Compiled Monday 11th March, 2019

Supplemental Material for manuscript **"Use of stochastic patch occupancy models in the California red-legged frog for Bayesian inference regarding past events and future persistence"**

Contents

Cross-references between the main manuscript and the supplemental material **SM 3**

Appendix S1. Additional methodological considerations for the study system **SM 4**
- Independence of the creeks . . . SM 4
- Closure assumption . . . SM 4

Appendix S2. Building the posterior distribution of the parameters from the general SPOM **SM 4**
- State transitions . . . SM 4
- Occupancy detection . . . SM 5
- Likelihood function . . . SM 5
- Estimating the shared parameters from their likelihood . . . SM 6

Appendix S3. Building the posterior distribution of the parameters under different hypotheses **SM 8**
- Likelihood functions . . . SM 8
- Estimating the parameters from their likelihood . . . SM 9

Appendix S4. Data imputation from the *a posteriori* estimates of the shared parameters **SM 10**
- Method . . . SM 10
- Results . . . SM 10

Appendix S5. Numerical computation of the posterior distribution **SM 11**
- Numerical computation of the posterior distribution of the shared parameters . . . SM 11
- Numerical computation of the posterior distribution under different hypotheses . . . SM 12
- Many-patches approximation . . . SM 14

Supplementary tables **SM 17**

Supplementary figures **SM 21**

List of Tables

Appendix S6 **Parameter estimation.** . . . SM 17
Appendix S7 **Hypothesis tests.** . . . SM 18
Appendix S8 **Model selection using the Akaike Information Criterion.** . . . SM 19
Appendix S9 **Parameter estimation under the exact and approximate algorithms, as a function of the number of states from each year retained in the approximation, m.** . . . SM 20

List of Figures

Appendix S10 **Bayesian parameter estimation of the mean dispersal distance (α^{-1}) of the California red-legged frog.** . . . . SM 21
Appendix S11 **Bayesian parameter estimation of the probability of detection (p), the extinction parameter (e), and the colonization parameter (c) of the California red-legged frog in Deer Creek, using an uninformative prior (eq. S2.9) for the missing occupancy in the initial year (1998).** . . . . SM 22
Appendix S12 **Accuracy of the Bayesian parameter estimation of the extinction and colonization parameters from Fig. 3.** . . . . SM 23
Appendix S13 **Accuracy of the Bayesian model comparison from Table Appendix S7.** . . . . SM 24
Appendix S14 **Bayesian estimation of the timing of the population reduction event, under two hypotheses.** . . . . SM 25
Appendix S15 **Bayesian estimation of model parameters under the *in situ* die-off hypothesis (H1) in Matadero Creek.** . . . . SM 26
Appendix S16 **Bayesian estimation of model parameters for the California red-legged frog, under the source population loss hypothesis instead of the *in situ* die-off hypothesis presented in Fig. 4.** . . . . SM 27
Appendix S17 **Estimated trajectories of the occupancy and segment occupancy of three creeks.** . . . . SM 28
Appendix S18 **Bayesian parameter estimation of the probability of detection (p), the extinction parameter (e), and the colonization parameter (c) of the California red-legged frog in the three creeks, using the many-patches approximate algorithm.** 29

Cross-references between the main manuscript and the supplemental material

The **methods** section refers to the following supplemental material:

- the DATA section refers to Appendix S1, to Table Appendix S19 (spreadsheet in a separate file), and to Table Appendix S20 (spreadsheet in a separate file)
- the GENERAL SPOM section refers to the entire Appendix S2, to eq. S2.5, and to eq. S5.7
- the HYPOTHETICAL CAUSES OF POPULATION DECLINE section refers to Appendix S3, to both eqs. S3.3 and S3.5, and to eq. S3.6
- the BAYESIAN PARAMETER ESTIMATION section refers to eq. S2.14, to Appendix S4, to eq. S3.11, to eqs. S5.4, S5.6 and S5.8, and to Supplementary file S21 (zip file with the MIDASPOM program in a separate file)
- the PARAMETRIZATION section refers to Appendix S2 and Appendix S3 (in particular to eqs. S2.9–S2.12 and eqs. S3.7–S3.9), and to Appendix S19 twice
- the HYPOTHESIS TESTING section refers to eqs. S3.7–S3.10
- the EXTINCTION PROBABILITY UNDER MANAGEMENT SCENARIOS section refers to eq. S2.14, to eqs. S3.2 and S3.4, and to Table Appendix S6 twice

The **results** section refers to the following supplemental material:

- the ESTIMATION OF SHARED MODEL PARAMETERS section refers to Fig. Appendix S10, to Fig. Appendix S11, and to Fig. Appendix S12
- the HYPOTHESIS TESTING section refers to Tables Appendix S7 and Appendix S8, and to Fig. Appendix S13
- the PARAMETER ESTIMATION section refers to Fig. Appendix S14 twice, to Fig. Appendix S15, and to Fig. Appendix S16
- the EXTINCTION PROBABILITY UNDER MANAGEMENT SCENARIOS section refers to Fig. Appendix S17

The **discussion** section refers to Fig. Appendix S10, and to Table Appendix S7

The **supporting elements** (Figures and Tables) refer to the following supplemental material:

- Table 1 refers to Data Table Appendix S19 (spreadsheet in a separate file), to eqs. S2.9–S2.12 and eqs. S3.7–S3.9, and to Appendix S2 and Appendix S3.
- the caption of Fig. 3 refers to eq. S2.12, to eq. S5.7, to eq. S2.10, and to Fig. Appendix S11
- the caption of Fig. 4 refers to eq. S3.7, to eq. S3.11, to Table Appendix S7, and to Figs. Appendix S15 and Appendix S16

Appendix S1. Additional methodological considerations for the study system

INDEPENDENCE OF THE CREEKS

We considered the three creeks independently for the analyses because the creeks displayed notable differences in creek characteristics, and because localized analyses facilitate actionable conservation management. Although San Francisquito Creek is beyond the annual migration distance for California red-legged frog (Fellers and Kleeman 2007), Matadero and Deer Creeks meet at a confluence, and some frog dispersal might occur between them. Nevertheless, the surveys suggest that occupancy in one creek has little influence on occupancy in the second: the nearest Matadero Creek segment to Deer Creek is usually occupied (14/20 years surveyed), but the closest Deer Creek segment is usually unoccupied (15/19 years). In addition, because Deer and Matadero Creeks differ in their biotic and abiotic characteristics (including predators and agricultural use), modeling Deer and Matadero Creek separately helps target best management strategies and locations for habitat enhancement.

CLOSURE ASSUMPTION

Following MacKenzie *et al.* (2003), we made a closure assumption, namely that segment occupancy remained constant between surveys of a given year. This assumption is likely appropriate because surveys were completed during summer months (June-August), when creeks are reduced to a series of pools with refugia. California red-legged frog movements >30m occurred only between the months of October and May in Point Reyes, California (Fellers and Kleeman 2007). In addition, the average time between the first and last survey of a segment in a given year was short (∼2 months), likely ensuring closure in most surveys.

Appendix S2. Building the posterior distribution of the parameters from the general SPOM

STATE TRANSITIONS

Let $\Theta = (e, c, \alpha, K_1, ..., K_N, d_{11}, ..., d_{NN})$ be the vector of model parameters, let $\mathbf{z_t} = (z_{1,t}, ..., z_{N,t})$ be the vector of true (hidden) segment occupancies at the beginning of year t (before the extinction phase), let $\mathbf{z'_t} = (z'_{1,t}, ..., z'_{N,t})$ be the vector of true segment occupancies after the extinction phase of year t, and let $\mathbf{z_{t+1}} = (z_{1,t+1}, ..., z_{N,t+1})$ be the vector of segment occupancies at the beginning of year $t+1$ (after the colonization phase of year t).

The probability that segment i is in state $z'_{i,t}$ depends only on its own extinction probability, E_i, and on its previous occupancy, $z_{i,t}$. Extinction of segment i is possible only if $z_{i,t} = 1$ and $z'_{i,t} = 0$ — that is, if $z_{i,t}(1 - z'_{i,t}) = 1$. Extinction then occurs with probability E_i. Non-extinction of segment i is possible under two cases. If $z_{i,t} = 1$ and $z'_{i,t} = 1$ — that is, if $z_{i,t}z'_{i,t} = 1$ — then non-extinction occurs with probability $1 - E_i$. If $z_{i,t} = 0$ and $z'_{i,t} = 0$ — that is, if $(1 - z_{i,t})(1 - z'_{i,t}) = 1$ — then non-extinction occurs with certainty (trivially). The case of $z_{i,t} = 0$ and $z'_{i,t} = 1$ is not permissible, because the extinction phase cannot convert a patch from unoccupied to occupied. Because extinction events in all segments are independent, the probability of a transition from state $z_{i,t}$ to state $z'_{i,t}$ is:

$$\Pr(\mathbf{z'_t}|\mathbf{z_t}, \Theta) = \prod_{i=1}^{N} \Big[z_{i,t}(1 - z'_{i,t})E_i + z_{i,t}z'_{i,t}(1 - E_i) + (1 - z_{i,t})(1 - z'_{i,t}) \Big]. \tag{S2.1}$$

This product proceeds over all N patches in the habitat.

The probability that segment i is in state $z_{i,t+1}$ depends on its colonization probability, $C_{i,t}$—which is a function of $\mathbf{z'_t}$, the occupancy of all other segments after the extinction phase—and on its previous occupancy, $z'_{i,t}$,

$$\Pr(\mathbf{z_{t+1}}|\mathbf{z'_t}, \Theta) = \prod_{i=1}^{N} \Big[(1 - z'_{i,t})z_{i,t+1}C_{i,t} + (1 - z'_{i,t})(1 - z_{i,t+1})(1 - C_{i,t}) + z'_{i,t}z_{i,t+1}\Big]. \tag{S2.2}$$

Here, analogously to the extinction phase, we are using the fact that $z'_{i,t} = 0$ and $z_{i,t+1} = 1$ produces the first term, $z'_{i,t} = 0$ and $z_{i,t+1} = 0$ the second term, $z'_{i,t} = 1$ and $z_{i,t+1} = 1$ the third term, and $z'_{i,t} = 1$ and $z_{i,t+1} = 0$ is impermissible.

We obtain the probability of transition from occupancy vector $\mathbf{z_t}$ to vector $\mathbf{z_{t+1}}$ by summing the product of the transition probabilities $\Pr(\mathbf{z_{t+1}}|\mathbf{z'_t}, \Theta)$ (eq. S2.1) and $\Pr(\mathbf{z'_t}|\mathbf{z_t}, \Theta)$ (eq. S2.2) over the set of all possible intermediate states $\mathbf{z'_t}$. This set has 2^N possible vectors, where N is the number of segments:

$$\Pr(\mathbf{z_{t+1}}|\mathbf{z_t}, \Theta) = \sum_{\mathbf{z'_t}} \Pr(\mathbf{z_{t+1}}|\Theta, \mathbf{z'_t}) \Pr(\mathbf{z'_t}|\Theta, \mathbf{z_t}). \tag{S2.3}$$

OCCUPANCY DETECTION

Because occupancy detection is imperfect, the values of the occupancy variables $z_{i,t}$ are not known. Rather, several possible states can match the observed data. As a result, to compute the probability of the observed data, we must compute probabilities for all possible values of the unknown occupancies.

