

Appendix E Description of procedures for estimation of annual removals

Equation 3 in the main text is based on the following logic, illustrated with a scalar equation for simplicity. Let

$$h_{(t)} = r_{(t)}n_{(t-1)}$$

where h_t is the number of animals removed from the population during $t - 1$ to t and r_t is the proportion of the population alive at $t - 1$ that are removed during $t - 1$ to t . Define the probability of mortality from natural causes as $1 - p$ so that survival excluding removal mortality is p . The number of animals that survive from one time step to the next is properly modeled as

$$n_{(t)} = (n_{(t-1)}p^q - h_{(t)})p^{1-q}$$

where q is the proportion of the time step elapsed before removals occur and $h_{(t)} \leq n_{(t-1)}p^q$.

Substituting $r_t n_{t-1}$ for h_t , it follows that

$$\begin{aligned} n_t &= (n_{t-1}p^q - r_t n_{t-1})p^{1-q} \\ &= n_{t-1}p^q p^{1-q} - r_t n_{t-1}p^{1-q} \\ &= n_{t-1}(p^q p^{1-q} - r_t p^{1-q}) \\ &= n_{t-1}(p - r_t p^{1-q}) \end{aligned}$$

Therefore, the probability that an individual survives removals and natural causes of mortality is $p - r_t p^{1-q}$. This expression allows us to use a time invariant p and a time varying expression r_t informed by data, as described below. This approach was applied to each of the model stages where p_i was an estimated parameter and $r_{(i,t)}$ was estimated from data as follows.

The vector of proportions of animals removed from each age class \mathbf{r}_t was estimated as the number of animals removed from each age class divided by the total number of animals in the population.

We first estimated the number of individuals in each state removed from at time t as

$\mathbf{y}_{\text{removals}(t)} \times \mathbf{r}_{\text{comp}(t)} \times \mathbf{r}_{\text{sero}(t)} \times \mathbf{r}_{\text{infect}(t)}$, where $\mathbf{y}_{\text{removals}(t)}$ is the number of removals at time t observed without error, $\mathbf{r}_{\text{comp}(t)}$ is a stochastic vector of proportions of the removals in each age class; \mathbf{r}_{sero} is a stochastic vector of proportions of age classes that were exposed and susceptible, and $\mathbf{r}_{\text{infect}(t)}$ is a stochastic vector assigning infection status to adult females. The vectors are stochastic because they are estimated from a sample of the individuals removed.

We estimated the proportion of removals in each age / sex class that occurred during time $t - 1 \rightarrow t$ as the four element vector $\tilde{\mathbf{r}}_{\text{comp}[t]}$,

$$\left[\tilde{\mathbf{r}}_{\text{comp}(t)} \mid \mathbf{y}_{\text{r.age}(t)} \right] = \text{Dirichlet} \left(y_{\text{r.age}(1,t)} + 1, y_{\text{r.age}(2,t)} + 1, y_{\text{r.age}(3,t)} + 1, y_{\text{r.age}(4,t)} + 1 \right) \quad (\text{E.1})$$

The data vector $\mathbf{y}_{\text{r.age}(t)}$ contains the number of juveniles ($y_{\text{r.age}(1,t)}$), yearling females ($y_{\text{r.age}(2,t)}$), adult females ($y_{\text{r.age}(3,t)}$) and males yearling and older ($y_{\text{r.age}(4,t)}$) in the annual sample of removals that occurred during $t - 1 \rightarrow t$. Equation E.1 calculates the posterior distribution of the age/sex composition of removals using a Dirichlet conjugate for the multinomial likelihood of the number of individuals in each age class and the total number sampled, assuming no prior information. We used $\tilde{\mathbf{r}}$ to compose an eight element vector for assigning age and sex of the total removals as

$$\mathbf{r}_{\text{comp}(t)} = \left(\tilde{r}_{\text{comp}(1,t)}, \tilde{r}_{\text{comp}(2,t)}, \tilde{r}_{\text{comp}(3,t)}, \tilde{r}_{\text{comp}(1,t)}, \tilde{r}_{\text{comp}(2,t)}, \tilde{r}_{\text{comp}(3,t)}, \tilde{r}_{\text{comp}(3,t)}, \tilde{r}_{\text{comp}(4,t)} \right)'$$

It is important to note that this vector does not sum to one because the *product* of the vectors $\mathbf{r}_{\text{comp}(t)} \cdot \mathbf{r}_{\text{sero}(t)} \cdot \mathbf{r}_{\text{infect}(t)}$ must sum to one.

The serological composition of removals was estimated from a sample of animals removed that

were tested for brucellosis. We composed a vector for assigning serological state to removals as

$$\mathbf{r}_{sero}(t) = (1 - r_{sero(1,t)}, 1 - r_{sero(2,t)}, 1 - r_{sero(3,t)}, r_{sero(4,t)}, r_{sero(5,t)}, r_{sero(6,t)}, r_{sero(7,t)}, 1)'$$

where $r_{sero(j,t)}$ is the proportion of age class j that were truly exposed, estimated using

$$\begin{aligned} & [r_{sero(j,t)}, z_{r.ser}(j,t), \pi, \rho | \mathbf{y}_{r.ser}(t)] \propto \\ & \text{binomial} \left(y_{r.ser}(j,1,t) \middle| \frac{\pi z_{r.ser}(j,t) + \rho (y_{r.ser}(j,2,t) - z_{r.ser}(j,t))}{y_{r.ser}(j,2,t)} \right) \text{binomial} (z_{r.ser}(j,t) | r_{sero(j,t)}, y_{r.ser}(j,2,t)) \times \\ & \text{beta} (r_{sero(j,t)} | 1, 1) \text{beta} (\pi | 476, 42) \text{beta} (\rho | 403, 15467) \quad (\text{E.2}) \end{aligned}$$

where $z_{r.ser}(j,t)$ is the true number of exposed animals of age class j removed at time t ; $y_{r.ser}(j,1,t)$ is the observed number of seropositive animals of age class j and $y_{r.ser}(j,2,t)$ is the number of animals in age class j that were tested. The parameter π specifies the probability of a positive test result for an animal that has been exposed and ρ gives the probability of a positive test result for an animal that has not been exposed. By including the probability of a true positive and the probability of a false positive in equation E.2 we account for uncertainty in the serological test as well as uncertainty created by sampling. Methods for developing informative prior distributions on π and ρ are given in the section *Parameter models*.

The proportion of seropositives that were infectious or recovered was not observed in the serology data. The proportion of adult seropositive females that were infectious was unknown. To include this latent quantity in the calculation of annual removals, we defined a vector,

$$\mathbf{r}_{infect}(t) = \left(1, 1, 1, 1, 1, \frac{n_{(6,t-1)}}{n_{(6,t-1)} + n_{(7,t-1)}}, \frac{n_{(7,t-1)}}{n_{(6,t-1)} + n_{(7,t-1)}}, 1 \right)'$$

This represents the assumption that the proportion of infectious animals removed was the same as

their proportion in the population. We calculated the proportion of state i removed during $t - 1$ as

$$r_{(i,t)} = \frac{y_{\text{removals}(t)} \cdot r_{\text{comp}(i,t)} \cdot r_{\text{sero}(i,t)} \cdot r_{\text{infect}(i,t)}}{n_{(i,t-1)}}.$$