Supporting Information. Daniel Fink, Tom Auer, Alison Johnston, Viviana Ruiz-Gutierrez, Wesley M. Hochachka, and Steve Kelling. 2020. Modeling avian full annual cycle distribution and population trends with citizen science data. *Ecological Applications.*

Appendix S1: Ensemble Design

The ensemble of stixels was designed as a Monte Carlo sample of 100 spatiotemporal partitions of the study extent with randomized spatial and temporal location and orientation. Because each set of stixels partitions the study extent, each unique location in space and time is a member of 100 different randomly located and oriented stixels. Averaging predictions at a given location and time across this randomized sample of stixels helps control for biases associated with the arbitrary partitioning of data into stixels.

Stixel size controls an important bias-variance tradeoff (Fink *et al* 2010; Fink *et al.* 2013). Stixel size needs to be chosen small enough to capture local predictor-response (i.e. species-environment) relationships, controlling the bias of base model estimates. Stixel size also needs to be chosen large enough to meet the minimum sample size requirements necessary for fitting the base models: this controls the variance when averaging across the ensemble. We began by specifying the temporal dimension of the partitions to be 366 days divided by 12 temporal partitions, equaling 30.5 contiguous days. We have found that an approximately 30 day window is small enough to capture a wide variety of migration patterns across a diverse set of terrestrial species using eBird data (Johnston *et al.* 2015; La Sorte *et al.* 2017). The spatial dimensions were adaptively sized to generate smaller stixels in regions with higher data density using QuadTrees (Samet 1984), a recursive partitioning algorithm. Adaptively sizing stixels as a function of spatial data density has been found to improve performance over fixed-size stixels (Fink *et al.* 2013) in the presence of data with variable spatial density (Fig S1). Note that for single spatiotemporal partition of the study exent, the same QuadTree spatial partition was used for each of the 12 months of the year. In QuadTrees, the splitting rule controls the recursion. The splitting rule was set to recursively split stixels with more than 15,000 checklists. Given high variation in data density, this splitting rule generated overly large stixels in data poor regions and extremely small stixels in areas of high data density. To prevent this, we constrained the partitioning to 1) not split stixels smaller than 5° longitude x 5° latitude regardless of the number of checklists within the stixel, and 2) forced stixels larger than 25° longitude x 25° latitude to split regardless of data density within the stixel.

We used the following procedure to randomize the location and orientation of each spatiotemporal partition of the study extent. First, the location of the QuadTree partition was randomized by jittering the latitude and longitude of the QuadTree origin. Additionally, the spatial orientation of the QuadTree partition was randomized by randomly rotating the resulting grid. Each spatiotemporal partition was temporally randomized by jittering the starting date of the temporal (monthly) grid division. Fig. S2 shows realizations of two randomly located adaptive spatial partitions used to define AdaSTEM stixels for the analysis.

The number of base models averaged across the stixel ensemble to estimate a prediction at a single location and time is called the *ensemble support*. Ensemble support is important because it controls inter-model variability when averaging across the ensemble. In this application, the maximum ensemble support is 100, the total number of randomized partitions. We selected this value based on both statistical considerations and our computational budget. When sample sizes are too small to fit base models, those stixels are dropped from the ensemble and the number of base model estimates available for ensemble averaging decreases, increasing the variance of the ensemble estimator. In general, ensemble support follows patterns of data density, filtered through a combination of the base model minimum sample size requirements and stixel geometry. We required an ensemble support of at least 75 to generate the weekly estimates of occurrence and abundance.

Literature Cited

Fink, D., Hochachka, W.M., Zuckerberg, B., Winkler, D.W., Shaby, B., Munson, M.A., Hooker, G., Riedewald, M., Sheldon, D. and Kelling, S., (2010). Spatiotemporal exploratory models for broad_scale survey data. *Ecological Applications*, 20(8), 2131-2147.

Fink, D., Damoulas, T. and Dave, J., (2013), July. Adaptive Spatio-Temporal Exploratory Models: Hemisphere-wide species distributions from massively crowdsourced eBird data. In: *Twenty-Seventh AAAI Conference on Artificial Intelligence* (AAAI-13) July 14–18, 2013 in Bellevue, Washington, USA.

Johnston, A., Fink, D., Reynolds, M.D., Hochachka, W.M., Sullivan, B.L., Bruns, N.E., Hallstein, E., Merrifield, M.S., Matsumoto, S. and Kelling, S. (2015). Abundance models improve spatial and temporal prioritization of conservation resources. *Ecological Applications*, 25(7), 1749-1756.

La Sorte, F.A., Fink, D., Blancher, P.J., Rodewald, A.D., Ruiz-Gutierrez, V., Rosenberg, K.V., Hochachka, W.M., Verburg, P.H. and Kelling, S., (2017). Global change and the distributional dynamics of migratory bird populations wintering in Central America. *Global Change Biology*, 23(12), 5284-5296.Mease *et al.* 2007

Samet, H. (1984). The quadtree and related hierarchical data structures. ACM Computing Surveys (CSUR), 16(2), 187-260.

