**Appendix S1**. Additional details on bird survey data sources, processing, and model output.

**METHODS**

The black rail survey data used in this study came from two sources: a structured nocturnal marsh bird survey (Allen et al. 2017, NJDEP ENSP 2021), and the eBird citizen science database (2008-2019; Sullivan et al. 2014). The structured survey data were provided by the New Jersey Department of Environmental Protection, Endangered and Nongame Species Program (C Davis, personal communication). The data were collected by state biologists and trained volunteers in 2015, 2016, 2018, and 2019 in coordination with local non-profit conservation groups (New Jersey Audubon and Conserve Wildlife Foundation of New Jersey). More details on this survey can be found in Allen et al. (2017) but are summarised here. Survey points were assigned to surveyors from a larger pool of mapped points which covered nearly all high-elevation tidal marsh habitats in the state accessible along roadsides or by boat. The choice of points to survey was driven partly by logistics, but also by the goal of covering as many points as possible spread evenly across the state’s coastline. Surveyors were instructed to visit points between 22:00 and 03:00, three times per season, from 29 April to 15 July. Surveys lasted 10 minutes and included 4 min of broadcasting black rail calls with a portable mp3 player and speaker. Surveys at some points in 2016 (n = 140) were longer, lasting 15 minutes, with 8 min of broadcasted black rail calls. The final data set yielded 373 surveyed points, some of which were surveyed in multiple years (see Results). The number of surveys per point within each year ranged from 1-6 (mean = 2.9).

We found high uncertainty in parameter estimates during exploratory analysis (see Figs. S2 and S3), likely due to low prevalence of rails encountered during surveys and the statistical phenomenon of class imbalance (Robinson et al. 2018). Therefore, we opted to incorporate additional data on black rail presence or absence from the eBird citizen science database (Sullivan et al. 2014). These data take the form of checklists, with counts of each species encountered, along with location and timing information, and effort variables including duration of count and distance traveled. These data, after filtering to improve compatibility with structured surveys, have been shown to improve the accuracy of species distribution models (Steen et al. 2019, Robinson et al. 2020). We obtained 194,260 ‘complete checklists’ with latitude and longitude information from 36,701 unique locations submitted within our study area. Complete checklists record every bird species encountered by an observer and therefore can be used as a source of both detection and non-detection data. We used the following criteria to stringently filter the pool of available checklists to a subset compatible with our existing dataset. We retained only stationary checklists, submitted in 2008 or later, between 1 April and 15 July, ≥ 5 hours from solar noon (i.e., ~17:00-07:00 EST), and lasting ≤ 5 hrs. If > 10 checklists were available from a given location and year, we randomly chose 10 to be included as the repeated visits for use in occupancy modelling (following Strimas-Mackey et al. 2020). This process resulted in a final total of 977 eBird survey locations added to our data set. Exploratory analysis revealed the eBird data covered a similar range of environmental covariates as did the structured surveys, except with a greater range of coverage by eBird surveys in areas with less high marsh (Fig. S1).

**RESULTS**

The full black rail survey data set included 373 state-coordinated survey locations and 977 eBird stationary checklist locations. Thirty percent of survey points and 23% of eBird locations were surveyed in more than one year. This resulted in 2075 total location and year combinations (500 from survey data and 1575 from eBird). Black rails were detected at 30 survey points (8%) and 10 eBird locations (1%). Only one standardised survey point and one eBird location had black rail detections in multiple years.

 Results from the occupancy model, including a comparison between a model with and without eBird data, can be found in Figs. S2-S4.

**LITERATURE CITED**

Allen MC, Tsipoura N, La Puma D (2017) Assessing the status of Eastern Black Rail (*Laterallus jamaicensis jamaicensis*) in New Jersey. *The Peregrine Observer* 2017: 180-185.

NJDEP ENSP [New Jersey Department of Environmental Protection Endangered and Nongame Species Program]. 2021. Black Rail survey 2015-2020. Trenton: New Jersey Department of Environmental Protection, Division of Fish and Wildlife - Endangered & Nongame Species Program.

Robinson OJ, Ruiz‐Gutierrez V, Fink D (2018) Correcting for bias in distribution modelling for rare species using citizen science data. *Diversity and Distributions* 24: 460-472.