Let $J_{i,t}$ be the the number of surveys in segment i and year t and let $Y_{i,j,t}$ be the observed occupancy in the jth survey of segment i in year t, where $Y_{i,j,t} = 0$ or 1 for all segments $1 \leq i \leq N$, surveys $1 \leq j \leq J_{i,t}$ and all years $1 \leq t \leq T$. At each survey, the probability of detection given species presence is p. Following MacKenzie *et al.* (2002, 2003):

$$\begin{aligned} &\Pr(Y_{i,1,t} = x_1, ..., Y_{i,J_{i,t},t} = x_{J_{i,t}}|z_{i,t} = 1) = \prod_{j=1}^{J_{i,t}} p^{x_j}(1-p)^{1-x_j}, \\ &\Pr(Y_{i,1,t} = x_1, ..., Y_{i,J_{i,t},t} = x_{J_{i,t}}|z_{i,t} = 0) = \begin{cases} 1, \text{if } x_1 = x_2 = ... = x_{J_{i,t}} = 0 \\ 0, \text{otherwise.} \end{cases} \end{aligned} \tag{S2.4}$$

Because we consider surveys to be independent, we obtain the probability of observing the $1 \times (\sum_{i=1}^{N} J_{i,t})$ vector $\mathbf{Y_t} = (Y_{1,1,t}, Y_{1,2,t}, ..., Y_{1,J_{1,t},t}, Y_{2,1,t}, ..., Y_{2,J_{2,t},t}, ..., Y_{N,J_{N,t},t})$ given true occupancy $\mathbf{z_t}$ by multiplying the probabilities across segments:

$$\Pr(\mathbf{Y_t}|\mathbf{z_t}) = \prod_{i=1}^{N} \prod_{j=1}^{J_{i,t}} \Pr(Y_{i,j,t}|z_{i,t}). \tag{S2.5}$$

LIKELIHOOD FUNCTION

For each year t, denote each possible state of the occupancy vector $\mathbf{z_t}$ by a number between 1 and 2^N. We denote by $\mathbf{q_t}$ the $2^N \times 1$ column vector containing the values of $\Pr(\mathbf{Y_t}|\mathbf{z_t})$ (computed from eq. S2.5), and by $D(\mathbf{q_t})$ the $2^N \times 2^N$ diagonal matrix in which elements on the diagonal correspond to $\mathbf{q_t}$; we further denote by $\phi_\mathbf{0}$ the 1×2^N row vector of the initial probabilities of each possible state $\mathbf{z_1}$ in the initial year (eqs. S2.9 and S2.10). The probability of the observed data $\mathbf{Y_1}$ in the first year is:

$$\begin{aligned}\Pr(\mathbf{Y_1}) &= \sum_{\mathbf{z_1}} \Pr(\mathbf{z_1}) \Pr(\mathbf{Y_1}|\mathbf{z_1}), \\ &= \phi_{\mathbf{0}} \mathbf{q_1}.\end{aligned} \tag{S2.6}$$

This sum proceeds over all 2^N possible occupancy states $\mathbf{z_1}$.

We denote by $\phi_{\mathbf{t}}(\Theta)$ the $2^N \times 2^N$ yearly transition matrix, where element $\phi_{tk\ell}$ corresponds to the probability of transition from state k at time t to state ℓ at time $t+1$, as computed from eq. S2.3. Given the model parameters Θ, the probability of the two observed vectors $\mathbf{Y_1}$ and $\mathbf{Y_2}$ is:

$$\begin{aligned}\Pr(\mathbf{Y_1}, \mathbf{Y_2}|\Theta) &= \sum_{\mathbf{z_1}} \sum_{\mathbf{z_2}} \Pr(\mathbf{z_1}) \Pr(\mathbf{Y_1}|\mathbf{z_1}) \Pr(\mathbf{z_2}|\mathbf{z_1}, \Theta) \Pr(\mathbf{Y_2}|\mathbf{z_2}) \\ &= \phi_{\mathbf{0}} D(\mathbf{q_1}) \phi_{\mathbf{1}}(\Theta) \mathbf{q_2}.\end{aligned} \tag{S2.7}$$

Similarly, we obtain the likelihood of the parameters given all observations $\mathbf{Y_1}, ..., \mathbf{Y_T}$:

$$\mathcal{L}(\Theta|\mathbf{Y_1}, ..., \mathbf{Y_T}) = \phi_{\mathbf{0}} \left[\prod_{t=1}^{T-1} D(\mathbf{q_t}) \phi_{\mathbf{t}}(\Theta) \right] \mathbf{q_T}. \tag{S2.8}$$

Note that eq. S2.8 is equal to eq. 5 from MacKenzie *et al.* (2003). Nevertheless, vectors $\phi_{\mathbf{0}}$ and $\mathbf{q_T}$, and matrices $D(\mathbf{q_t})$ and $\phi_{\mathbf{t}}(\Theta)$ have different expressions, owing to the fact that our model differs from that of MacKenzie *et al.* (2003) in terms of extinction and colonization dynamics.

ESTIMATING THE SHARED PARAMETERS FROM THEIR LIKELIHOOD

With the likelihood function of the shared parameters (eq. 1), and assuming they have specified prior distributions, we can obtain parameter estimates and credible intervals by computing their posterior distribution using Bayes' theorem. Elements of vector $\phi_{\mathbf{1997}}$ that correspond to possible states $\mathbf{z_{1997}}$ lie in the open interval $(0, 1)$, and their value reflects the prior probability of the states in 1997. We consider either an uninformative or an informative prior. Our uninformative prior is a discrete uniform distribution over the set $\mathcal{A}_{1997}$ of all possible states in 1997 (2^m elements, where m is the number of patches with uncertain occupancy in 1997). Denoting by $\phi_{1997,k}$ the prior probability of state k,

$$\phi_{1997,k} = \begin{cases} \dfrac{1}{2^m}, \text{ if } k \in \mathcal{A}_{1997} \\ 0, \text{otherwise.} \end{cases} \tag{S2.9}$$

Our informative prior corresponds has the property that, for each patch with missing data, the occupancy follows a Bernoulli distribution with parameter equal to the mean occupancy of non-missing data in $\mathbf{Y_{1997}}$, denoted by $\bar{z}_{1997}$. As a result, given a state k with m patches with missing data, its probability of having exactly m_0 specific patches where the missing data is 0 (unoccupied) and $m_1 = m - m_0$ remaining patches where the missing data is 1 (occupied), is

$$\phi_{1997,k} = (1 - \bar{z}_{1997})^{m_0} \bar{z}_{1997}^{m_1}. \tag{S2.10}$$

We consider a uniform prior for the mean dispersal distance α^{-1}, measured in meters,

$$\Pr(\alpha^{-1}) = \begin{cases} 1, \text{if } \alpha^{-1} \in [50, 500] \\ 0, \text{otherwise.} \end{cases} \tag{S2.11}$$

We consider uniform priors for the model parameters (p, e, c),

$$\Pr(p, e, c) = \begin{cases} \dfrac{1}{1.5}, \text{ if } p \in [0,1], e \in [0,1], \text{and } c \in [0, 1.5] \\ 0, \text{otherwise.} \end{cases} \tag{S2.12}$$

Note that because p is a probability, a prior defined on $[0, 1]$ covers its entire range. Similarly, although e is not a probability, because we set $K = 1$ in the general SPOM, $E = e/K = e$, and e is also equivalent to a probability and restricted to the interval $[0, 1]$. c is not a probability, and could be greater than 1. In the case of the California red-legged frog, we found that setting a maximum of 1.5 for the prior was enough to cover the parameter space where the posterior is large (see Fig. 3). In order to accommodate other uses, we allow the range of the prior to be set by the user in our implementation MIDASPOM.

Multiplying the prior probability $\Pr(\mathbf{\Theta_0})$ of the model parameters (eqs. S2.11, S2.12) by the likelihood of the parameters given the occupancy dataset between 1997 and 2016 (eq. 1), where the probability of all possible states in 1997 corresponds either to eq. S2.9 or to eq. S2.10, we obtain the posterior distribution of the parameters given the observed data:

$$\Pr(\mathbf{\Theta_0}|\mathbf{Y_{1997}}, ..., \mathbf{Y_{2016}}) \propto \Pr(\mathbf{\Theta_0})\mathcal{L}(\mathbf{\Theta_0}|\mathbf{Y}_{1997}, ..., \mathbf{Y}_{2016}). \tag{S2.13}$$

When a single dataset $\mathbf{Y_{1997}}, ..., \mathbf{Y_{2016}}$ is considered, all parameters are estimated jointly, and the mode of the joint posterior distribution is then used to obtain maximum *a posteriori* estimates $\widetilde{\alpha^{-1}}$, $\tilde{e}$, $\tilde{c}$, in each creek; the 2.5% and 97.5% quantiles of the marginal posterior distributions are used to construct 95% credible intervals. When M datasets $\mathbf{Y^X} = (\mathbf{Y^X_{1997}}, ..., \mathbf{Y^X_{2016}})$ are considered (e.g., $M = 3$ independent creeks $\mathbf{Y^1}$, $\mathbf{Y^2}$, and $\mathbf{Y^3}$), the dispersal distance is estimated first, because it is assumed to be a property of the species and thus the same for all datasets. Because we assume the datasets to be independent, the joint likelihood of the parameters of the datasets is the product of the likelihoods of each dataset. The posterior distribution of α is then obtained by multiplying the likelihood of the parameters by their prior distribution, and integrating over all possible values of parameters e and c:

$$\Pr(\alpha^{-1}|\mathbf{Y^1}, ..., \mathbf{Y^M}) \int\int_{(e=0,c=0)}^{(1,1.5)} \prod_{X=1}^{M} [\Pr(\mathbf{\Theta_0})\mathcal{L}(\mathbf{\Theta_0}|\mathbf{Y^X})]. \tag{S2.14}$$

The mode of the posterior distribution is then used to obtain a maximum *a posteriori* estimate $\widetilde{\alpha^{-1}}$. Parameters e and c are then estimated independently for each dataset assuming a mean dispersal distance of $\widetilde{\alpha^{-1}}$ using eq. S2.13.

The numerical evaluation of the function proceeded as described in Appendix S5, using eq. S5.2. We built an approximate likelihood function, denoted by $\mathcal{L}(\mathbf{\Theta_0}|\mathbf{Y_{1997}}, ..., \mathbf{Y_{2016}})$ (eq. S5.7; see Appendix S5).

Appendix S3. Building the posterior distribution of the parameters under different hypotheses

LIKELIHOOD FUNCTIONS

In this section, we derive the likelihood of the model parameters Θ_h used for Bayesian inference under each hypothesis (eq. 2). We can divide the likelihood computation into two parts: the likelihood for the years 1902 to t_e, and the likelihood for the years t_e to 1997. The likelihood after t_e does not depend on the parameters under each hypothesis Θ_h, so its expression is similar under the two hypotheses:

$$\left[\prod_{t=t_e}^{1996} D(\mathbf{q_t})\phi_\mathbf{t}(\mathbf{\Theta_0})\right]\mathbf{q_{1997}}, \tag{S3.1}$$

where $\phi_\mathbf{t}(\mathbf{\Theta_0})$ is computed as in eq. S2.8. Because the product in eq. S3.1 proceeds over years prior to the onset of data collection (1997), all states are equally likely and the elements of $\mathbf{q_t}$ are all equal to $1/2^N$ and are constant through time. Thus, $D(\mathbf{q_t}) = (1/2^N)\mathbf{I}$ in eq. S3.1, where $\mathbf{I}$ is the $2^N \times 2^N$ identity matrix. $\mathbf{q_{1997}}$ is computed as in eq. S2.8.

