confounding environmental factors that affect species detectability or its relationship with the landscape (Fink, Auer, et al., 2023; Johnston et al., 2018). Despite its informative use, these status and trends are yet not available for all species in eBird and fit discrete-time stochastic exponential models as the underlying ecological process of the number of individuals reported and analyzed (Fink, Johnston, et al., 2023).

We aim to provide a rapid assessment of short trends of local persistence probability () by incorporating population dynamics robust models (Dennis & Ponciano, 2014; Humbert et al., 2009) withing the risk-based population viability monitoring framework (Staples et al., 2005) using time series data from eBird. We begin by organizing eBird data records in a time series of weekly high counts estimates, filtering by best practices for bias reduction (Backstrom et al., 2024; Johnston et al., 2021; Kelling et al., 2019; Strimas-Mackey et al., 2023). We selected a species with long-term monitoring efforts that varies in population dynamics in space, including recent expansions of its breeding ranges (Machado-Stredel et al., 2024; Poli et al., 2020). This example species provides benchmark population trends and seasonal dynamics; here, we focused on the new northern limit of the breeding range since 2018. Then, we estimated the local persistence probability for every week since 2020 by simulating trajectories of the population for a period of 2-years (104 weeks). In every week point of abundance estimation, we updated the model fitted, the estimated parameters of the population models for projected trajectories, and the probability to persist above a preestablished threshold after 2 years of stochastic dynamic. We provide an annotated program in R of our approach that could be adjusted and applicable to other species and locations (link code for peer review attached).

**2 | METHODS**

To conduct our analysis, we used the R software v4.3.3 (R Core Team, 2024). Specifically, we filtered the eBird data with the packages *auk* (Strimas-Mackey et al., 2018) and *tidyverse* (Wickham et al., 2019); we generated the hexagonal grid with the package *dggridR* (Barnes, 2023) as suggested in Strimas-Mackey et al. (2023); and we adjusted the R programs published by Humbert et al. (2009) and Dennis & Ponciano (2014), which require the package *MASS* (Venables & Ripley, 2002). In addition to figures using basic R or *ggplot2* (Wickham, 2016) within *tidyverse* (Wickham et al., 2019), we used the package *maps* (Becker et al., 2022) and *gridExtra* (Auguie, 2017) to load maps and generate composite figures. Code is available in FigShare (link available as accepted – code for peer review attached).

**2.1 | The state-space models**

We denoted random variable processes with upper case, while their realization (data) and the constants with lower case. Thus, let be the latent population of unobserved abundance at time (generally year), the logarithmic abundance (), and the sampling estimates of log-abundance at time . The ecological process could have different forms, here we applied two cases: density-independent (EGSS) and density-dependent (OUSS). Both models are a special case of the discrete stochastic GSS population model (Dennis et al., 2006; Dennis & Ponciano, 2014). We computed the Restricted Maximum Likelihood Estimates (REMLEs) for both models, given the improvement in statistical properties when comparing with maximum likelihood estimation (Dennis et al., 2006; Dennis & Ponciano, 2014; Humbert et al., 2009).

**2.2.1 | Exponential Growth State-Space (EGSS)**

We begin with the first model in most population ecology classes, the exponential growth, which could serve as a null hypothesis of density dependence models (Dennis et al., 2006). The exponential stochastic process for is , where defines the finite rate of change (; being the instantaneous rate of change), is the initial population (), and is the environmental stochasticity or process noise at time , which follows a normal distribution with mean and variance (). On the logarithmic scale (), the discrete-time process can be defined as a continuous diffusion process with the stochastic differential equation , where is a random perturbation representing the environmental stochasticity at the Itô log-transformation, or the intensity of noise scaled by , with (Humbert et al., 2009). The EGSS model includes a component in sampling times not equally spaced between times and (), and the log-abundance observed is defined as , where is the observation error and follows a normal distribution with mean and variance (). This model has four unknown parameters: (the trend parameter), (variation of process noise), (variation of observation noise), and (the initial log-abundance population).