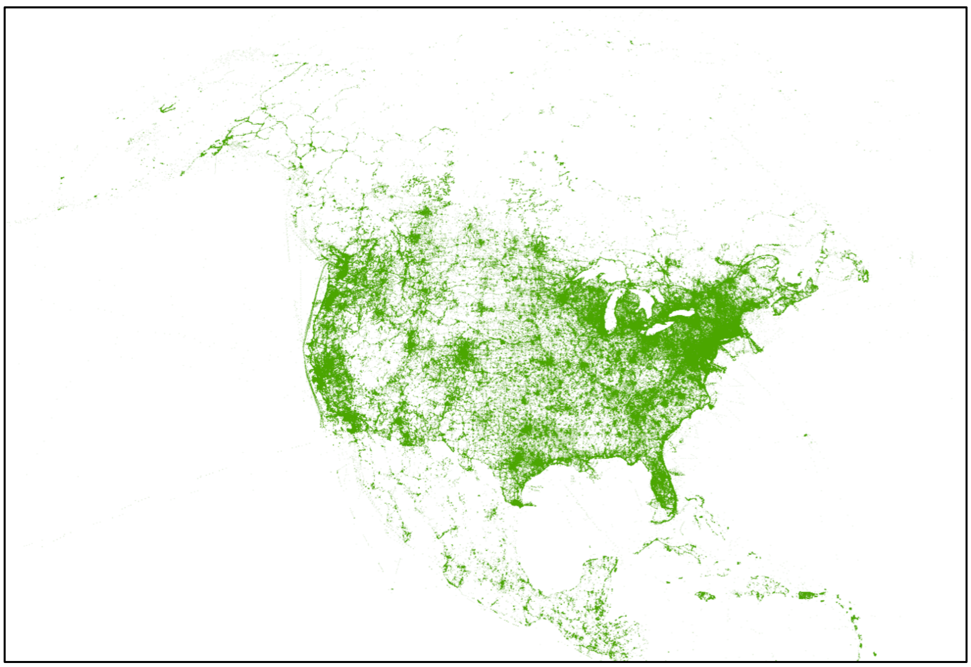

Figure S1: eBird checklist locations between January 1, 2004 to December 31, 2016. The spatial density of the 11.7 million checklists used in the analysis vary significantly across the North American continent.

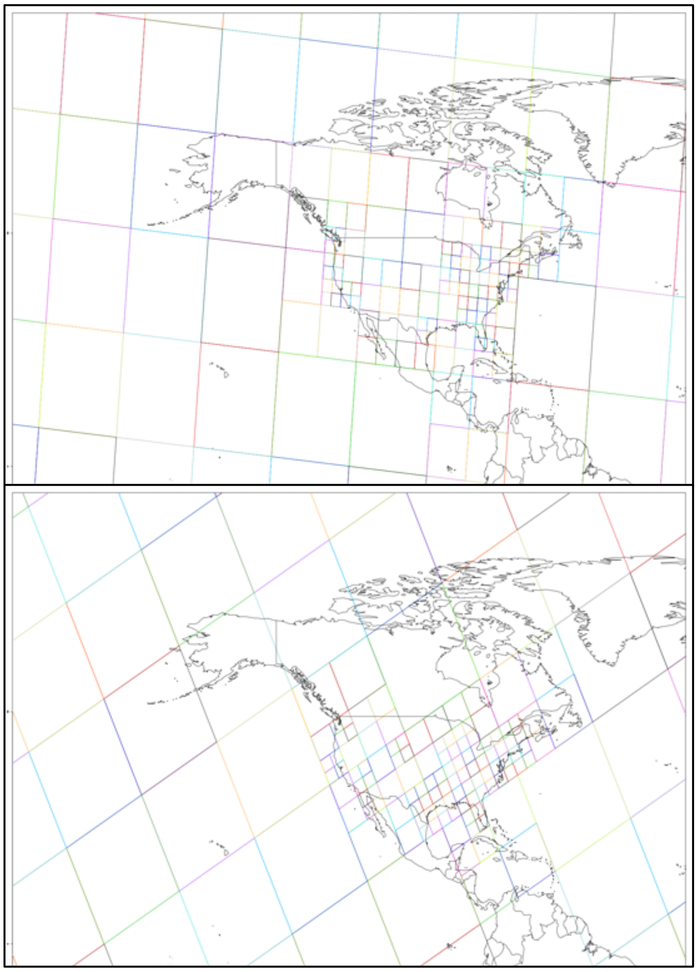

Figure S2: Two realizations of randomly located adaptive spatial partitions used to define AdaSTEM stixels. These image shows how stixel size adapts to data density and how stixel location is randomized through jittering and rotation. One hundred randomized adaptive partitions were used for the analysis. For each spatial partition, the temporal partition is fixed to 30.5 days, but has a randomized starting date.