The likelihood for the years 1902 to $t_e = t_D$ under hypothesis 1 depends both on $\mathbf{\Theta_0} = (e, c, \alpha)$ and $\mathbf{\Theta_1} = K_D$,

$$\phi_\mathbf{1902}\prod_{t=1902}^{t_D-1} D(\mathbf{q_t})\phi_\mathbf{t}(\mathbf{\Theta_0}, \mathbf{\Theta_1}). \tag{S3.2}$$

To compute $\phi_\mathbf{t}(\mathbf{\Theta_0}, \mathbf{\Theta_1})$, we first compute $\Pr(\mathbf{z'_t}|\mathbf{z_t}, \mathbf{\Theta_0}, \mathbf{\Theta_1})$ from eq. S2.1 using $(\mathbf{\Theta_0}, \mathbf{\Theta_1})$ in place of Θ, and $\Pr(\mathbf{z_{t+1}}|\mathbf{z'_t}, \mathbf{\Theta_0}, \mathbf{\Theta_1})$ from eq. S2.2 using $(\mathbf{\Theta_0}, \mathbf{\Theta_1})$ in place of Θ and setting $C_{i,t} = c\sum_{j=1,j\neq i}^{N}\exp(-\alpha d_{ij})K_D z'_{j,t}$. We then compute $\Pr(\mathbf{z_{t+1}}|\mathbf{z_t}, \mathbf{\Theta_0}, \mathbf{\Theta_1})$ from eq. S2.3 using $\Pr(\mathbf{z'_t}|\mathbf{z_t}, \mathbf{\Theta_0}, \mathbf{\Theta_1})$ in place of $\Pr(\mathbf{z'_t}|\mathbf{z_t}, \Theta)$ and $\Pr(\mathbf{z_{t+1}}|\mathbf{z'_t}, \mathbf{\Theta_0}, \mathbf{\Theta_1})$ in place of $\Pr(\mathbf{z_{t+1}}|\mathbf{z'_t}, \Theta)$. We finally compute $\phi_\mathbf{t}(\mathbf{\Theta_0}, \mathbf{\Theta_1})$ from eq. S2.8 using $\Pr(\mathbf{z_{t+1}}|\mathbf{z_t}, \mathbf{\Theta_0}, \mathbf{\Theta_1})$ in place of $\Pr(\mathbf{z_{t+1}}|\mathbf{z_t}, \Theta)$. Because the product in eq. S3.2 proceeds over years prior to the onset of data collection (1997), $D(\mathbf{q_t}) = (1/2^N)\mathbf{I}$ in eq. S3.2. Combining eqs. S3.1 and S3.2 leads to the likelihood of the parameters under hypothesis 1,

$$\mathcal{L}(\mathbf{\Theta_1}, \mathbf{z_{1902}}|\mathbf{Y_{1997}}, \mathbf{\Theta_0}) = \phi_\mathbf{1902}\left[\prod_{t=1902}^{t_D-1} D(\mathbf{q_t})\phi_\mathbf{t}(\mathbf{\Theta_0}, \mathbf{\Theta_1})\right]\left[\prod_{t=t_D}^{1996} D(\mathbf{q_t})\phi_\mathbf{t}(\mathbf{\Theta_0})\right]\mathbf{q_{1997}}. \tag{S3.3}$$

Similarly, the likelihood for the years 1902 to $t_e = t_L$ under hypothesis 2 depends both on $\mathbf{\Theta_0} = (e, c, \alpha)$ and $\mathbf{\Theta_2} = (K_L, d_L)$,

$$\phi_\mathbf{1902}\prod_{t=1902}^{t_L-1} D(\mathbf{q_t})\phi_\mathbf{t}(\mathbf{\Theta_0}, \mathbf{\Theta_2}). \tag{S3.4}$$

To compute $\phi_\mathbf{t}(\mathbf{\Theta_0}, \mathbf{\Theta_2})$, we first compute $\Pr(\mathbf{z'_t}|\mathbf{z_t}, \mathbf{\Theta_0})$ from eq. S2.1 using $\mathbf{\Theta_0}$ in place of Θ, and $\Pr(\mathbf{z_{t+1}}|\mathbf{z'_t}, \mathbf{\Theta_0}, \mathbf{\Theta_2})$ from eq. S2.2 using $(\mathbf{\Theta_0}, \mathbf{\Theta_2})$ in place of Θ and using $C_{i,t} = c\Big[\sum_{j=1,j\neq i}^{N}\exp(-\alpha d_{ij})Kz'_{j,t} + \exp(-\alpha d_L)K_L\Big]$. We then compute $\Pr(\mathbf{z_{t+1}}|\mathbf{z_t}, \mathbf{\Theta_0}, \mathbf{\Theta_2})$ from eq. S2.3 using $\Pr(\mathbf{z'_t}|\mathbf{z_t}, \mathbf{\Theta_0})$ in place of $\Pr(\mathbf{z'_t}|\mathbf{z_t}, \Theta)$ and $\Pr(\mathbf{z_{t+1}}|\mathbf{z'_t}, \mathbf{\Theta_0}, \mathbf{\Theta_2})$ in place of $\Pr(\mathbf{z_{t+1}}|\mathbf{z'_t}, \Theta)$. We finally compute $\phi_\mathbf{t}(\mathbf{\Theta_0}, \mathbf{\Theta_2})$ from eq. S2.8 using $\Pr(\mathbf{z_{t+1}}|\mathbf{z_t}, \mathbf{\Theta_0}, \mathbf{\Theta_2})$ in place of $\Pr(\mathbf{z_{t+1}}|\mathbf{z_t}, \Theta)$. Because the product in eq. S3.4 proceeds over years prior to the onset of data collection (1997), $D(\mathbf{q_t}) = (1/2^N)\mathbf{I}$ in eq. S3.4. $\mathbf{q_{1997}}$ is computed as in eq. S2.8. Combining eqs. S3.1

and S3.4 leads to the likelihood of the parameters under hypothesis 2,

$$\mathcal{L}(\mathbf{\Theta_2}, \mathbf{z_{1902}}|\mathbf{Y_{1997}}, \mathbf{\Theta_0}) = \boldsymbol{\phi_{1902}} \left[\prod_{t=1902}^{t_L-1} D(\mathbf{q_t})\boldsymbol{\phi_t}(\mathbf{\Theta_0}, \mathbf{\Theta_2})\right] \left[\prod_{t=t_L}^{1996} D(\mathbf{q_t})\boldsymbol{\phi_t}(\mathbf{\Theta_0})\right] \mathbf{q_{1997}}. \tag{S3.5}$$

Note that the likelihood for the years t_e to 1997 under the null hypothesis is equal to that under hypothesis 1 with $K_D = K$, in which case $(\mathbf{\Theta_0}, \mathbf{\Theta_1})$ can be reduced to $\mathbf{\Theta_0}$,

$$\begin{aligned} \mathcal{L}(\mathbf{z_{1902}}|\mathbf{Y_{1997}}, \mathbf{\Theta_0}) &= \boldsymbol{\phi_{1902}} \left[\prod_{t=1902}^{t_D-1} D(\mathbf{q_t})\boldsymbol{\phi_t}(\mathbf{\Theta_0})\right] \left[\prod_{t=t_D}^{1996} D(\mathbf{q_t})\boldsymbol{\phi_t}(\mathbf{\Theta_0})\right] \mathbf{q_{1997}}, \\ &= \boldsymbol{\phi_{1902}} \left[\prod_{t=1902}^{1996} D(\mathbf{q_t})\boldsymbol{\phi_t}(\mathbf{\Theta_0})\right] \mathbf{q_{1997}}. \end{aligned} \tag{S3.6}$$

ESTIMATING THE PARAMETERS FROM THEIR LIKELIHOOD

With the likelihood function of the parameters under each hypothesis, and assuming the parameters have specified prior distributions, we can obtain parameter estimates and credible intervals by computing their posterior distributions using Bayes' theorem. We consider a log-uniform prior between 0.1 and 100 for K_D (resp. K_L),

$$\Pr(K_D = x) = \begin{cases} \dfrac{1}{x[\ln(100) - \ln(0.1)]}, \text{ if } K_D \in [0.1, 100] \\ 0, \text{otherwise}, \end{cases} \tag{S3.7}$$

and a uniform prior between 200 and 4000 for d_L,

$$\Pr(d_L = x) = \begin{cases} \dfrac{1}{3800}, \text{ if } d_L \in [200, 4000] \\ 0, \text{otherwise}. \end{cases} \tag{S3.8}$$

We consider a uniform prior between 1902 and 1982 for t_D and t_L:

$$\Pr(t_D = x) = \begin{cases} \dfrac{1}{1982 - 1902}, \text{ if } t_D \in [1902, 1982] \\ 0, \text{otherwise}. \end{cases} \tag{S3.9}$$

We consider a discrete uniform distribution over the set of all $2^N - 1$ possible non-empty occupancy states for $\mathbf{z_{1882}}$

$$\phi_{1882,k} = \begin{cases} 0,\ k \text{ such that } \mathbf{z_{1882}} = (0, 0, ..., 0) \\ \dfrac{1}{2^N - 1}, \text{otherwise}. \end{cases} \tag{S3.10}$$

We multiply the prior distribution of the parameters (the product of eqs. S3.7 and S3.9 under hypothesis 1, and of eqs. S3.7, S3.8 and S3.9 under hypothesis 2) with the likelihood function under the hypothesis (eq. S3.3 or S3.5), assuming the initial occupancy follows eq. S3.10, to obtain the posterior distribution of Θ_h:

$$\Pr(\mathbf{\Theta_h}, \mathbf{z_{1882}}|\mathbf{Y_{1997}}) \propto \Pr(\mathbf{\Theta_h})\mathcal{L}(\mathbf{\Theta_h}, \mathbf{z_{1882}}|\mathbf{Y_{1997}}). \tag{S3.11}$$

The mode of the posterior distribution is used as a maximum *a posteriori* estimate of a parameter, $\tilde{K}_D$ and $\tilde{t}_D$ under hypothesis 1, and $\tilde{K}_L$, $\tilde{d}_L$ and $\tilde{t}_L$ under hypothesis 2; the 2.5% and 97.5% quantiles of the posterior distributions are used as 95% credible intervals. Note that even though our method provides a posterior distribution for the occupancy in the initial year $\mathbf{z_{1882}}$, this value is not of interest, and we will thus simply integrate the joint posterior distribution of the other parameters over all possible values of $\mathbf{z_{1882}}$. The numerical evaluation of the function proceeded as described in Appendix S5, using eq. S5.4 under hypothesis 1 and eq. S5.6 under hypothesis 2.

Similarly to what was done to approximate the likelihood of the general SPOM, we built an approximate likelihood function for each hypothesis, denoted by $\tilde{\mathcal{L}}(\mathbf{\Theta_h}, \mathbf{z_{1882}}|\mathbf{Y_{1997}})$ (eq. S5.8; see derivation in Appendix S5), where $h = 1$ under hypothesis 1, and $h = 2$ under hypothesis 2. This approximate likelihood only considers the most likely occupancy states instead of all possible states. See Fig. Appendix S13 for an assessment of the accuracy of the model testing using the approximate likelihood. The exact likelihood was used to infer parameters from Matadero and Deer Creeks, while the approximate likelihood was used to infer parameters from San Francisquito Creek.

Appendix S4. Data imputation from the *a posteriori* estimates of the shared parameters

METHOD

An interesting use of the posterior distribution of model parameters is for imputation of missing data; this computation makes it possible, for example, to track temporal changes in patch occupancy. The missing data can be imputed using the maximum *a posteriori* estimates of α^{-1}, p, e, and c (from eqs. S2.13 and S2.14) $\widetilde{\alpha^{-1}}$, $\tilde{p}$, $\tilde{e}$ and $\tilde{c}$. To perform the imputation, for all years $t = 1998, ..., 2016$, we compute the probability vector ψ_t for all possible states in year t:

$$\psi_t = \phi_{1997} \prod_{s=1997}^{t} D(\mathbf{q}_t)\phi_s(\widetilde{\mathbf{\Theta_0}})\mathbf{q}_t, \tag{S4.1}$$

where matrices $\phi_t(\widetilde{\mathbf{\Theta_0}})$ are computed from eq. S2.3 with $\widetilde{\mathbf{\Theta_0}} = (\widetilde{\alpha^{-1}}, \tilde{e}, \tilde{c})$, vectors $\mathbf{q}_t$ are computed from eq. S2.5, and the initial probability ϕ_{1997} is computed assuming a discrete uniform distribution over the set of all possible states (eq. S2.9). ψ_t is a $1 \times 2N$ vector. Then, for each year t, the imputed state is that corresponding to $\max(\psi_t)$.

