Appendix 1 – Example of a single simulation realization, and summary of 1,000 repeated simulations

Here, we illustrate a single realization of a simulation that was used to test our statistical model. As described in the main text of the manuscript, we simulated trajectories at each of 15 colonies as a combination of two temporally autocorrelated environmental processes: 1) an "unshared" environmental covariate that varied independently among colonies, and 2) a "shared" environmental covariate that affected all colonies simultaneously, imposing correlation among colony trajectories. In Figure S1.1, we illustrate a single realization of a simulation in which colonies were affected by both environmental covariates; thus, colonies partially share a temporal trajectory. We randomly assigned each colony to be surveyed in 2-6 randomly selected years across the 50-year period, with at least one survey occurring in the first and last 10 years of the simulation. Observation error was simulated according to equation 2 of the manuscript, with parameters chosen from the posterior mean from the empirical analysis for storm-petrels, where $\log(\varepsilon_{i,t}) \sim Normal(-0.95 + 0.86 \times \log(E_{i,t}), 0.45^2)$. Counts were simulated as Poisson random variables using equation 1. The magnitude of observation error was then retained for 65% of surveys, to replicate the empirical data where estimates of error were not available for all surveys.

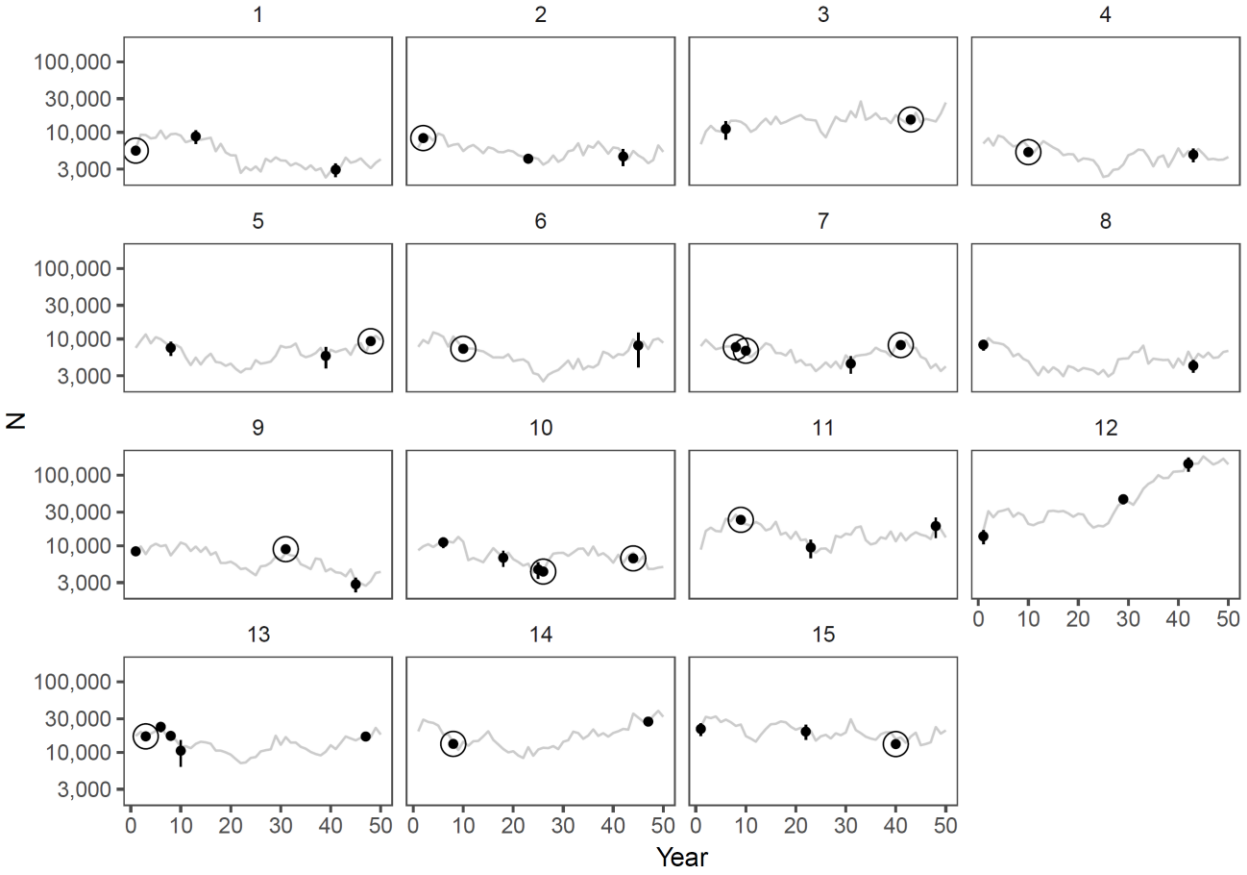

Figure S1.1. Example of simulated dynamics and observed counts at each of 15 colonies. In this example, colonies shared a time-varying environmental driver, resulting in partially shared temporal colony trajectories. Black dots indicate survey counts, and error bars indicate counts for which we assumed the

magnitude of survey error was known. Open circles around dots indicate that survey error was unknown, and therefore would be estimated as part of the model-fitting process.