**2.2.2 | Ornstein-Uhlenbeck State-Space (OUSS)**

The ecological process in the GSS model for is defined as , where and are constants representing population growth rate and density dependence, and is the environmental stochasticity or process noise (see above and Dennis et al., 2006). This model on the logarithmic scale (), becomes linear and follows an autoregressive model of order 1: , where , and is a constant that represents the strength of density dependence (Dennis et al., 2006; Ponciano et al., 2018). The probability distribution of this logarithmic abundance is normal, with mean and variance changing as a function of time. However, if the strength of density dependence () ranges in absolute values (), the long-run probability distribution of log-abundance approaches a time-independent normal stationary distribution (), with mean and variance .

Note that a special case of this model, when (which means ) is the density independent EGSS (above), without an equilibrium and an additional unknown parameter () (Dennis et al., 2006; Humbert et al., 2009). Also note that the stochastic GSS model does not include the deterministic “carrying capacity”, but a stationary distribution mean representing a long-term expected population size where the population fluctuates with some variation (Dennis et al., 2006). Estimated or observed abundance arose from the unobserved state time-series and observation error as ; . The GSS population model incorporates these principles and has four unknown parameters under stationary distribution: , , , and (Dennis et al., 2006). However, if the observation data commenced before the assumed stationary distribution mean, the initial population become an additional unknown parameter. In case the initial population is known, as in translocation experiments, the observed data for initial population is the same realization of the latent variable () and could be treated as a known parameter (Dennis et al., 2006; Dennis & Ponciano, 2014).

As occurred in the EGSS model, the logarithmic transformation in the GSS discrete model opens the opportunity for estimating the infinitesimal mean and variance under a diffusion process, the Ornstein-Uhlenbeck diffusion process in the OUSS population model for unequal sampling (Dennis & Ponciano, 2014). The OUSS model adds a component in sampling times not equally spaced , where the observation error keeps a normal distribution and the same unknown parameter (), and the underlying unobserved population follows a continuous-time version of the GSS model. With the strength of density dependence parameter () ranging positive values (), the dynamic of the population is stationary. The relationship between parameters from discrete equal sampling GSS model to continuous unequal sampling OUSS model is: , , , and (Dennis & Ponciano, 2014). Assuming a stationary distribution for the initial log-abundance population in the time-series (), the OUSS model has four unknown parameters: , , , and .

These parameters relate with the GSS model parameters in stationary distribution dynamics as: , , , and ; the normal stationary probability distribution has mean and variance (Dennis & Ponciano, 2014). If the initial log-abundance of the population does not meet this assumption (e.g., it is under transition growth), a nonstationary distribution should be modeled with a different maximum likelihood estimation. The normal transition in nonstationary cases has a mean and variance , adding an extra parameter to estimate (). We modeled the dynamic of the populations for density independence (including initial nonstationary distribution) as EGSS, while stationary distributions as OUSS.

**2.2 | Time-series and population viability monitoring from eBird data**

We constructed the time-series eBird data with the maximum count reported per week in sampling units of 100 km2, assuming that this maximum count in eBird is the potential minimum number of individuals in each sampling unit representing populations (Poli et al., 2020). We used Snail Kite (*Rostrhamus sociabilis*) in Florida as our example species. Snail Kite has been extensively monitored in Florida U.S. (Reichert et al., 2020), reflecting spatial movement and lower genetic diversity across wetlands (Poli et al., 2022; Reichert et al., 2020; Robertson et al., 2018). We estimated abundance dynamics from eBird data, accessing records and observer counts through May 2024 from <https://ebird.org/data/download>; based on temporal sampling bias (Backstrom et al., 2024), we concentrate our time-series from 2018 (Supporting Information; Figure SI-1). We constraint the variation of sampling effort in this semi-structured platform following elsewhere recommendations (Backstrom et al., 2024; Johnston et al., 2021; Kelling et al., 2019; Strimas-Mackey et al., 2023). For example, we were only interested in complete checklists from "traveling" or "stationary" protocols within a sampling effort of ≤ 5 hours, ≤ 5 km, ≤ 10 observers, assigning each record to a hexagonal grid of 100 km2 area, assuming population estimation from the maximum observation counts per day in these fixed spatial sampling units (Johnston et al., 2021; Strimas-Mackey et al., 2023). In addition, we limited the records to diurnal events, between 06:00 and 18:00.