We can obtain the distribution of the proportion of segments occupied in year t, denoted by r. To do so, for each year t and for r, we sum the elements of ϕ_t corresponding to occupancy vectors with a proportion of occupied segments r.

RESULTS

Over the time frame of the study, all creeks declined in proportion of occupied segments (Fig. Appendix S17(a), (c), (e)). Matadero and Deer Creeks had a higher proportion of segments occupied by *R. draytonii* until 2002-2003, with probably more than 80% occupancy. They then experienced a decline between 2004 and 2007, and have had 30-70% occupancy since 2007. Occupancy in San Francisquito Creek decreased continuously between 1997 and 2007, and likely totally disappeared in 2008.

Although declines in proportion of occupied segments are similar in Matadero and Deer Creeks, their occupancy dynamics were different (Fig. Appendix S17(b) and (d)). In Matadero Creek, segments 5 to 10 became extinct in 2005 and most likely stayed extinct until 2016, while other segments mostly stayed occupied during that period (Fig. Appendix S17(b)). Such dynamics are expected, because of the relatively small extinction and colonization parameters in Matadero Creek. As a result, segment extinction is unlikely, but once it occurs, because colonization is also unlikely, unoccupied segments tend to stay unoccupied for a long time. In contrast, in Deer Creek, segments 2 to 8 have been periodically switching from occupied to unoccupied every 1-4 years (Fig. Appendix S17(d)). This is expected due to the large extinction and colonization parameters estimated that lead to a rapid turnover of occupancy.

Occupancy dynamics in San Francisquito Creek show a long persistence (10 years) of populations in the middle of the creek (segment 9; Fig. Appendix S17(f)), and a gradual disappearance of other populations, with occasional sporadic colonizations (e.g., segment 5 in 2003) and recolonizations (e.g., segment 19 in 1999) of neighboring segments. Such dynamics are expected, because of the large extinction rate that leads to a steady decline of occupancy, and because of the moderate colonization parameter and small dispersal distance that only enable occasional colonizations of segments close to the few occupied ones.

Appendix S5. Numerical computation of the posterior distribution

In this appendix, we present the numerical computation of the posterior distribution of the parameters (eqs. S2.13 and S3.11). Our implementation of the method is written in C, using the BLAS library for numerical algebra computations and the MPI library for parallel computing, and is available under the GNU General Public License.

NUMERICAL COMPUTATION OF THE POSTERIOR DISTRIBUTION OF THE SHARED PARAMETERS

In order to compute the posterior distribution of the parameters (eq. S2.13) across the range of the prior distributions of continuous variables (eqs. S2.12 and S2.11), we evaluate the likelihood function from eq. S2.8 on a regular grid for (e, c, α^{-1}) with a given resolution (default is 0.01 for e and c, and 25 for α^{-1}). We obtain values $\mathcal{L}\big((\frac{a}{100}, \frac{b}{100}, 50 + 25\beta)|\mathbf{Y_1}, ..., \mathbf{Y_T}\big)$ for all integers a from 0 to 100, integers b from 0 to 150, and integers β from 0 to 18.

Because the prior probabilities of e and c are constant (eq. S2.12) across the range considered ([0,1]), and because that of α is 1/450 (eq. S2.11) across the range considered ([50,500]), from eq. S2.13 the posterior probability is proportional solely to the likelihood function multiplied by a factor $1/450$. Computing the proportionality constant thus enables us to obtain the posterior distribution. Because the posterior distribution has an integral of 1, the proportionality constant corresponds to $1/450$ multiplied by the integral L of the likelihood. We approximate this integral by numerical integration over the grid, using the trapezoid rule,

$$
\begin{aligned}
\hat{L} = (0.01 \times \frac{1}{150} \times 25)\Bigg\{ & \frac{1}{8}\Big[\mathcal{L}\left((0,0,50)|\mathbf{Y_1}, ..., \mathbf{Y_T}\right) + \mathcal{L}\left((1,0,50)|\mathbf{Y_1}, ..., \mathbf{Y_T}\right) + \mathcal{L}\left((0,1.5,50)|\mathbf{Y_1}, ..., \mathbf{Y_T}\right) \\
& + \mathcal{L}\left((1,1.5,50)|\mathbf{Y_1}, ..., \mathbf{Y_T}\right) + \mathcal{L}\left((0,0,500)|\mathbf{Y_1}, ..., \mathbf{Y_T}\right) + \mathcal{L}\left((1,0,500)|\mathbf{Y_1}, ..., \mathbf{Y_T}\right) \\
& + \mathcal{L}\left((0,1.5,500)|\mathbf{Y_1}, ..., \mathbf{Y_T}\right) + \mathcal{L}\left((1,1.5,500)|\mathbf{Y_1}, ..., \mathbf{Y_T}\right)\Big] \\
& + \frac{1}{12}\sum_{a=1}^{99}\Bigg[\mathcal{L}\left(\left(\frac{a}{100}, 0, 50\right)|\mathbf{Y_1}, ..., \mathbf{Y_T}\right) + \mathcal{L}\left(\left(\frac{a}{100}, 1.5, 50\right)|\mathbf{Y_1}, ..., \mathbf{Y_T}\right) \\
& + \mathcal{L}\left(\left(\frac{a}{100}, 0, 500\right)|\mathbf{Y_1}, ..., \mathbf{Y_T}\right) + \mathcal{L}\left(\left(\frac{a}{100}, 1.5, 500\right)|\mathbf{Y_1}, ..., \mathbf{Y_T}\right)\Bigg] \\
& + \frac{1}{12}\sum_{b=1}^{149}\Bigg[\mathcal{L}\left(\left(0, \frac{b}{100}, 50\right)|\mathbf{Y_1}, ..., \mathbf{Y_T}\right) + \mathcal{L}\left(\left(1, \frac{b}{100}, 50\right)|\mathbf{Y_1}, ..., \mathbf{Y_T}\right) \\
& + \mathcal{L}\left(\left(0, \frac{b}{100}, 500\right)|\mathbf{Y_1}, ..., \mathbf{Y_T}\right) + \mathcal{L}\left(\left(1, \frac{b}{100}, 500\right)|\mathbf{Y_1}, ..., \mathbf{Y_T}\right)\Bigg]
\end{aligned}
$$

$$+\frac{1}{12}\sum_{\beta=1}^{17}\Big[\mathcal{L}\left((0,0,50{+}25\beta)\,|\mathbf{Y_1},...,\mathbf{Y_T}\right)+\mathcal{L}\left((1,0,50{+}25\beta)\,|\mathbf{Y_1},...,\mathbf{Y_T}\right)$$

$$+\mathcal{L}\left((0,1.5,50{+}25\beta)\,|\mathbf{Y_1},...,\mathbf{Y_T}\right)+\mathcal{L}\left((1,1.5,50{+}25\beta)\,|\mathbf{Y_1},...,\mathbf{Y_T}\right)\Big]$$

$$+\frac{1}{6}\sum_{a=1}^{99}\sum_{b=1}^{149}\left[\mathcal{L}\left(\left(\frac{a}{100},\frac{b}{100},50\right)|\mathbf{Y_1},...,\mathbf{Y_T}\right)+\mathcal{L}\left(\left(\frac{a}{100},\frac{b}{100},500\right)|\mathbf{Y_1},...,\mathbf{Y_T}\right)\right]$$

$$+\frac{1}{6}\sum_{a=1}^{99}\sum_{\beta=1}^{17}\left[\mathcal{L}\left(\left(\frac{a}{100},0,50{+}25\beta\right)|\mathbf{Y_1},...,\mathbf{Y_T}\right)+\mathcal{L}\left(\left(\frac{a}{100},1.5,50{+}25\beta\right)|\mathbf{Y_1},...,\mathbf{Y_T}\right)\right]$$

$$+\frac{1}{6}\sum_{b=1}^{149}\sum_{\beta=1}^{17}\left[\mathcal{L}\left(\left(0,\frac{b}{100},50{+}25\beta\right)|\mathbf{Y_1},...,\mathbf{Y_T}\right)+\mathcal{L}\left(\left(1,\frac{b}{100},50{+}25\beta\right)|\mathbf{Y_1},...,\mathbf{Y_T}\right)\right]$$

$$+\sum_{a=1}^{99}\sum_{b=1}^{149}\sum_{\beta=1}^{17}\mathcal{L}\left(\left(\frac{a}{100},\frac{b}{100},50{+}25\beta\right)|\mathbf{Y_1},...,\mathbf{Y_T}\right)\Bigg\}. \tag{S5.1}$$

Finally we compute a numerical approximation of the posterior probability of the parameters,

$$\Pr(\Theta|\mathbf{Y_1},...,\mathbf{Y_T})=\frac{\mathcal{L}(\Theta|\mathbf{Y_1},...,\mathbf{Y_T})}{\hat{L}}, \tag{S5.2}$$

for Θ values from the grid, where the likelihood $\mathcal{L}$ comes from eq. S2.8. Note that because we used uninformative priors, the prior probability terms in the numerator and denominator cancel out in eq. S5.2 and the posterior probability depends only on the likelihood function.

NUMERICAL COMPUTATION OF THE POSTERIOR DISTRIBUTION UNDER DIFFERENT HYPOTHESES

We similarly obtain the posterior distribution of the parameters under each hypothesis (eq. S3.11) across the range of the prior distributions of continuous variables (eqs. S3.7, S3.8, and S3.9), by evaluating the likelihood function from eq. S3.3 or S3.5 on a regular grid for parameters K_D and t_D, or K_L, d_L, and t_L. Because the prior distributions of K_D and K_L are log-uniform, we evaluate the likelihoods of parameters $\log_{10}(K_D)$ and $\log_{10}(K_L)$, so as to obtain a regular grid (default resolution of 0.02 for $\log_{10}(K_D)$ and $\log_{10}(K_L)$, 200 for d_L, and 5 for t_D and t_L). The likelihood can be used to compute the joint posterior distributions of $\log_{10}(K_D)$ and t_D, and that of $\log_{10}(K_L)$, d_L, and t_L.

Under hypothesis 1, because the prior probability of $\log_{10}(K_D)$ is uniform (eq. S3.7) across the range considered ($[\log_{10}(0.1),\log_{10}(100)]=[-1,2]$), we evaluate the likelihood function from eq. S3.3 at values $\log_{10}(K_D)=\frac{a-50}{50}$, for all integers a ranging from 0 to 150. In addition, because the prior probability of t_D is uniform (eq. S3.9) across the range [1902,1982], we evaluate the likelihood function from eq. S3.3 at values $t_D=5b+1902$, for all integers b ranging from 0 to 16. The prior distribution of $\log_{10}(K_D)$ is $\frac{1}{3}$ over the interval considered ($[-1,2]$), and the prior distribution of t_D is $\frac{1}{81}$ (from eq. S3.9) over the range considered ([1902,1982]). Note that the prior distribution of the initial occupancy $\mathbf{z_{1902}}$ (eq. S3.10) is already included in the likelihood function (eq. S3.3). Thus, similarly to the derivation of eq. S5.2, we compute an approximation of the integral of the likelihood of $(\ln(K_D),t_D,\mathbf{z_{1902}})$ by numerical integration over the grid,

$$\hat{L_1}=\Big(0.02{\times}5\Big)\Bigg\{\frac{1}{4}\Big[\mathcal{L}\Big((10^{-1},1902,\mathbf{z_{1902}})|\mathbf{Y_{1997}}\Big)+\mathcal{L}\Big((10^{-1},1982,\mathbf{z_{1902}})|\mathbf{Y_{1997}}\Big)$$

$$
\begin{aligned}
&+\mathcal{L}\Big((10^{2},1902,\mathbf{z_{1902}})|\mathbf{Y_{1997}}\Big)+\mathcal{L}\Big((10^{2},1982,\mathbf{z_{1902}})|\mathbf{Y_{1997}}\Big)\Big]\\
&+\frac{1}{4}\sum_{a=1}^{149}\Big[\mathcal{L}\Big((10^{\frac{a-50}{50}},1902,\mathbf{z_{1902}})|\mathbf{Y_{1997}}\Big)+\mathcal{L}\Big((10^{\frac{a-50}{50}},1982,\mathbf{z_{1902}})|\mathbf{Y_{1997}}\Big)\Big]\\
&+\frac{1}{4}\sum_{b=1}^{15}\Big[\mathcal{L}\Big((10^{-1},5b{+}1902,\mathbf{z_{1902}})|\mathbf{Y_{1997}}\Big)+\mathcal{L}\Big((10^{2},5b{+}1902,\mathbf{z_{1902}})|\mathbf{Y_{1997}}\Big)\Big]\\
&+\sum_{a=1}^{149}\sum_{b=1}^{15}\mathcal{L}\Big((10^{\frac{a-50}{50}},5b{+}1902,\mathbf{z_{1902}})|\mathbf{Y_{1997}}\Big)\Big\}. \qquad \text{(S5.3)}
\end{aligned}
$$

Finally we compute a numerical approximation of the posterior probability of the parameters under hypothesis 1,

$$
\Pr(\mathbf{\Theta_1},\mathbf{z_{1902}}|\mathbf{Y_{1997}})=\frac{\mathcal{L}(\mathbf{\Theta_1},\mathbf{z_{1902}}|\mathbf{Y_{1997}})}{\hat{L_1}}, \qquad \text{(S5.4)}
$$

for $\mathbf{\Theta_1}$ values from the grid.