Robinson OJ, Ruiz‐Gutierrez V, Reynolds MD, Golet GH, Strimas‐Mackey M, Fink D (2020) Integrating citizen science data with expert surveys increases accuracy and spatial extent of species distribution models. *Diversity and Distributions* 26: 976-986.

Steen VA, Elphick CS, Tingley MW (2019) An evaluation of stringent filtering to improve species distribution models from citizen science data. *Diversity and Distributions* 25:1857-1869.

Strimas-Mackey M, Hochachka WM, Ruiz-Gutierrez V, Robinson OJ, Miller ET, Auer T, Kelling S, et al. (2020) Best Practices for Using eBird Data. Version 1.0. Ithaca, USA: Cornell Lab of Ornithology. URL <https://doi.org/10.5281/zenodo.3620739>.

Sullivan BL, Aycrigg JL, Barry JH, Bonney RE, Bruns N, Cooper CB, Damoulas T, et al. (2014) The eBird enterprise: an integrated approach to development and application of citizen science. *Biological Conservation* 169: 31-40.

**Table S1.** Estimates of ‘coarse scale’ (*Ac*) and ‘fine scale’ (*Af*) area of occurrence for black rails (*Laterallus jamaicensis jamaicensis*) in New Jersey, USA saltmarshes (see Methods in main text).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Ownership** | ***Ac*** | **Ac 95% CI** | ***d†*** | ***r‡*** | ***Af*** |
| All | 6915 | [3746, 12144] | 200 | 3 | 1650.8 |
|  |  |  | 300 | 3 | 733.7 |
|  |  |  | 400 | 3 | 412.7 |
|  |  |  | 500 | 3 | 264.1 |
|  |  |  | 200 | 4 | 2201.1 |
|  |  |  | 300 | 4 | 978.3 |
|  |  |  | 400 | 4 | 550.3 |
|  |  |  | 500 | 4 | 352.2 |
|  |  |  | mean | 892.9 |
| Federal | 1652 | [874, 5427] | 200 | 3 | 394.4 |
|  |  |  | 300 | 3 | 175.3 |
|  |  |  | 400 | 3 | 98.6 |
|  |  |  | 500 | 3 | 63.1 |
|  |  |  | 200 | 4 | 525.8 |
|  |  |  | 300 | 4 | 233.7 |
|  |  |  | 400 | 4 | 131.5 |
|  |  |  | 500 | 4 | 84.1 |
|  |  |  | mean | 213.3 |
| State | 3089 | [1664, 5427] | 200 | 3 | 737.4 |
|  |  |  | 300 | 3 | 327.8 |
|  |  |  | 400 | 3 | 184.4 |
|  |  |  | 500 | 3 | 118.0 |
|  |  |  | 200 | 4 | 983.3 |
|  |  |  | 300 | 4 | 437.0 |
|  |  |  | 400 | 4 | 245.8 |
|  |  |  | 500 | 4 | 157.3 |
|  |  |  | mean | 398.9 |
| County | 79 | [42, 144] | 200 | 3 | 18.9 |
|  |  |  | 300 | 3 | 8.4 |
|  |  |  | 400 | 3 | 4.7 |
|  |  |  | 500 | 3 | 3.0 |
|  |  |  | 200 | 4 | 25.1 |
|  |  |  | 300 | 4 | 11.2 |
|  |  |  | 400 | 4 | 6.3 |
|  |  |  | 500 | 4 | 4.0 |
|  |  |  | mean | 10.2 |

*†* Estimated range of typical maximum distances (m) from which black rails can be detected. The lower bound is half the recommended 400 m between-point distance in standardized surveys for black rails (Conway 2011: <https://doi.org/10.1675/063.034.0307>). The upper bound is based on estimated distances reported by observers during standardized surveys used in our study.

*‡*Estimated typical home range size (3-4 ha) for black rails from Eddleman et al. (2020; <https://doi.org/10.2173/bow.blkrai.01>).