Although the high counts per week provide impressive data to estimate parameters, predict latent abundance, and their 95% confidence intervals for time-series at each spatial sampling unit, we applied first the risk-based population viability monitoring for the recently established northern population at the Payne’s Prairie State Park wetland, in Alachua County north-central Florida (Poli et al., 2020). After Hurricane Irma flooded this ~85 km2 wetland system in 2017, Snail kites established a nesting population in 2018, with a steady increase in abundance (Fink, Auer, et al., 2023) tightly correlating weekly high eBird counts with systematic airboat counts and long-term banding monitoring during this year of establishment (Poli et al., 2020). This range extension reflects changes in wetland conditions that matched with resources availability of novel invasive prey (Machado-Stredel et al., 2024). The resource availability by invasive apple snails have also modified morphological traits (Cattau et al., 2017; Machado-Stredel et al., 2024), driving increase in population growth for the northern Snail kites’ population declines in southern populations (Cattau et al., 2016).

To conduct the population viability monitoring (Staples et al., 2005) for Snail kites in Payne’s Prairie with eBird data, we modeled the trend of the population for the first two years of establishment (2018-2019) using the EGSS model (See SI Section S2 for a graphical step by step, Figures SI-2 to SI-5). With the parameters estimated in this first model, we simulated 10000 trajectories for the following two years (104 weeks) and count the number of trajectories that ended below an assumed critical number of individuals of ¾ the mean of abundance observed in eBird between 2018-2024 (). Then, each new observation by week was included, the model fitted, and new projections of trajectories were simulated, storing the probability of declining below our threshold value (*i.e.*, risk to (quasi)extinction). We fitted in each iteration the OUSS, but if the trend parameter , we selected EGSS because it is an indication of density independence ().

**3 | RESULTS**

The eBird record history of Snail Kite in Florida includes 438 sampling units of 100 km2 (figure). After filtering by good practices of sampling effort (Johnston et al., 2021; Strimas-Mackey et al., 2023), we used the time-series for ## sampling units with more than 100 checklists that might reflect different dynamics. For example, a single cell concentrated most of the observations (>7,000). This cell overlaps with most of the Payne’s Prairie State Park wetland



**Figure 1**. Snail Kite population abundance in Payne’s Prairie, north-central Florida. **(a)** Map of spatial sampling effort, each hexagonal cell corresponds to an area of ~100 km2, and the color identify sites with higher sampling in eBird (log10 scale). A single cell in northcentral Florida (~Payne’s Prairie Preserve State Park) concentrated 7,141 checklists (intense yellow color). **(b)** Records of Snail Kite in Payne’s Prairie, showing some popular visited sites within this sampling unit. **(c)** Time-series abundance observed (blue circles, ) and estimated (red squares and solid line, ), with their bootstrapped 95% confidence intervals (gray lines). Estimations with restricted maximum likelihood under Ornstein-Uhlenbeck State-Space model. This model estimates parameters of mean stationary log-abundance ((confidence intervals )) considering the rate of approaching to stationary distribution under density dependence (), variability of the process noise (), and variability of observation noise ().

A graph of a number of numbers

Description automatically generated with medium confidence

**Table 1**. Restricted maximum likelihood estimates (and 95% confidence intervals) for the parameters in the Ornstein-Uhlenbeck State Space model for two subspecies of Apolinar’s Wren in five spatiotemporal sampling unites within the Colombian Eastern Andes. The páramo subspecies (*C. a. hernandezi*) is represented in the Chisacá Lakes system sampling unit, the other four sampling units represent the nominal subspecies (*C. a. apolinari*).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Population (*n* sampling events) |  |  |  |  |
| Chisacá Lakes system (407) | 1.4332  (1.385, 1.472) | 0.0844  (0.468, 9.654) | 0.0163  (3.08 \* 10-6, 0.491) | 0.3692  (0.168, 0.505) |
| Bogotá wetlands (171) | 0.2966  (0.237, 0.352) | 1.1036  (0.401, 19.054) | 0.4613  (0.120, 7.476) | 1.92 \* 10-4  (1.47 \* 10-9, 0.091) |
| Fuquene Lake system (123) | 0.8497  (0.746, 0.936) | 0.8626  (0.793, 15.632) | 0.6696  (0.582, 6.419) | 0.1677  (2.07 \* 10-5, 0.555) |
| Tota Lake system – north (101) | 1.1863  (1.152, 1.340) | 0.9497  (1.094, 18.633) | 1.0037  (1.225, 23.050) | 6.30 \* 10-6  (4.59 \* 10-9, 0.241) |
| Tota Lake system – south (100) | 0.8892  (0.757, 0.936) | 0.3698  (0.159, 12.749) | 0.2352  (0.060, 2.436) | 0.0897  (1.14 \* 10-6, 0.464) |