Similarly, under hypothesis 2, because the prior probability of $\log_{10}(K_L)$ is uniform (eq. S3.7) across the range considered $[\log_{10}(0.1),\log_{10}(100)]$, we evaluate the likelihood function from eq. S3.5 at values $\log_{10}(K_L)=\frac{a-50}{50}$, for all integers a ranging from 0 to 150. In addition, because the prior probability of t_L has a discrete uniform distribution (eq. S3.9) across the range [1902,1982], we evaluate the likelihood function from eq. S3.5 at values $t_L=5b+1902$, for all integers b ranging from 0 to 16. Finally, because the prior probability of d_L is uniform (eq. S3.8) across the range [200,4000], we evaluate the likelihood function from eq. S3.5 at values $d_L=200\gamma$, where γ ranges from 1 to 20. The prior density of $\log_{10}(K_L)$ is $\frac{1}{3}$ over the interval considered ($[-1,2]$), the prior density of t_L is $\frac{1}{81}$ (from eq. S3.9) over the range considered ([1902,1982]), and the prior density of d_L is $\frac{1}{3800}$ (from eq. S3.8) over the range considered ([200,4000]). Thus, similarly to the derivation of eq. S5.3, we compute an approximation of the integral of the likelihood of $(\ln(K_L),d_L,t_L,\mathbf{z_{1997}})$ by numerical integration over the grid,

$$
\begin{aligned}
\hat{L_2}=\Big(0.02{\times}200{\times}5\Big)\Big\{&\frac{1}{8}\Big[\mathcal{L}\Big((10^{-1},200,1902,\mathbf{t_{1902}})|\mathbf{Y_{1997}}\Big)+\mathcal{L}\Big((10^{2},200,1902,\mathbf{t_{1902}})|\mathbf{Y_{1997}}\Big)\\
&+\mathcal{L}\Big((10^{-1},4000,1902,\mathbf{t_{1902}})|\mathbf{Y_{1997}}\Big)+\mathcal{L}\Big((10^{2},4000,1902,\mathbf{t_{1902}})|\mathbf{Y_{1997}}\Big)\\
&+\mathcal{L}\Big((10^{-1},200,1982,\mathbf{t_{1902}})|\mathbf{Y_{1997}}\Big)+\mathcal{L}\Big((10^{2},200,1982,\mathbf{t_{1902}})|\mathbf{Y_{1997}}\Big)\\
&+\mathcal{L}\Big((10^{-1},4000,1982,\mathbf{t_{1902}})|\mathbf{Y_{1997}}\Big)+\mathcal{L}\Big((10^{2},4000,1982,\mathbf{t_{1902}})|\mathbf{Y_{1997}}\Big)\Big]\\
&+\frac{1}{12}\sum_{a=1}^{149}\Big[\mathcal{L}\Big((10^{\frac{a-50}{50}},200,1902,\mathbf{t_{1902}})|\mathbf{Y_{1997}}\Big)+\mathcal{L}\Big((10^{\frac{a-50}{50}},4000,1902,\mathbf{t_{1902}})|\mathbf{Y_{1997}}\Big)\\
&+\mathcal{L}\Big((10^{\frac{a-50}{50}},200,1982,\mathbf{t_{1902}})|\mathbf{Y_{1997}}\Big)+\mathcal{L}\Big((10^{\frac{a-50}{50}},4000,1982,\mathbf{t_{1902}})|\mathbf{Y_{1997}}\Big)\Big]\\
&+\frac{1}{12}\sum_{\gamma=2}^{19}\Big[\mathcal{L}\Big((10^{-1},200\gamma,1902,\mathbf{t_{1902}})|\mathbf{Y_{1997}}\Big)+\mathcal{L}\Big((10^{2},200\gamma,1902,\mathbf{t_{1902}})|\mathbf{Y_{1997}}\Big)\\
&+\mathcal{L}\Big((10^{-1},200\gamma,1982,\mathbf{t_{1902}})|\mathbf{Y_{1997}}\Big)+\mathcal{L}\Big((10^{2},200\gamma,1982,\mathbf{t_{1902}})|\mathbf{Y_{1997}}\Big)\Big]\\
&+\frac{1}{12}\sum_{b=1}^{15}\Big[\mathcal{L}\Big((10^{-1},200,5b{+}1902,\mathbf{t_{1902}})|\mathbf{Y_{1997}}\Big)+\mathcal{L}\Big((10^{2},200,5b{+}1902,\mathbf{t_{1902}})|\mathbf{Y_{1997}}\Big)\\
&+\mathcal{L}\Big((10^{-1},4000,5b{+}1902,\mathbf{t_{1902}})|\mathbf{Y_{1997}}\Big)+\mathcal{L}\Big((10^{2},4000,5b{+}1902,\mathbf{t_{1902}})|\mathbf{Y_{1997}}\Big)\Big]
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{1}{6}\sum_{a=1}^{149}\sum_{\gamma=2}^{19}\Big[\mathcal{L}\Big((10^{\frac{a-50}{50}},200\gamma,1902,\mathbf{t_{1902}})|\mathbf{Y_{1997}}\Big)+\mathcal{L}\Big((10^{\frac{a-50}{50}},200\gamma,1982,\mathbf{t_{1902}})|\mathbf{Y_{1997}}\Big)\Big]\\
&+\frac{1}{6}\sum_{a=1}^{149}\sum_{b=1}^{15}\Big[\mathcal{L}\Big((10^{\frac{a-50}{50}},200,5b{+}1902,\mathbf{t_{1902}})|\mathbf{Y_{1997}}\Big)+\mathcal{L}\Big((10^{\frac{a-50}{50}},4000,5b{+}1902,\mathbf{t_{1902}})|\mathbf{Y_{1997}}\Big)\Big]\\
&+\frac{1}{6}\sum_{\gamma=2}^{19}\sum_{b=1}^{15}\Big[\mathcal{L}\Big((10^{-1},200\gamma,5b{+}1902,\mathbf{t_{1902}})|\mathbf{Y_{1997}}\Big)+\mathcal{L}\Big((10^{2},200\gamma,5b{+}1902,\mathbf{t_{1902}})|\mathbf{Y_{1997}}\Big)\Big]\\
&+\sum_{a=1}^{149}\sum_{\gamma=2}^{19}\sum_{b=1}^{15}\mathcal{L}\Big((10^{\frac{a-50}{50}},200\gamma,5b{+}1902,\mathbf{t_{1902}})|\mathbf{Y_{1997}}\Big)\Big\}.
\end{aligned}
\tag{S5.5}
$$

Finally we compute a numerical approximation of the posterior probability of the parameters under hypothesis 2,

$$
\Pr(\mathbf{\Theta_2},\mathbf{z_{1902}}|\mathbf{Y_{1997}})=\frac{\mathcal{L}(\mathbf{\Theta_2},\mathbf{z_{1902}}|\mathbf{Y_{1997}})}{\hat{L_2}},
\tag{S5.6}
$$

for $\mathbf{\Theta_2}$ values from the grid.

Note that there are no parameters to estimate under hypothesis 0. Consequently, we do not need a numerical computation and can directly compute the likelihood function from eq. S3.6.

MANY-PATCHES APPROXIMATION

When the number of patches N becomes large, vectors $\mathbf{q_t}$ and matrices $\phi_\mathbf{t}(\Theta)$—which respectively have dimensions $2^N \times 1$ and $2^N \times 2^N$—become too large to compute. In order to solve this issue, we have implemented a sparse approximation algorithm based on the algorithm from Reichel *et al.* (2015) (Algorithm 1). Algorithm 1 consists in approximating the smallest elements of $\mathbf{q_t}$ to zero, that is, to consider that the unlikeliest occupancy states are impossible. The algorithm uses the m most likely states at the beginning of each year, where m is an input parameter, to compute approximate values of vectors $\mathbf{q_t}$ and matrices $\phi_\mathbf{t}(\Theta)$ with reduced dimensions.

Algorithm 1 Many-patches likelihood approximation when values of $\mathbf{q_t}$ are different

for $t := 1$ to T **do**
    Compute vector $\mathbf{q_t}$ from eq. S2.5
    Sort the elements of $\mathbf{q_t}$ from greatest to smallest, and store the corresponding states in variables $\mathbf{z^1}, \mathbf{z^2}, ..., \mathbf{z^{2^N}}$
    Set $A_t := \{\mathbf{z^1}, \mathbf{z^2}, ..., \mathbf{z^m}\}$
Set $\mathcal{A}^* := \bigcup_{t=1}^{T} A_t$, and $m^* := |\mathcal{A}^*|$
for $k := 1$ to m^* **do**
    Compute vector $\phi_\mathbf{k} = \big(\Pr(\mathbf{z_{t+1}}{=}\mathbf{z^1}|\Theta,\mathbf{z_t}{=}\mathbf{z_k}), \Pr(\mathbf{z_{t+1}}{=}\mathbf{z^2}|\Theta,\mathbf{z_t}{=}\mathbf{z_k}), ..., \Pr(\mathbf{z_{t+1}} = \mathbf{z^{2^N}}|\Theta,\mathbf{z_t} = \mathbf{z_k})\big)$ from eq. S2.2
    Sort the elements of $\phi_\mathbf{k}$ from greatest to smallest, and store the corresponding states in variables $\mathbf{z'^1}, \mathbf{z'^2}, ..., \mathbf{z'^{2^N}}$
    Set $E_k := \{\mathbf{z'^1}\}$
Set $\mathcal{A}^{**} := \mathcal{A}^* \cup \left(\bigcup_{k=1}^{m^*} E_k\right)$, and $m^{**} := |\mathcal{A}^{**}|$
Compute $m^{**} \times 1$ vectors $\tilde{\phi_\mathbf{0}}$ and $\tilde{\mathbf{q_t}}$, and $m^{**} \times m^{**}$ matrix $\tilde{\phi_\mathbf{t}}(\Theta)$, using states in $\mathcal{A}^{**}$

The first *for loop* (lines 1–4 in Algorithm 1) finds the m most likely states at the beginning of each year t (sets A_t) given the observed occupancies $\mathbf{Y_t}$. The next step (line 5 in Algorithm 1) computes the union of these sets, $\mathcal{A}^*$, which is at most of size $\min(mT, 2^N)$, if sets A_t are all disjoint. The second *for loop* (lines 6–9 in Algorithm 1) finds the most likely states after the extinction phase of year t, E_k, starting from each state k from A_t. The next step (line 10) computes the union of these sets and $\mathcal{A}^*$, $\mathcal{A}^{**}$, which is at most of size $\min(2mT, 2^N)$, if sets A_t and E_t are all disjoint. This set $\mathcal{A}^{**}$ is the final set of states that are used for the computation of all quantities (line 11 in Algorithm 1). The size of this set, m^{**}, depends on the parameter m; if $m << 2^N$, then $m^{**} << 2^N$ and the approximate likelihood will be much faster to compute than the exact likelihood.