**Table S1 (continued).** Estimates of ‘coarse scale’ (*Ac*) and ‘fine scale’ (*Af*) area of occurrence for black rails (*Laterallus jamaicensis jamaicensis*) in New Jersey, USA saltmarshes (see Methods).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Ownership**  | ***Ac*** | **95% CI** | ***d*** | ***r*** | ***Af*** |
| City/Municipal | 353 | [189, 623] | 200 | 3 | 84.3 |
|  |  |  | 300 | 3 | 37.5 |
|  |  |  | 400 | 3 | 21.1 |
|  |  |  | 500 | 3 | 13.5 |
|  |  |  | 200 | 4 | 112.4 |
|  |  |  | 300 | 4 | 49.9 |
|  |  |  | 400 | 4 | 28.1 |
|  |  |  | 500 | 4 | 18.0 |
|  |  |  | mean | 45.6 |
| Non-profit | 303 | [161, 543] | 200 | 3 | 72.3 |
|  |  |  | 300 | 3 | 32.1 |
|  |  |  | 400 | 3 | 18.1 |
|  |  |  | 500 | 3 | 11.6 |
|  |  |  | 200 | 4 | 96.4 |
|  |  |  | 300 | 4 | 42.9 |
|  |  |  | 400 | 4 | 24.1 |
|  |  |  | 500 | 4 | 15.4 |
|  |  |  | mean | 39.1 |
| Private | 1421 | [768, 2519] | 200 | 3 | 339.2 |
|  |  |  | 300 | 3 | 150.8 |
|  |  |  | 400 | 3 | 84.8 |
|  |  |  | 500 | 3 | 54.3 |
|  |  |  | 200 | 4 | 452.3 |
|  |  |  | 300 | 4 | 201.0 |
|  |  |  | 400 | 4 | 113.1 |
|  |  |  | 500 | 4 | 72.4 |
|  |  |  | mean | 183.5 |

*†* Estimated range of typical maximum distances (m) from which black rails can be detected. The lower bound is half the recommended 400 m between-point distance in standardized surveys for black rails (Conway 2011: <https://doi.org/10.1675/063.034.0307>). The upper bound is based on estimated distances reported by observers during standardized surveys used in our study.

*‡*Estimated typical home range size (3-4 ha) for black rails from Eddleman et al. (2020; <https://doi.org/10.2173/bow.blkrai.01>).

**Table S2.** Comparison of model coefficients from the Bayesian occupancy models describing predicted eastern black rail (*Laterallus jamaicensis jamaicensis*) distribution within saltmarshes in New Jersey, USA; models treating year as a categorical (top; DIC = 503.6) and continuous (bottom; DIC = 512.4) variable are compared.**†**