We fit the statistical model described in the manuscript to these simulated data, allowing us to estimate the temporal trajectories at each colony (shown in Figure S1.2).

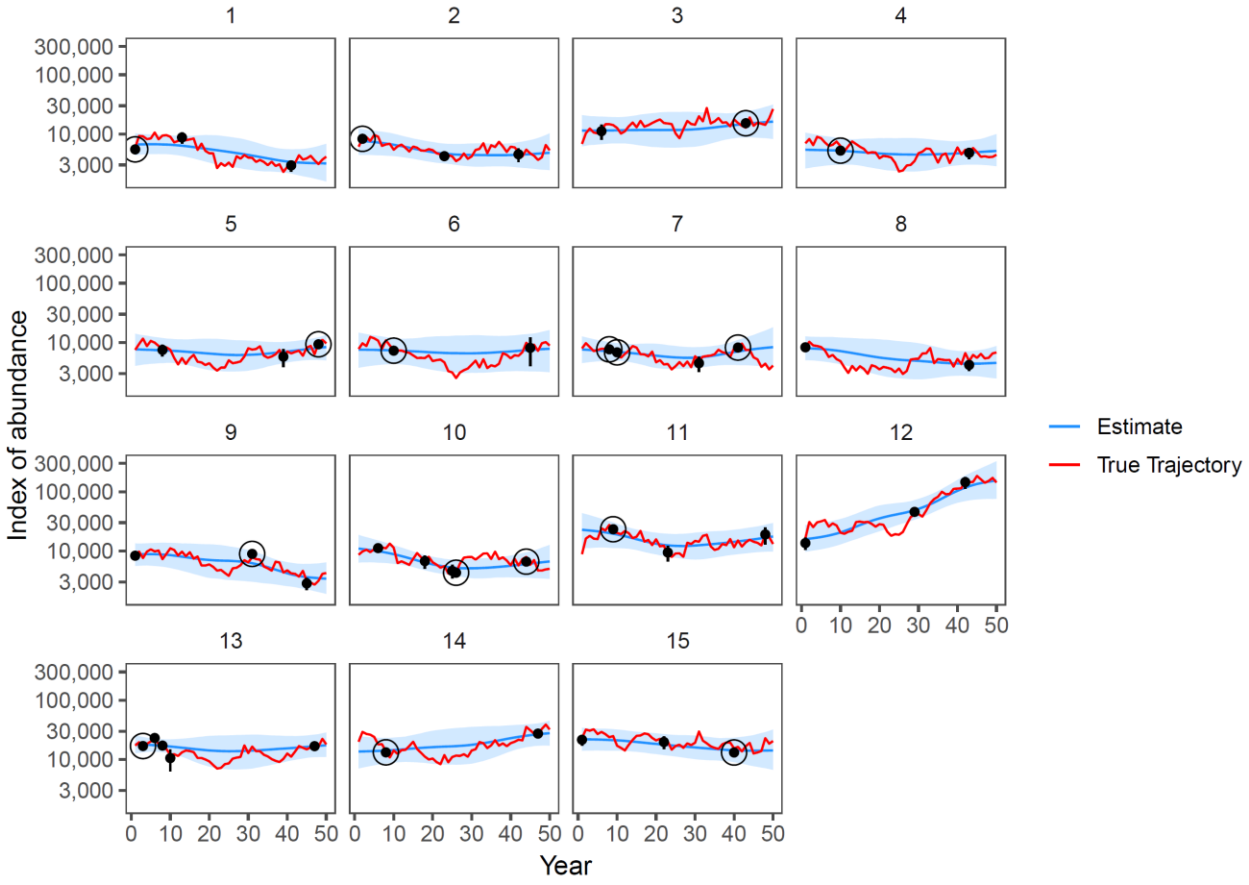

Figure S1.2. Example of estimated temporal trajectories at each of the 15 simulated colonies. Blue lines indicate posterior median estimate, and blue ribbons represent 95% equal-tailed credible intervals around the prediction. Red lines represent the true colony-level population trajectories. Dots indicate survey counts and are interpreted as in Figure S1.1.

We used estimates of colony-level annual abundance (Figure S1.2) to calculate annual estimates of the regional population total. The trend in the estimated regional population was calculated using equation 7 in the manuscript. Figure S1.3 compares the estimated regional population trajectory (blue line) to the true population trajectory (red line), as well as a smooth fit to the true population trajectory (black line).