The approximate likelihood of the parameters given all observations $\mathbf{Y_1}, ..., \mathbf{Y_T}$ is thus:

$$\tilde{\mathcal{L}}(\Theta|\mathbf{Y_1}, ..., \mathbf{Y_T}) = \tilde{\phi_{\mathbf{0}}} \left[\prod_{t=1}^{T-1} D(\tilde{\mathbf{q_t}}) \tilde{\phi_{\mathbf{t}}}(\Theta) \right] \tilde{\mathbf{q_T}}. \tag{S5.7}$$

The posterior distributions for the three creeks under the many-patches approximate algorithm are presented in Fig. Appendix S18, along with the exact posterior distributions for the two creeks with the smallest number of segments, Matadero and Deer Creek. Corresponding point estimates and credible intervals are presented in Table Appendix S9. We can see that even for very small values of m ($m = 2$ and $m = 5$), results are very close to those under the exact algorithm. This is due to the fact that few of occupancy states are likely at each step, resulting in many values close to 0 in vectors $\mathbf{q_t}$ and matrices $\phi_{\mathbf{t}}(\Theta)$.

When all elements of $\mathbf{q_t}$ are equal due to the absence of surveys, such as during the time period before sampling, we cannot reduce the number of occupancy states by approximating the smallest elements of $\mathbf{q_t}$ to zero (Algorithm 1). Instead, we approximate the smallest elements of $\phi_{\mathbf{t}}(\Theta)$ to zero (Algorithm 2), by computing the most likely occupancy trajectories from the prior occupancy probability $\phi_{\mathbf{0}}$.

Algorithm 2 Many-patches likelihood approximation when values of $\mathbf{q_t}$ are equal

```
Compute $\phi_{\mathbf{0}}$ from eq. S2.9 or S2.10
Sort the elements of $\phi_{\mathbf{0}}$ from greatest to smallest, and store the corresponding states in variables $\mathbf{z_0^1}, \mathbf{z_0^2}, ..., \mathbf{z_0^{2^N}}$
Set $\mathcal{A}_0 := \{\mathbf{z_0^1}, \mathbf{z_0^2}, ..., \mathbf{z_0^m}\}$
for $k := 1$ to $m$ do
    for $t := 0$ to $T-1$ do
        Compute vector $\phi'_{t,k} = \big(\Pr(\mathbf{z'_t}{=}\mathbf{z_t^1}|\Theta, \mathbf{z_t}{=}\mathbf{z_t^k}), \Pr(\mathbf{z'_t}{=}\mathbf{z_t^2}|\Theta, \mathbf{z_t}{=}\mathbf{z_t^k}), ..., \Pr(\mathbf{z'_t} = \mathbf{z_t^{2^N}}|\Theta, \mathbf{z_t} = \mathbf{z_t^k})\big)$ as in eqs. S3.3, S3.5, or S3.6
        Sort the elements of $\phi'_{t,k}$ from greatest to smallest, and store the corresponding states in variables $\mathbf{z'^1_t}, \mathbf{z'^2_t}, ..., \mathbf{z'^{2^N}_t}$
        Set $E_{t,k} := \{\mathbf{z'^1_t}\}$
        Compute vector $\phi_{t+1,k} = \big(\Pr(\mathbf{z_{t+1}}{=}\mathbf{z_t^1}|\Theta, \mathbf{z'_t}{=}\mathbf{z'^1_t}), \Pr(\mathbf{z_{t+1}}{=}\mathbf{z_t^2}|\Theta, \mathbf{z'_t}{=}\mathbf{z'^1_t}), ..., \Pr(\mathbf{z_{t+1}}{=}\mathbf{z_t^{2^N}}|\Theta, \mathbf{z_t}{=}\mathbf{z'^1_t})\big)$ as in eqs. S3.3, S3.5, or S3.6
        Sort the elements of $\phi_{t+1,k}$ from greatest to smallest, and store the corresponding states in variables $\mathbf{z^1_{t+1}}, \mathbf{z^2_{t+1}}, ..., \mathbf{z^{2^N}_{t+1}}$
        Set $C_{t+1,k} := \{\mathbf{z^1_{t+1}}\}$
Set $\mathcal{A}^* := \mathcal{A}_0 \cup \left[\bigcup_{k=1}^{m} \bigcup_{t=0}^{T-1} (E_{t,k} \cup C_{t+1,k})\right]$, and $m^* := |\mathcal{A}^*|$
Compute $m^* \times 1$ vectors $\tilde{\phi_{\mathbf{0}}}$ and $\tilde{\mathbf{q_t}}$, and $m^* \times m^*$ matrix $\tilde{\phi_{\mathbf{t}}}(\Theta)$, using states in $\mathcal{A}^*$
```

The first steps (lines 1–3 in Algorithm 2) find the m most likely initial states given the prior occupancy in the first year, $\phi_{\mathbf{0}}$. The *for loop* (line 4–11 in Algorithm 2) computes the most likely succession of occupancy states, starting from each of these m states, by computing the probability of each state after each extinction of colonization phase of years 1 to T. The next step (line 12) computes the union of all sets of states, $\mathcal{A}^*$, which is at most of size $\min(2mT, 2^N)$, if sets are all disjoint. This set $\mathcal{A}^*$ is the final set of states that are used for the computation of all quantities (line 13 in Algorithm 2). The size of this set, m^*, depends on the parameter m; if $m << 2^N$, then $m^* << 2^N$ and the approximate likelihood will be much faster to compute than the exact likelihood.

The approximate likelihood of the parameters given the prior probability of the initial state $\phi_{\mathbf{0}}$ and the survey in the first sampled year $\mathbf{Y_1}$ is thus:

$$\tilde{\mathcal{L}}(\Theta|\mathbf{Y_1}) = \tilde{\phi_{\mathbf{0}}} \left[\prod_{t=1}^{T-1} D(\mathbf{q_t}) \tilde{\phi_{\mathbf{t}}}(\Theta) \right] \tilde{\mathbf{q_T}}. \tag{S5.8}$$

The approximate likelihood function from eq. S5.8 was used in place of eqs. S3.3, S3.5, and S3.6 to compute the likelihood of the different hypotheses in the case of San Francisquito Creek, which has many segments. The accuracy of the likelihood estimation was assessed using parametric bootstrapping: 100 simulations were performed under the most likely model from Table Appendix S7 (hypothesis 1, *in situ* die-off), and Algorithm 2 was used to compare the models. Results are presented in Fig. Appendix S13(c); consistent with the results from Table Appendix S7, the null model was rejected in favor of hypothesis 1 ($\log_{10}(K_{0,1}) < -1$) in all simulations, while there was little support for hypothesis 2 ($|\log_{10}(K_{0,2})| < 0.5$) in any simulation.

Literature Cited

Fellers, G. M., and P. M. Kleeman, 2007 California red-legged frog (*Rana draytonii*) movement and habitat use: implications for conservation. Journal of Herpetology 41: 276–286.

MacKenzie, D. I., J. D. Nichols, J. E. Hines, M. G. Knutson, and A. B. Franklin, 2003 Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. Ecology 84: 2200–2207.

MacKenzie, D. I., J. D. Nichols, G. B. Lachman, S. Droege, J. A. Royle, *et al.*, 2002 Estimating site occupancy rates when detection probabilities are less than one. Ecology 83: 2248–2255.

Reichel, K., V. Bahier, C. Midoux, N. Parisey, J.-P. Masson, *et al.*, 2015 Interpretation and approximation tools for big, dense Markov chain transition matrices in population genetics. Algorithms for Molecular Biology 10: 31.

Supplementary tables

Table Appendix S6: **Parameter estimation.** Point estimates mentioned as $>x$ indicate that the posterior probability of the parameter plateaus for all values larger than x (see Figs. 4 and Appendix S16).

	Hypothesis	**Parameter**	**Estimate**	95% credible intervals
Shared parameters	-	α	175	[125,425]
		MATADERO CREEK		
	-	p	0.77	[0.69,0.82]
	-	e	0.12	[0.06,0.24]
	-	c	0.46	[0.22,0.96]
		DEER CREEK		
	-	p	0.75	[0.64,0.81]
	-	e	0.39	[0.21, 0.52]
	-	c	1.34	[0.75, 1.87]
		SAN FRANCISQUITO CREEK		
	-	p	0.69	[0.57,0.77]
	-	e	0.47	[0.33,0.62]
	-	c	0.81	[0.43,1.30]
Hypothesis parameters		MATADERO CREEK		
	H0: no change	-	-	-
	H1: *in situ* die-off	K_D	> 1.26	[1.05,100]
		t_D	1982	[1902,1982]
	H2: source population loss	K_L	> 0.13	[0.13,100]
		d_L	200	[200,3600]
		t_L	1902	[1902,1982]
		DEER CREEK		
	H0: no change	-	-	-
	H1: *in situ* die-off	K_D	> 1.45	[1.38, 100]
		t_D	1982	[1912, 1982]
	H2: source population loss	K_L	> 66.1	[0.18, 100]
		d_L	200	[200,3400]
		t_L	1982	[1902, 1982]
		SAN FRANCISQUITO CREEK		
	H0: no change	-	-	-
	H1: *in situ* die-off	K_D	100	[41.69,100]
		t_D	1982	[1977,1982]
	H2: source population loss	K_L	0.1	[0.1,79.43]
		d_L	4000	[400,4000]
		t_L	1902	[1902,1982]

Table Appendix S7: **Hypothesis tests.** In each creek, the hypothesis with substantial or strong evidence is highlighted in bold.

Numerator	**Denominator**	$\log_{10}$(**Bayes factor**)[a]
	MATADERO CREEK	
H0: no change	H1: *in situ* die-off	0.058
H0: no change	H2: source population loss	-0.011
H1: *in situ* die-off	H2: source population loss	-0.069
	DEER CREEK	
H0: no change	**H1: *in situ* die-off**	**-0.944***
H0: no change	H2: source population loss	-0.387
H1: *in situ* die-off	H2: source population loss	**0.558***
	SAN FRANCISQUITO CREEK	
H0: no change	**H1: *in situ* die-off**	**-44.616****
H0: no change	H2: source population loss	0.087
H1: *in situ* die-off	H2: source population loss	**44.680**

[a] Bayes factors are computed from eq. 3, using the likelihood from eq. 2 for Matadero and Deer Creeks and from eq. S5.8 for San Francisquito Creek.
* Substantial evidence
** Strong evidence

Table Appendix S8: **Model selection using the Akaike Information Criterion.** The AIC is computed from the maximum likelihood of the parameters $\mathcal{L}(\Theta_h, \mathbf{z_{1902}}|\mathbf{Y_{1997}})$ as AIC=$2k - 2\ln\mathcal{L}(\Theta_h, \mathbf{z_{1902}}|\mathbf{Y_{1997}})$, where k is the number of model parameters and h is the hypothesis. For Matadero and Deer Creeks, likelihoods for hypotheses H0, H1, and H2 are computed from eqs. S3.6, S3.3, and S3.5, respectively; for San Francisquito Creek, likelihoods for hypotheses H0, H1, and H2 are computed from eq. S5.8. The number of parameters is 0 for H0, 2 for H1, and 3 for H2. This analysis provides an alternative to the Bayes factor model selection presented in Table Appendix S7.