**DISCUSSION**

It works. It differs from previous. We have opportunities.

For instance, state-space models could be extended to intrinsic conditional autoregressive models with parameters of variation of estimates for neighboring populations (Besag et al. 1991, Bled et el. 2013).

Test hypotheses: different populations have different dynamic, thus a particular SS case (EGSS, OUSS.NoSt, OUSS.St). Further hypotheses could include movement in Snail Kite?

Different community science data can be integrated with rigorous traditional surveys to estimate abundance in an effective way (Zhao et al., 2024). The problem arises when there are not traditional surveys to integrate community science data. Nonetheless,

Spatiotemporal subsampling in hexagonal grids to eBird data provide a measure of abundance density from community science data. We extracted the maximum value observed per day within a sampling unit, which provide a way to include more data in the time-series construction.

For example, population trends from eBird have used discrete-time stochastic exponential growth rate models (Fink, Johnston, et al., 2023). We used a similar model, EGSS, as our null hypothesis for density dependence. Including the density dependence estimation in the population growth model for stationary and nonstationary trends (OUSS) we improve decisions?

Our practical approach follows the probability distribution suggested in the original formulation of GSS (Dennis et al., 2006). Nonetheless, the model could be easily modified, like changing the observation process from a log-normal distribution to a Poisson distribution with scaling parameters linking detection probabilities or other covariates (Zhao et al., 2024). In addition, include immigration and emigration to open populations monitored by community science data?

The Kunming-Montreal Global Biodiversity Framework Goal A states that “*by 2050, the extinction rate and risk of all species are reduced tenfold and the abundance of native wild species is increased to healthy and resilient levels*”. It might be difficult to monitoring the reduction of risk of all species, or the increase of native wild species. Community science data provide an integrative approach for our society. Here we show that although eBird abundances underestimate the number of individuals, it is possible to track the population dynamics and estimate risk of extinction through time.

**ACKNOLEDGEMENTS**

We are in dept with the many volunteers that report under good practices data to eBird, they are making the differences with the power of people. O.A.C. received departmental funding at the University of Florida, such as the College of Agricultural and Life Science Dean’s Award, the Doris and Earl Lowe and Verna Lowe CALS Scholarship, as well as the SNRE Robin E. Nadeau Graduate Research Award.

**REFERENCES**

Amano, T., Lamming, J. D. L., & Sutherland, W. J. (2016). Spatial gaps in global biodiversity information and the role of citizen science. *BioScience*, *66*(5), 393–400. https://doi.org/10.1093/biosci/biw022

Auger‐Méthé, M., Newman, K., Cole, D., Empacher, F., Gryba, R., King, A. A., Leos‐Barajas, V., Mills Flemming, J., Nielsen, A., Petris, G., & Thomas, L. (2021). A guide to state–space modeling of ecological time series. *Ecological Monographs*, *91*(4), e01470. https://doi.org/10.1002/ecm.1470

Auguie, B. (2017). *gridExtra: Miscellaneous Functions for “Grid” Graphics*. R Package version 2.3. https://CRAN.R-project.org/package=gridExtra

Backstrom, L. J., Callaghan, C. T., Worthington, H., Fuller, R. A., & Johnston, A. (2024). Estimating sampling biases in citizen science datasets. *Ibis*, ibi.13343. https://doi.org/10.1111/ibi.13343

Barnes, S. K. (2023). *dggridR: Discrete Global Grids*. https://CRAN.R-project.org/package=dggridR

Bart, J., Collins, B., & Morrison, R. I. G. (2003). Estimating population trends with a linear model. *The Condor*, *105*, 367–372.

