G.M.Berberich, C.F. Dormann, D. Klimetzek, M.B. Berberich, N.J. Sanders & A.M. Ellison

Appendix S1

Detailed methods of simulating how many observers would be needed to estimate accurately the number of nests in a site

We started with the assumption that each observer had a probability P_d of detecting a given RWA. Because we observed that some nests consistently were detected (or overlooked), we defined P_c to be the probability that observer i detected a RWA nest that was not detected by observer j, $i > j$. P_c is “complementarity for zeros”, i.e., it is a conditional probability of finding a nest where the previous observer did not: $P_c = P(i+1 = 1 \mid i = 0)$. For a series of n observers $\mathbf{i} = \{i_1, i_2, \ldots i_n\}$ visiting the same site s, the probability that a given nest has been overlooked is determined recursively:

$$
\begin{aligned}
&P(i_1 = 0) = 1 - P_d;\\
&P(i_2 = 0, i_1 = 0) = P(i_2 = 0 \mid i_1 = 0)\, P(i_1 = 0) = (1-P_c)\,(1-P_d)\\
&P(i_3 = 0, i_2 = 0, i_1 = 0) = \mathrm{P}(i_3 = 0 \mid i_2 = 0, i_1 = 0) = (1-P_c)\,(1-P_c)\,(1-P_d) \qquad \text{(B1)}\\
&\ldots\\
&\mathrm{P}(i_n = 0, \ldots i_1 = 0) = (1-P_c)^{n-1}(1-P_d).
\end{aligned}
$$

Intuitively, equation B1 means that the probability of n observers overlooking a nest is dependent on the detection probability of the first, and the complementarity score of all subsequent observers. We had an estimate of the (average) detection probability from our initial maximum-likelihood computations (P_d = 0.42; Table 1), so we computed, for any pair of observers, the proportion of visits the second observer found a nest that was previously undetected. On average, this quantity is $P_c = P(i_2 = 1 \mid i_1 = 0) = 0.25$. In other words, 65% of the effort of each additional observer could be considered to be redundant (wasted). The quantity P_c quantified the correlation between observers and could not be expressed in terms of P_d. As both P_d and P_c were estimated from the data and hence were random variables, we bootstrapped the above function using random draws from the observed values of P_d and P_c to compute the variance. To compute P_d, we counted, for any pair of observers, how often a 0 of observer A was complemented by a 1 of observer B. Finally, we noted that as the number of observers, n, increased, the probability of overlooking any individual nest decreased.

With these estimates in hand, we then asked: how many observers would be needed to come within x nests of the true number of nests, N_s, i.e., to reduce the probability of overlooking a nest $P(i_n = 0, \ldots i_1 = 0) = (1–P_c)^{n–1}(1–P_d)$ to less than a fixed quantity (e.g., 10%). The inset in Fig. 3 (main text) shows the bootstrapped probability of overlooking a nest with indicated targets at 10%, 5%, and 1% (horizontal dashed lines in the inset to Fig. 3 in the main text). To determine these values, we randomly drew an observer (with a detection probability determined from Table 1), drew a second observer randomly, looked up the overlooking rate for the second observer given the first observer (computed from the data), then drew another and so on. This simulation was repeated 1000 times (R code provided in Appendix S2).