Hypothesis	**AIC**
MATADERO CREEK	
H0: no change	**11.86**
H1: *in situ* die-off	14.74
H2: source population loss	17.00
DEER CREEK	
H0: no change	18.31
H1: *in situ* die-off	**14.96**
H2: source population loss	17.07
SAN FRANCISQUITO CREEK	
H0: no change	342.90
H1: *in situ* die-off	**130.09**
H2: source population loss	348.90

Table Appendix S9: **Parameter estimation under the exact and approximate algorithms, as a function of the number of states from each year retained in the approximation, m.** See Fig. Appendix S18.

	Algorithm	Estimate	95%CI
		p	
MATADERO CREEK	**Exact**	**0.77**	**[0.69,0.82]**
	Approximate $m = 2$	0.78	[0.69,0.82]
	Approximate $m = 5$	0.77	[0.68,0.82]
	Approximate $m = 10$	0.77	[0.69,0.82]
	Approximate $m = 20$	0.77	[0.69,0.82]
		e	
	Exact	**0.12**	**[0.06,0.24]**
	Approximate $m = 2$	0.10	[0.04,0.17]
	Approximate $m = 5$	0.10	[0.04,0.19]
	Approximate $m = 10$	0.11	[0.05,0.20]
	Approximate $m = 20$	0.11	[0.05,0.22]
		c	
	Exact	**0.46**	**[0.22,0.96]**
	Approximate $m = 2$	0.36	[0.17,0.71]
	Approximate $m = 5$	0.38	[0.17,0.76]
	Approximate $m = 10$	0.40	[0.18,0.81]
	Approximate $m = 20$	0.42	[0.20,0.86]
		p	
DEER CREEK	**Exact**	**0.75**	**[0.64,0.81]**
	Approximate $m = 2$	0.79	[0.69,0.84]
	Approximate $m = 5$	0.77	[0.67,0.83]
	Approximate $m = 10$	0.76	[0.65,0.82]
	Approximate $m = 20$	0.75	[0.65,0.82]
		e	
	Exact	**0.39**	**[0.21,0.52]**
	Approximate $m = 2$	0.38	[0.21,0.53]
	Approximate $m = 5$	0.35	[0.19,0.50]
	Approximate $m = 10$	0.39	[0.19,0.53]
	Approximate $m = 20$	0.39	[0.21,0.52]
		c	
	Exact	**1.34**	**[0.75,1.87]**
	Approximate $m = 2$	1.40	[0.82,1.92]
	Approximate $m = 5$	1.19	[0.67,1.76]
	Approximate $m = 10$	1.34	[0.67,1.87]
	Approximate $m = 20$	1.34	[0.75,1.87]
		p	
SAN FRANCISQUITO CREEK	Approximate $m = 2$	0.67	[0.56,0.75]
	Approximate $m = 5$	0.69	[0.58,0.77]
	Approximate $m = 10$	0.69	[0.57,0.76]
	Approximate $m = 20$	0.69	[0.57,0.77]
		e	
	Approximate $m = 2$	0.45	[0.33,0.58]
	Approximate $m = 5$	0.51	[0.36,0.63]
	Approximate $m = 10$	0.50	[0.35,0.63]
	Approximate $m = 20$	0.47	[0.33,0.62]
		c	
	Approximate $m = 2$	0.95	[0.60,1.31]
	Approximate $m = 5$	0.99	[0.56,1.34]
	Approximate $m = 10$	0.92	[0.52,1.33]
	Approximate $m = 20$	0.81	[0.43,1.30]

Supplementary figures

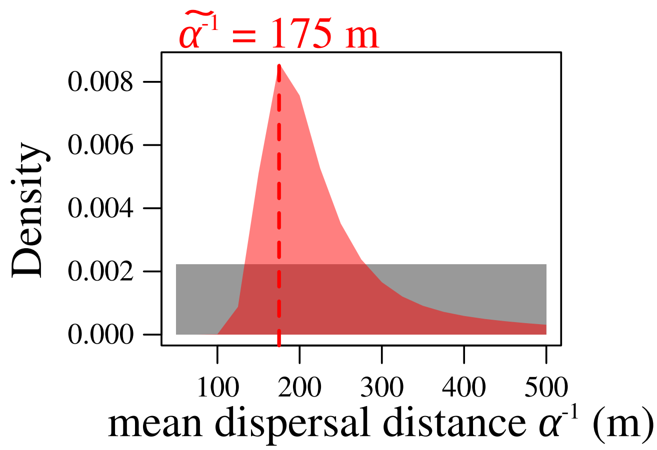

Figure Appendix S10: **Bayesian parameter estimation of the mean dispersal distance** (α^{-1}) **of the California red-legged frog.** The gray area represents the prior distribution, the red area represents the posterior distribution of the parameters given the observed data in the three creeks (computed from eq. S2.13). This posterior distribution was computed simultaneously with that of other parameters presented in Fig. 3.

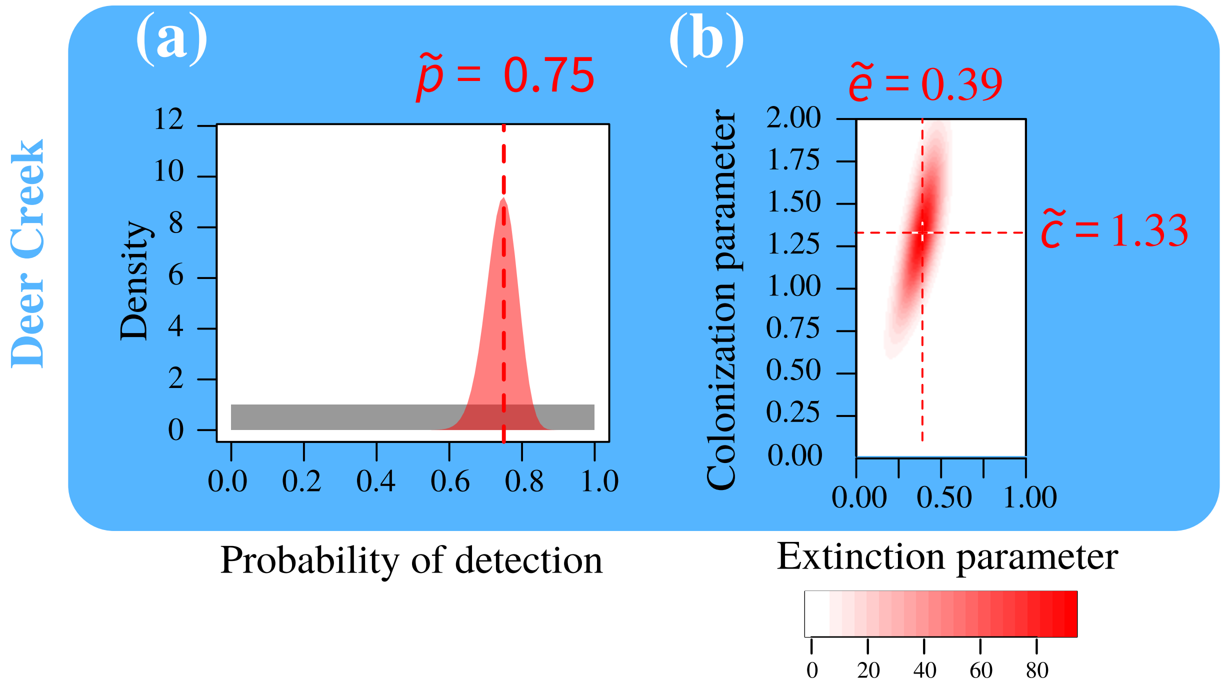

Figure Appendix S11: **Bayesian parameter estimation of the probability of detection** (p)**, the extinction parameter** (e)**, and the colonization parameter** (c) **of the California red-legged frog in Deer Creek, using an uninformative prior (eq. S2.9) for the missing occupancy in the initial year (1998).** Note that there is no missing data in the initial year in Matadero and San Francisquito Creeks, so this type of analysis is not needed for these creeks. The figure is analogous to Fig. 3(c)-(d).

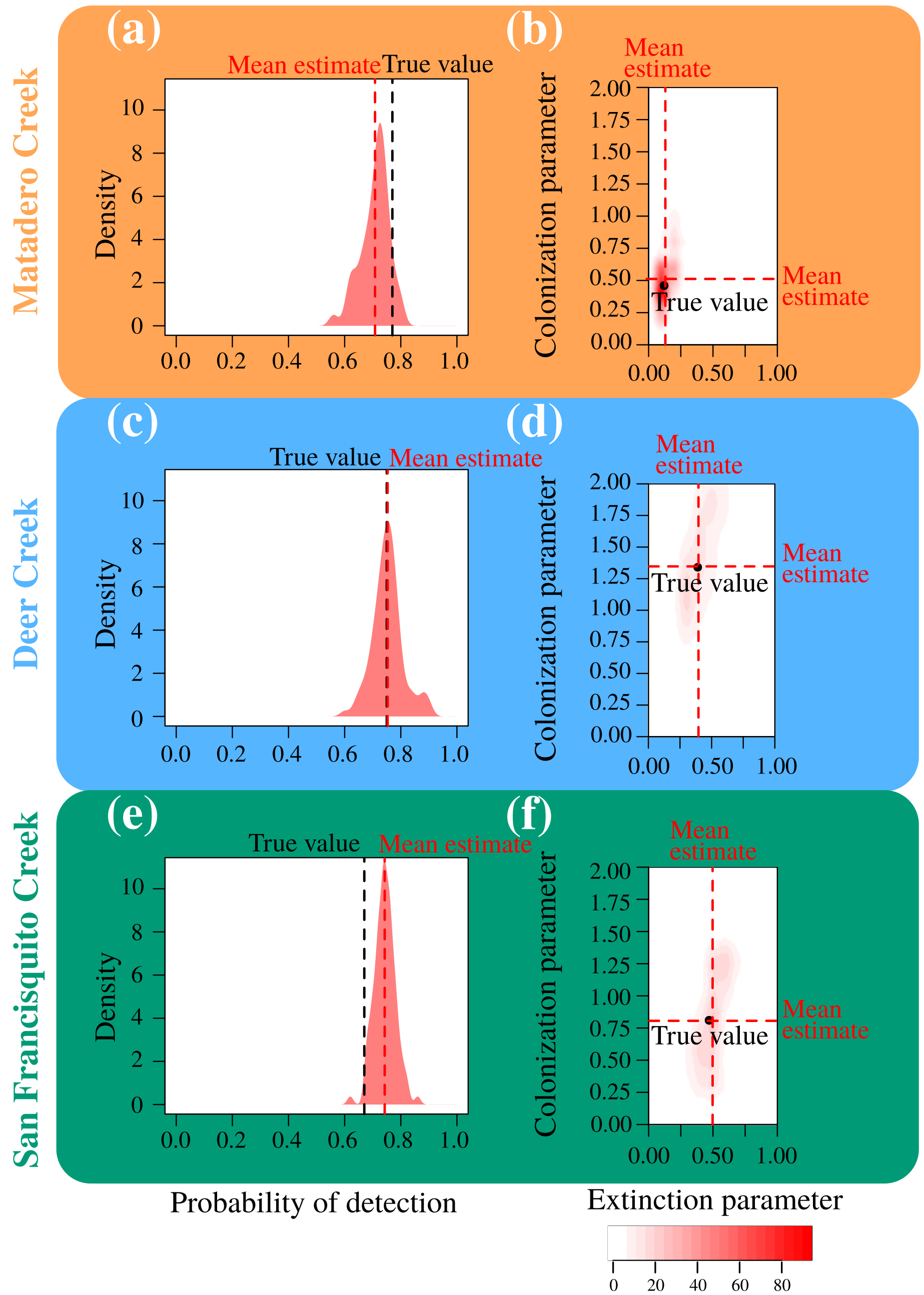

Figure Appendix S12: **Accuracy of the Bayesian parameter estimation of the extinction and colonization parameters from Fig. 3.** (a) Matadero Creek. (b) Deer Creek. (c) San Francisquito Creek. Shades of gray represent the density of point estimates $\tilde{e}$ and $\tilde{c}$, estimated using a Gaussian kernel density estimate with a bandwidth of 0.05. We performed 100 Monte Carlo simulations of patch occupancy data. Simulations started from the initial occupancy $\mathbf{Y_{1997}}$, and patch occupancies of the following years were successively drawn from the set of possible occupancies using probability transitions from eq. S2.2, with extinction and colonization parameters corresponding to the maximum *a posteriori* estimates of $\tilde{e}$ and $\tilde{c}$: 0.12 and 0.46 in Matadero Creek, 0.39 and 1.34 in Deer Creek, 0.47 and 0.81 in San Francisquito Creek (from Fig. 3). Colored dots represent the true values of e and c used for the simulations; dashed lines represent the mean $\tilde{e}$ and $\tilde{c}$ across the 100 simulations. Point estimates of e and c in (a)-(c) are similar to that in Fig. 3(c), (f), (i), and the distribution of point estimates from the simulations match the posterior distribution obtained from the actual dataset. The results support the accuracy of the credible intervals provided in Fig. 3.