Becker, R. A., Wilks, A. R., Brownrigg, R., Minka, T. P., & Dechmyn, A. (2022). *maps: Draw Geographical Maps*. R Package version 3.4.1. https://CRAN.R-project.org/package=maps

Blake, J. G., & Loiselle, B. A. (2024). Sharp declines in observation and capture rates of Amazon birds in absence of human disturbance. *Global Ecology and Conservation*, *51*, e02902. https://doi.org/10.1016/j.gecco.2024.e02902

Callaghan, C. T., Santini, L., Spake, R., & Bowler, D. E. (2024). Population abundance estimates in conservation and biodiversity research. *Trends in Ecology & Evolution*, S0169534724000375. https://doi.org/10.1016/j.tree.2024.01.012

Campbell, C. J., Barve, V., Belitz, M. W., Doby, J. R., White, E., Seltzer, C., Di Cecco, G., Hurlbert, A. H., & Guralnick, R. (2023). Identifying the identifiers: How iNaturalist facilitates collaborative, research-relevant data generation and why it matters for biodiversity science. *BioScience*, *73*(7), 533–541. https://doi.org/10.1093/biosci/biad051

Cattau, C. E., Fletcher Jr, R. J., Kimball, R. T., Miller, C. W., & Kitchens, W. M. (2017). Rapid morphological change of a top predator with the invasion of a novel prey. *Nature Ecology & Evolution*, *2*(1), 108–115. https://doi.org/10.1038/s41559-017-0378-1

Cattau, C. E., Fletcher, R. J., Reichert, B. E., & Kitchens, W. M. (2016). Counteracting effects of a non‐native prey on the demography of a native predator culminate in positive population growth. *Ecological Applications*, *26*(7), 1952–1968. https://doi.org/10.1890/15-1020.1

Dénes, F. V., Silveira, L. F., & Beissinger, S. R. (2015). Estimating abundance of unmarked animal populations: Accounting for imperfect detection and other sources of zero inflation. *Methods in Ecology and Evolution*, *6*(5), 543–556. https://doi.org/10.1111/2041-210X.12333

Dennis, B., Munholland, P. L., & Scott, J. M. (1991). Estimation of growth and extinction parameters for endangered species. *Ecological Monographs*, *61*(2), 115–143. https://doi.org/10.2307/1943004

Dennis, B., & Ponciano, J. M. (2014). Density‐dependent state‐space model for population‐abundance data with unequal time intervals. *Ecology*, *95*(8), 2069–2076. https://doi.org/10.1890/13-1486.1

Dennis, B., Ponciano, J. M., Lele, S. R., Taper, M. L., & Staples, D. F. (2006). Estimating density dependence, process noise, and observation error. *Ecological Monographs*, *76*(3), 323–341.

Dennis, B., Ponciano, J. M., & Taper, M. L. (2010). Replicated sampling increases efficiency in monitoring biological populations. *Ecology*, *91*(2), 610–620. https://doi.org/10.1890/08-1095.1

Di Cecco, G. J., Barve, V., Belitz, M. W., Stucky, B. J., Guralnick, R. P., & Hurlbert, A. H. (2021). Observing the observers: How participants contribute data to iNaturalist and implications for biodiversity science. *BioScience*, *71*(11), 1179–1188. https://doi.org/10.1093/biosci/biab093

Evans, K. O., Davis, J. B., & Wang, G. (2023). Assessing long‐term dynamics of non‐breeding Brown Pelican ( *Pelecanus occidentalis* ) populations using Christmas Bird Count data. *Ibis*, *165*(1), 55–68. https://doi.org/10.1111/ibi.13105

Fink, D., Auer, A., Johnston, A., Strimas-Mackey, M., Ligocki, S., Robinson, O., Hochachka, W., Jaromczyk, L., Crowley, C., Dunham, K., Stillman, A., Davies, I., Rodewald, A. D., Ruiz-Gutierrez, V., & Wood, C. (2023). *eBird Status and Trends, Data Version: 2022; Released: 2023. Cornell Lab of Ornithology, Ithaca, New York.* https://doi.org/10.2173/ebirdst.2022

Fink, D., Johnston, A., Strimas‐Mackey, M., Auer, T., Hochachka, W. M., Ligocki, S., Oldham Jaromczyk, L., Robinson, O., Wood, C., Kelling, S., & Rodewald, A. D. (2023). A double machine learning trend model for citizen science data. *Methods in Ecology and Evolution*, *14*(9), 2435–2448. https://doi.org/10.1111/2041-210X.14186

Gotelli, N. J. (2008). *A primer of ecology* (Fourth Edition.). Sinauer Associates, Inc.