| **Year treatment** | **Sub-model** | **Variable** | **Posterior****median** | **2.5%** | **97.5%** | **Rhat** |
| --- | --- | --- | --- | --- | --- | --- |
| **Categorical** | ***p*** | Intercept | 0.4440 | -1.4534 | 2.3986 | 1.002 |
|  |  | Ordinal date | -0.0095 | -0.0222 | 0.0288 | 1.002 |
|  |  | Duration (h) | 0.2436 | -0.3264 | 0.8809 | 1.000 |
|  | ***ψ*** | Intercept (Year 2015) | -3.4461 | -4.3333 | -2.6242 | 1.000 |
|  |  | Year 2008 | 0.2412 | -2.5537 | 2.1793 | 1.000 |
|  |  | Year 2009 | 1.5348 | -0.1079 | 3.0569 | 1.000 |
|  |  | Year 2010 | -0.2711 | -3.0076 | 1.5592 | 1.000 |
|  |  | Year 2011 | -0.3455 | -3.0260 | 1.4818 | 1.000 |
|  |  | Year 2012 | -2.6134 | -7.5367 | 0.3835 | 1.000 |
|  |  | Year 2013 | -2.8525 | -7.7845 | -0.1335 | 1.001 |
|  |  | Year 2014 | -0.8879 | -3.6010 | 0.8628 | 1.001 |
|  |  | Year 2016 | 0.5510 | -0.2132 | 1.3254 | 1.000 |
|  |  | Year 2017 | -0.5638 | -2.3388 | 0.8171 | 1.000 |
|  |  | Year 2018 | -1.3339 | -3.1813 | 0.0206 | 1.004 |
|  |  | Year 2019 | -3.6524 | -8.1157 | -1.1472 | 1.005 |
|  |  | Shrub-scrub | -2.1539 | -7.3314 | 2.5475 | 1.000 |
|  |  | High marsh | 3.1515 | 1.5808 | 4.8134 | 1.000 |
|  |  | Terrestrial border | 1.4024 | -4.2878 | 6.8702 | 1.000 |
|  |  | Low-int. developed | -1.9387 | -4.6885 | 0.5588 | 1.000 |
|  |  | Impoundments | 0.4883 | -2.8669 | 3.3286 | 1.000 |
|  |  | **Deviance** | 286.4 | 252.2 | 333.5 | 1.000 |
| **Continuous** | ***p*** | Intercept | 0.8789 | -0.8515 | 2.7837 | 1.005 |
|  |  | Date | -0.0127 | -0.0250 | -0.0014 | 1.004 |
|  |  | Duration | 0.1680 | -0.3911 | 0.7475 | 1.001 |
|  | ***ψ*** | Intercept | -3.6918 | -4.3722 | -3.0455 | 1.002 |
|  |  | Year | -0.1008 | -0.2304 | 0.0076 | 1.001 |
|  |  | Shrub-scrub | -1.7414 | -6.7884 | 3.0507 | 1.001 |
|  |  | High marsh | 3.1265 | 1.6062 | 4.6941 | 1.001 |
|  |  | Terrestrial border | 1.6580 | -3.9965 | 7.1676 | 0.999 |
|  |  | Low-int. developed | -1.6378 | -4.3318 | 0.7645 | 1.001 |
|  |  | Impoundments | 0.8553 | -2.3678 | 3.6263 | 0.999 |
|  |  | **Deviance** | 287.6 | 252.4 | 335.2 | 1.002 |

**†** Posterior medians and credible intervals were generated 12000 total draws from the posterior. Models are based on the combined eBird and survey data set. Pearson correlation between mapped grid cell predictions from the two models was 0.997.

**Figure S1.** Environmental covariates at eastern black rail (*Laterallus jamaicensis jamaicensis*) structured survey points, eBird stationary checklist locations, and all locations within emergent tidal marshes (i.e., at all 30 x 30 m grid cells). The covariates are the proportional cover of impoundments within 500 m (imp.500), low-intensity development within 100 m (ldev.100), high marsh habitat within 500 m (sharp\_hm.500), terrestrial border habitat within 500 m (sharp\_tb.500), and scrub-shrub habitat within 500 m (ss.500).



**Figure S2.** Coefficients from Bayesian occupancy models describing eastern black rail (*Laterallus jamaicensis jamaicensis*) habitat suitability within saltmarshes in New Jersey, USA. Yellow symbols represent the model based only on the structured survey data, while green symbols represent the model that also included data from stationary eBird checklists. Parameter estimate posterior medians (points) and credible intervals (thick lines – 80%; thin lines – 95%) were generated using the same MCMC settings as described in the Methods, but with 12000 total post-burn-in draws from the posterior.



**Figure S3.** Predicted area of occurrence from Bayesian occupancy models describing eastern black rail (*Laterallus jamaicensis jamaicensis*) habitat suitability within saltmarshes in New Jersey, USA. Yellow symbols represent the model based only on the structured survey data, while green symbols represent the model that also included data from stationary eBird checklists. These represent ‘coarse scale’ estimates (*Ac*), or the area from which black rails can be detected by an observer (see Methods in main text). Parameter estimate posterior medians (points) and credible intervals (thick lines – 80%; thin lines – 95%) are shown. See Fig. S2 for MCMC settings.



**Figure S4.** Predicted occupancy (red) and detection (blue) probabilities from a Bayesian occupancy model of eastern black rail (*Laterallus jamaicensis jamaicensis*) habitat suitability within New Jersey (USA) saltmarshes. The model is based on integrated data from both structured surveys and stationary eBird checklists. Shaded areas represent 80% and 95% credible intervals. Dark and light shading represent 80% and 95% credible intervals, respectively. See Fig. S2 for MCMC settings.