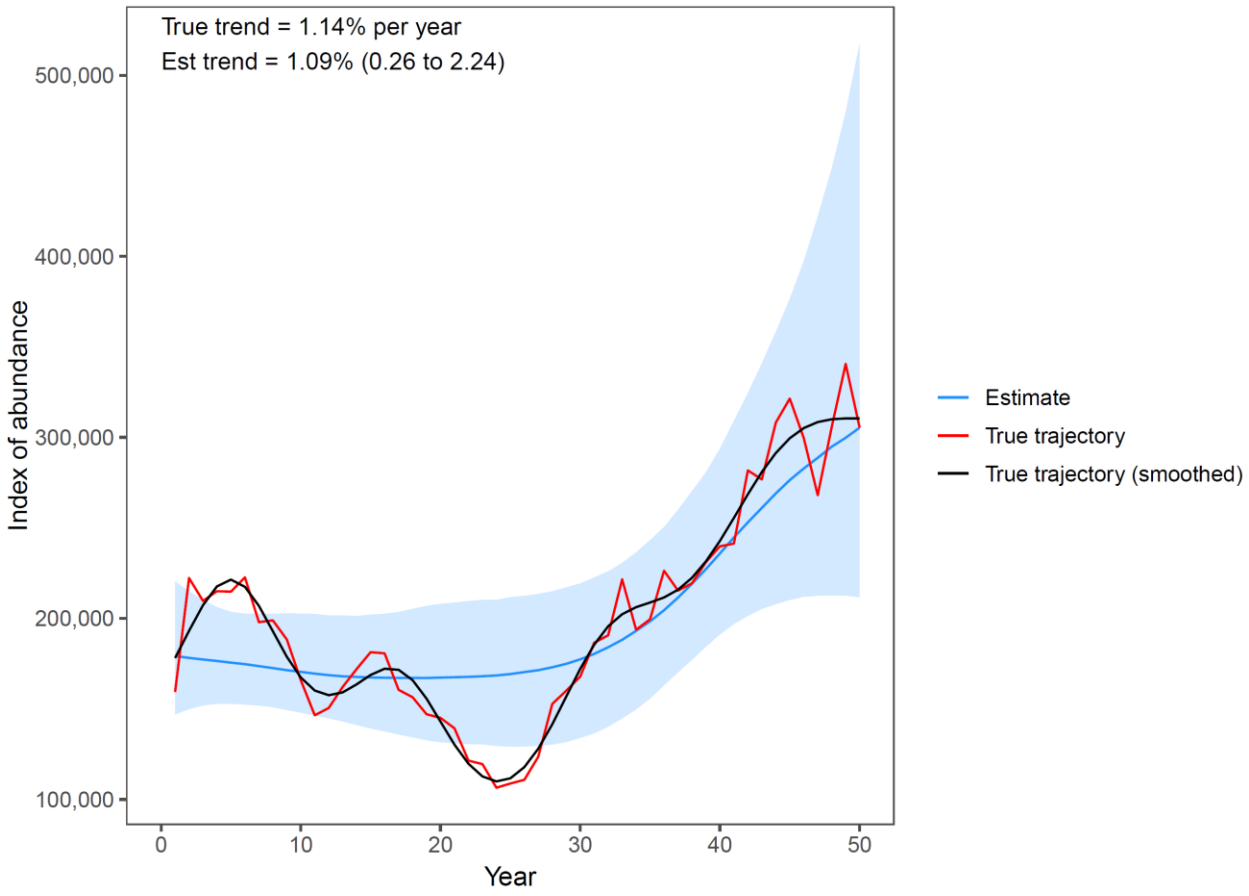

Figure S1.3. Example of regional population trajectories at each of the 15 simulated colonies. Blue lines indicate posterior median estimate, and blue ribbons represent 95% equal-tailed credible intervals around the prediction. "True trend" is calculated as the geometric mean rate of change in the smoothed true trajectory (black line), while estimated trend is calculated as the geometric mean rate of change in the estimated trajectory (blue line).

We repeated this process 1,000 times and compared true population trends to those estimated from our statistical model. We compared trend estimates calculated for the entire 50-year simulation (bottom row of Figure S1.4), as well as trend estimates for the final 10 years (top row of Figure S1.4). Results of these comparisons are illustrated in Figure S1.4 and Table S1.1.