Humbert, J., Scott Mills, L., Horne, J. S., & Dennis, B. (2009). A better way to estimate population trends. *Oikos*, *118*(12), 1940–1946. https://doi.org/10.1111/j.1600-0706.2009.17839.x

Johnston, A., Fink, D., Hochachka, W. M., & Kelling, S. (2018). Estimates of observer expertise improve species distributions from citizen science data. *Methods in Ecology and Evolution*, *9*(1), 88–97. https://doi.org/10.1111/2041-210X.12838

Johnston, A., Hochachka, W. M., Strimas‐Mackey, M. E., Ruiz Gutierrez, V., Robinson, O. J., Miller, E. T., Auer, T., Kelling, S. T., & Fink, D. (2021). Analytical guidelines to increase the value of community science data: An example using eBird data to estimate species distributions. *Diversity and Distributions*, *27*(7), 1265–1277. https://doi.org/10.1111/ddi.13271

Kelling, S. T., Johnston, A., Bonn, A., Fink, D., Ruiz-Gutierrez, V., Bonney, R., Fernandez, M., Hochachka, W. M., Julliard, R., Kraemer, R., & Guralnick, R. (2019). Using semistructured surveys to improve citizen science data for monitoring biodiversity. *BioScience*, *69*(3), 170–179. https://doi.org/10.1093/biosci/biz010

Lees, A. C., Haskell, L., Allinson, T., Bezeng, S. B., Burfield, I. J., Renjifo, L. M., Rosenberg, K. V., Viswanathan, A., & Butchart, S. H. M. (2022). State of the World’s birds. *Annual Review of Environment and Resources*, *47*(1), 231–260. https://doi.org/10.1146/annurev-environ-112420-014642

Machado-Stredel, F., Atauchi, P. J., Nuñez-Penichet, C., Cobos, M. E., Osorio-Olvera, L., Khalighifar, A., Peterson, A. T., & Fletcher Jr., R. J. (2024). The roles of abiotic and biotic factors in driving range shifts: An invasive *Pomacea* snail facilitates *Rostrhamus sociabilis* (Snail Kite) northward range expansion. *Ornithology*, *141*, ukae019. https://doi.org/10.1093/ornithology/ukae022

Neate-Clegg, M. H. C., Horns, J. J., Adler, F. R., Kemahlı Aytekin, M. Ç., & Şekercioğlu, Ç. H. (2020). Monitoring the world’s bird populations with community science data. *Biological Conservation*, *248*, 108653. https://doi.org/10.1016/j.biocon.2020.108653

Neate‐Clegg, M. H. C., Stanley, T. R., Şekercioğlu, Ç. H., & Newmark, W. D. (2021). Temperature‐associated decreases in demographic rates of Afrotropical bird species over 30 years. *Global Change Biology*, *27*(10), 2254–2268. https://doi.org/10.1111/gcb.15567

Poli, C. L., Dudek, S. J., Jeffrey, B. M., Robertson, E. P., & Fletcher, R. J. (2020). Recent breeding range expansion of the endangered Snail Kite (*Rostrhamus sociabilis*) in Florida. *The Wilson Journal of Ornithology*, *132*(1), 183. https://doi.org/10.1676/1559-4491-132.1.183

Poli, C. L., Robertson, E. P., Martin, J., Powell, A. N., & Fletcher, R. J. (2022). An invasive prey provides long-lasting silver spoon effects for an endangered predator. *Proceedings of the Royal Society B: Biological Sciences*, *289*(1977), 20220820. https://doi.org/10.1098/rspb.2022.0820

Ponciano, J. M., Taper, M. L., & Dennis, B. (2018). Ecological change points: The strength of density dependence and the loss of history. *Theoretical Population Biology*, *121*, 45–59. https://doi.org/10.1016/j.tpb.2018.04.002