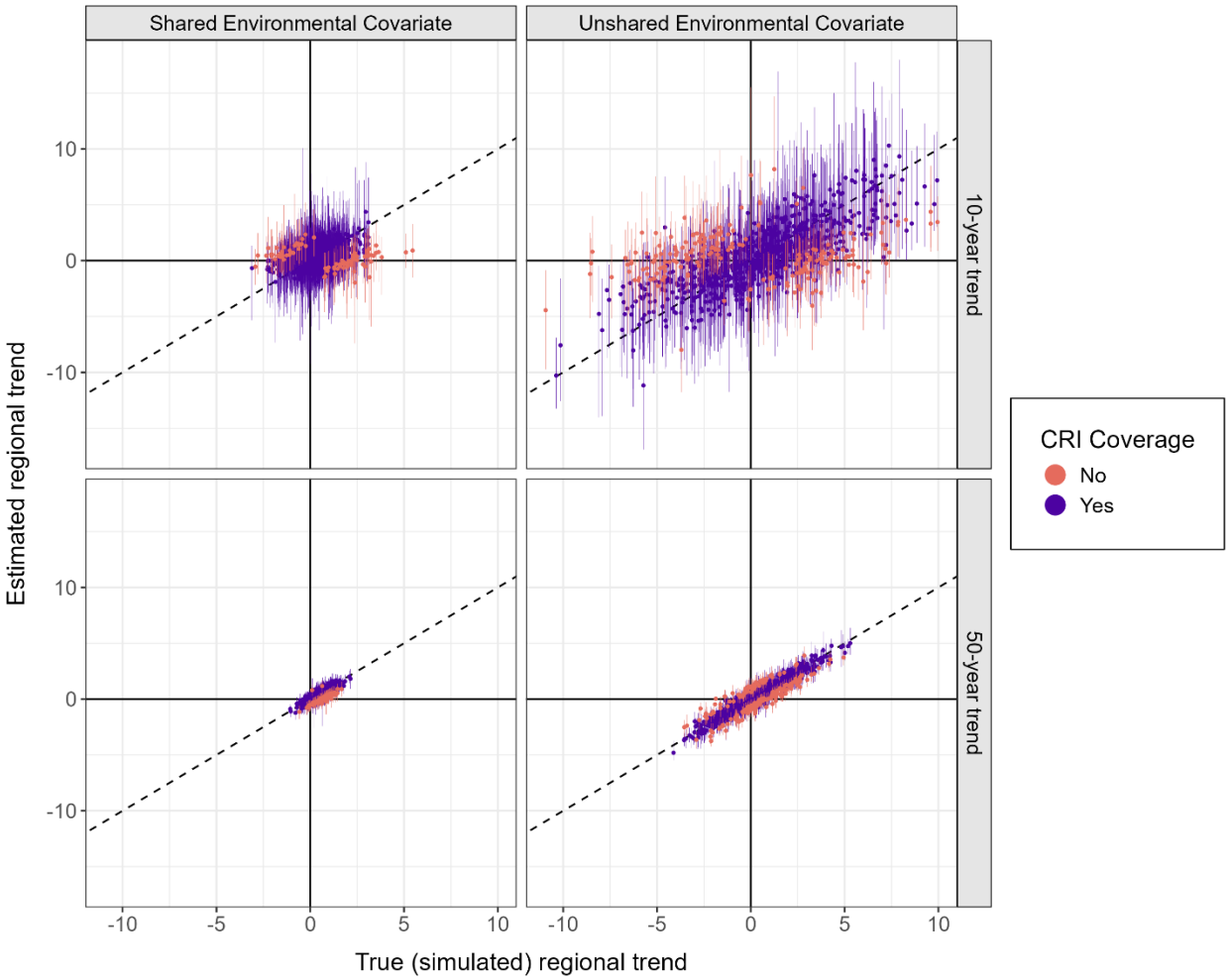

Figure S1.4. Results of 1,000 simulated hypothetical scenarios of population change derived using a Bayesian GAMM approach based on 2-6 imprecise surveys from 15 colonies over a 50-year time period. Credible interval (CRI) coverage indicates whether the 95% credible interval of the estimated regional trend (vertical error bars) captured the "true" simulated regional trend for the most recent 10-year interval (top panels) or a full 50-year interval (bottom panels), under simulations where population dynamics at each colony were affected by an autocorrelated shared environmental covariate imposing correlations among the colony trajectories (left panels) or an unshared environmental covariate (right panels).

Table S1.1. Summarized results of 1,000 simulated hypothetical scenarios of population change derived using a Bayesian GAMM approach based on 2-6 imprecise surveys from 15 colonies over a 50-year time period. Credible interval (CRI) coverage describes the proportion of simulations where the 95% credible interval of the estimated regional trend contained the "true" simulated regional trend.

Scenario	Trend Length	Mean Bias in Trend Estimates	Standard Deviation of Bias in Trend Estimates	95% CRI

SD of shared environmental driver = 0 (no shared trajectories)	50 years	-0.14	0.26	0.91
	10 years	-0.03	1.12	0.91
SD of shared environmental driver = 0.1 (partially shared trajectories)	50 years	-0.10	0.53	0.82
	10 years	0.09	2.63	0.79