Ponciano, J. M., Taper, M. L., Dennis, B., & Lele, S. R. (2009). Hierarchical models in ecology: Confidence intervals, hypothesis testing, and model selection using data cloning. *Ecology*, *90*(2), 356–362. https://doi.org/10.1890/08-0967.1

R Core Team. (2024). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org

Reichert, B. E., Cattau, C. E., Fletcher, R. J., Sykes Jr., P. W., Rodgers Jr., J. A., & Bennetts, R. E. (2020). Snail Kite (*Rostrhamus sociabilis*), version 1.0. In *Birds of the World (A. F. Poole, Editor)*. Cornell Lab of Ornithology. https://doi.org/10.2173/bow.snakit.01

Robertson, E. P., Fletcher, R. J., & Austin, J. D. (2018). Microsatellite polymorphism in the endangered snail kite reveals a panmictic, low diversity population. *Conservation Genetics*, *19*(2), 337–348. https://doi.org/10.1007/s10592-017-1003-1

Robinson, W. D., Errichetti, D., Pollock, H. S., Martinez, A., Stouffer, P. C., Shen, F.-Y., & Blake, J. G. (2021). Big bird plots: Benchmarking Neotropical bird communities to address questions in ecology and conservation in an era of rapid change. *Frontiers in Ecology and Evolution*, *9*, 697511. https://doi.org/10.3389/fevo.2021.697511

Rosenberg, K. V., Dokter, A. M., Blancher, P. J., Sauer, J. R., Smith, A. C., Smith, P. A., Stanton, J. C., Panjabi, A., Helft, L., Parr, M., & Marra, P. P. (2019). Decline of the North American avifauna. *Science*, *366*(6461), 120–124. https://doi.org/10.1126/science.aaw1313

Staples, D. F., Taper, M. L., & Shepard, B. B. (2005). Risk‐based viable population monitoring. *Conservation Biology*, *19*(6), 1908–1916. https://doi.org/10.1111/j.1523-1739.2005.00283.x

Strimas-Mackey, M., Hochachka, W. M., Ruiz Gutierrez, V., Robinson, O. J., Miller, E. T., Auer, T., Kelling, S., Fink, D., & Johnson, A. (2023). *Best practices for using eBird data*. https://doi.org/10.5281/zenodo.3620739

Strimas-Mackey, M., Miller, E., & Hochachka, W. (2018). *auk: eBird Data Extraction and Processing in R*. https://cornelllabofornithology.github.io/auk/

Sullivan, B. L., Aycrigg, J. L., Barry, J. H., Bonney, R. E., Bruns, N., Cooper, C. B., Damoulas, T., Dhondt, A. A., Dietterich, T., Farnsworth, A., Fink, D., Fitzpatrick, J. W., Fredericks, T., Gerbracht, J., Gomes, C., Hochachka, W. M., Iliff, M. J., Lagoze, C., La Sorte, F. A., … Kelling, S. (2014). The eBird enterprise: An integrated approach to development and application of citizen science. *Biological Conservation*, *169*, 31–40. https://doi.org/10.1016/j.biocon.2013.11.003

Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S* (Fourth Edition). Springer International Publishing.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag. https://CRAN.R-project.org/package=terra

Wickham, H. ., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686.

Zhao, Q., Latif, Q. S., Nuse, B. L., Pavlacky, D. C., Kilner, C. L., Ryder, T. B., & Latimer, C. E. (2024). Integrating counts from rigorous surveys and participatory science to better understand spatiotemporal variation in population processes. *Methods in Ecology and Evolution*, 2041-210X.14368. https://doi.org/10.1111/2041-210X.14368

Ziolkowski, D. J., Lutmerding, M., English, W. A., Aponte, V. I., & Hudson, M.-A. R. (2023). *North American Breeding Bird Survey Dataset 1966—2022: U.S. Geological Survey data release*. U.S. Geological Survey. https://doi.org/10.5066/P9GS9K64

**Supporting Information for**

**Risk-based viable population monitoring from community science data**

Orlando Acevedo-Charry1,2,3, \*, José M. Ponciano4, Caroline L. Poli5, Brian M. Jeffrey2, Robert J. Fletcher Jr. 2, María Angela Echeverry-Galvis6, Bette A. Loiselle2,7, Scott K. Robinson3, Miguel A. Acevedo2

1School of Natural Resources and Environmental, University of Florida, Gainesville, FL 32611, USA

2Department of Wildlife Ecology and Conservation, University of Florida, Gainesville, FL 32611, USA

3Florida Museum of Natural History, University of Florida, Gainesville, FL 32611, USA

4Department of Biology, University of Florida, Gainesville, FL 32611, USA

5Florida Fish and Wildlife Conservation Commission (FWC), FWC Fish and Wildlife Research Institute, Saint Petersburg, FL 33701, USA

6Departamento de Ecología y Territorio, Pontificia Universidad Javeriana, Bogotá D.C., Colombia

7Center for Latin American Studies and Tropical Conservation Development Program, University of Florida, Gainesville, FL 32611, USA

**\*Corresponding author.** [**acevedocharry@gmail.com**](mailto:acevedocharry@gmail.com)

**Section S1: Temporal bias of eBird records for Snail kites in Florida**

**A graph of a number of kites

Description automatically generated**

**Figure SI-1.** Histogram of the number of eBird checklists per year for Florida Snail kites. Red lines indicate our temporal sampling bias threshold (more than 1000 checklists per year), concentrating our sampling between January 2018 and May 2024.

**Section S2: Step by step of a Population Viability Monitoring (or risk-based viable population monitoring) for Snail kites abundance records in eBird**

1. Fit an EGSS model for the first two years (2018-2019).

A graph of a number of people

Description automatically generated with medium confidence

**Figure SI-2**. (*LEFT*) eBird observation (blue circles) and Exponential Growth State-Space model prediction (red triangles) for Snail kites in Payne’s Prairie, Alachua County, north-central Florida, during the first two years (2018-2019; weeks 1 to 104 in the time series). Black dots represent observations following standardized methods for 2018 (Poli et al., 2020). (*RIGTH*) With this first model, we can identify an underestimation of the abundance by a factor of 4.2 during the first year of standardized monitoring.

1. With the model parameters, project trajectories for the following two years (104 weeks) and estimate probability of going below a threshold (e.g., 1/2 mean of the Population abundance of all time-series).

A graph of a graph with numbers and a line

Description automatically generated with medium confidence

**Figure SI-3**. First projection of the population of Snail Kite in Payne’s Prairie wetland for two years after the first two years of establishment (2018-2019). The figure depicts 1000 trajectories that end after two years of simulation below (red lines) or above (gray lines) of 6 individuals (our assumed for pseudo extinction).

1. Add the next week of observation, estimate again the model parameters, and project again two years ahead (104 weeks), estimating probability of going below a threshold. From this point on, we can fit OUSS, but if the trend parameter , we selected EGSS because it is an indication of density independence ().

**A graph of a number of bird species

Description automatically generated with medium confidence**

**Figure SI-4**. New fitted model, including observation. Although the OUSS model was fitted, the trend parameter indicated density independence, suggesting the use of the EGSS model. With the new parameters estimated for the EGSS model, the population trajectories are simulated 2 years ahead (104 weeks), estimating an update probability to end below a threshold (**P(ext risk)**).

1. Iterate the process for the following periods. We can subsample some moments to illustrate the iteration, for instance in weeks 104 (December 2019), 129 (June 2020), 154 (December 2020), 179 (June 2021), 206 (December 2021), 231 (June 2022), 256 (November 2022), 281 (May 2023), and 306 (November 2023). The model selected in each period projection is also stored (see week 206-December 2021).

**A group of graphs showing different types of plants

Description automatically generated with medium confidence**

**Figure SI-5**. Nine weeks across the time-series of Snail kites’ abundance in Payne’s Prairie, illustrating the population viability monitoring methodology. The models (OUSS or EGSS) are fitted in each period, with . Use of the EGSS model depends on assessing if the trend parameter in the OUSS indicated density independence (). With the new parameters estimated in each period, the population trajectories are simulated 2 years ahead (104 weeks), and probability to end below a threshold is estimated (**P(ext risk)**).