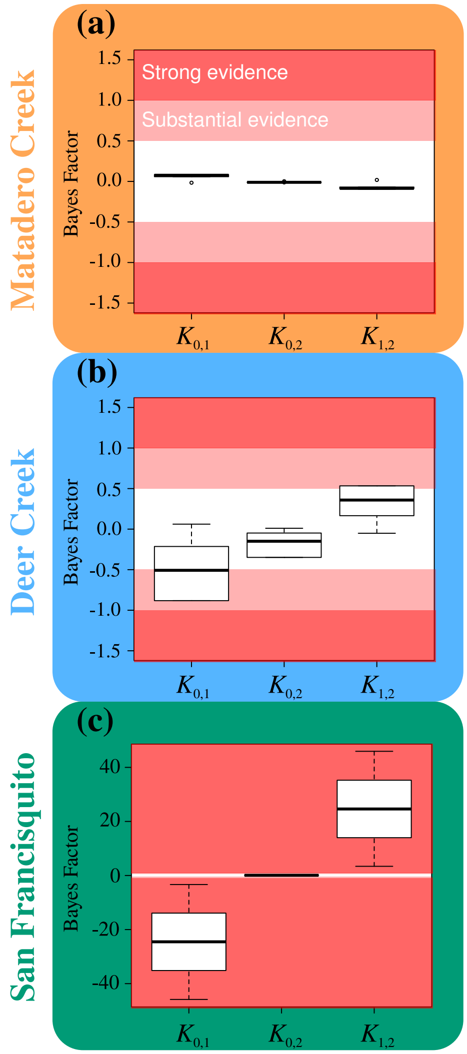

Figure Appendix S13: **Accuracy of the Bayesian model comparison from Table Appendix S7.** (a) Matadero Creek. (b) Deer Creek. (c) San Francisquito Creek. Boxplots represent the Bayes factors of hypotheses 0 and 1 ($K_{0,1}$), 0 and 2 ($K_{0,1}$), and 1 and 2 ($K_{0,1}$). For each creek, we performed 100 Monte Carlo simulations of patch occupancy data under the most likely hypothesis as determined by the Bayes factors between the three 3 models (Table Appendix S7). Simulations started from a random initial occupancy $\mathbf{z_{1882}}$, and patch occupancies of the following years were successively drawn from the set of possible occupancies using probability transitions from eq. S2.2, with extinction and colonization parameters corresponding to the maximum *a posteriori* estimates of $\tilde{e}$ and $\tilde{c}$: 0.12 and 0.46 in Matadero Creek, 0.39 and 1.34 in Deer Creek, 0.47 and 0.81 in San Francisquito Creek (from Fig. 3). For Matadero Creek, simulations were done under hypothesis 0 (null hypothesis), and the model likelihoods were computed from eqs. S3.3, S3.5, and S3.6. We assumed that the segment sizes K were constant through time and equal to 1 and that no source population was present during the following 115 years. For Deer Creek, simulations were done under hypothesis 1 (increased *in situ* die-off), and the model likelihoods were computed from eqs. S3.3, S3.5, and S3.6. For San Francisquito Creek, simulations were done under hypothesis 1 (increased *in situ* die-off), and the model likelihoods were computed from eq. S5.8. We assumed that the segment sizes K were equal to their maximum *a posteriori* estimate $\tilde{K_D} = 100$ before the event occurring at $\tilde{t}_D = 1982$, and equal to 1 after 1982.

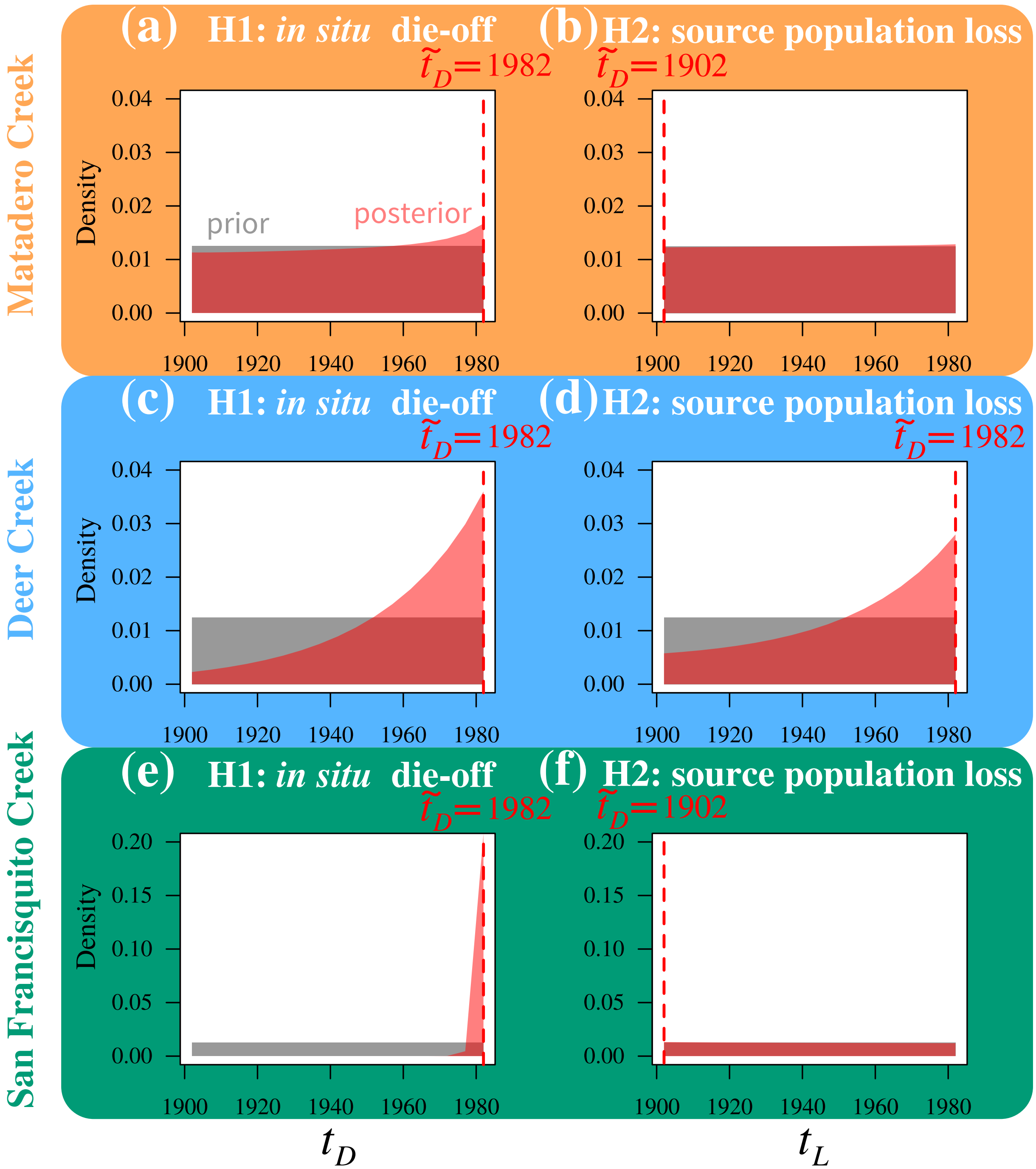

Figure Appendix S14: **Bayesian estimation of the timing of the population reduction event, under two hypotheses.** (a) *In situ* die-off hypothesis, Matadero Creek. (b) Habitat loss hypothesis, Matadero Creek. (c) *In situ* die-off hypothesis, Deer Creek. (d) Habitat loss hypothesis, Deer Creek. (e) *In situ* die-off hypothesis, San Francisquito Creek. (f) Habitat loss hypothesis, San Francisquito Creek. Under hypothesis 1, the model parameter t_D corresponds to the timing of the event increasing *in situ* die-off. Under hypothesis 2, the model parameter t_L corresponds to the timing of the loss of a source population. These posterior distributions were computed simultaneously with that of other parameters presented in Figs. 4 and Appendix S16.

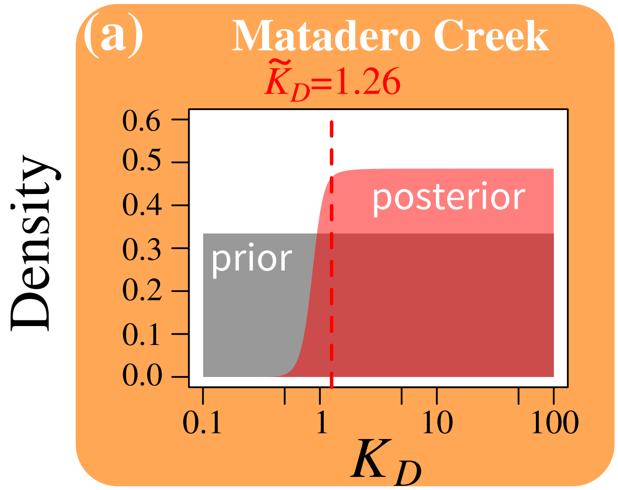

Figure Appendix S15: **Bayesian estimation of model parameters under the *in situ* die-off hypothesis (H1) in Matadero Creek.** Figure design matches that of Fig. 4.

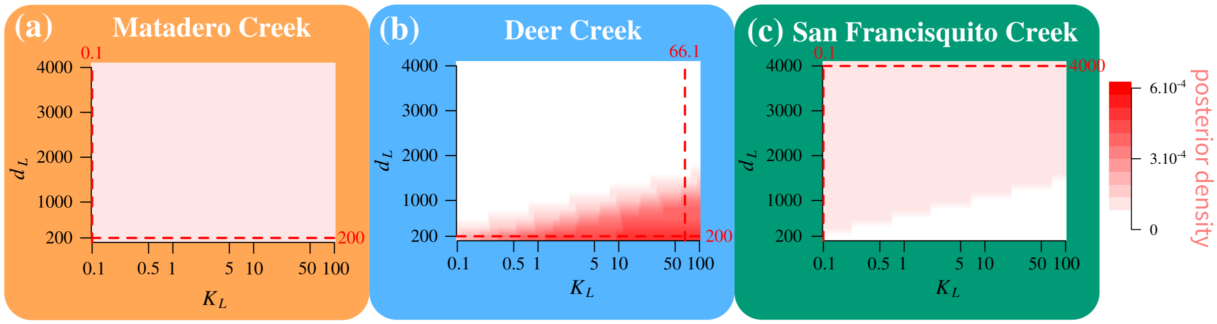

Figure Appendix S16: **Bayesian estimation of model parameters for the California red-legged frog, under the source population loss hypothesis instead of the *in situ* die-off hypothesis presented in Fig. 4.** (a) Matadero Creek. (b) Deer Creek. (c) San Francisquito Creek. The model parameters K_L and d_L correspond to the population size and the distance to the creek of a source population. The shades of red represent the joint posterior probability of K_L and d_L (see scale on the right). Other parameters of the SPOM appear in Table 1.

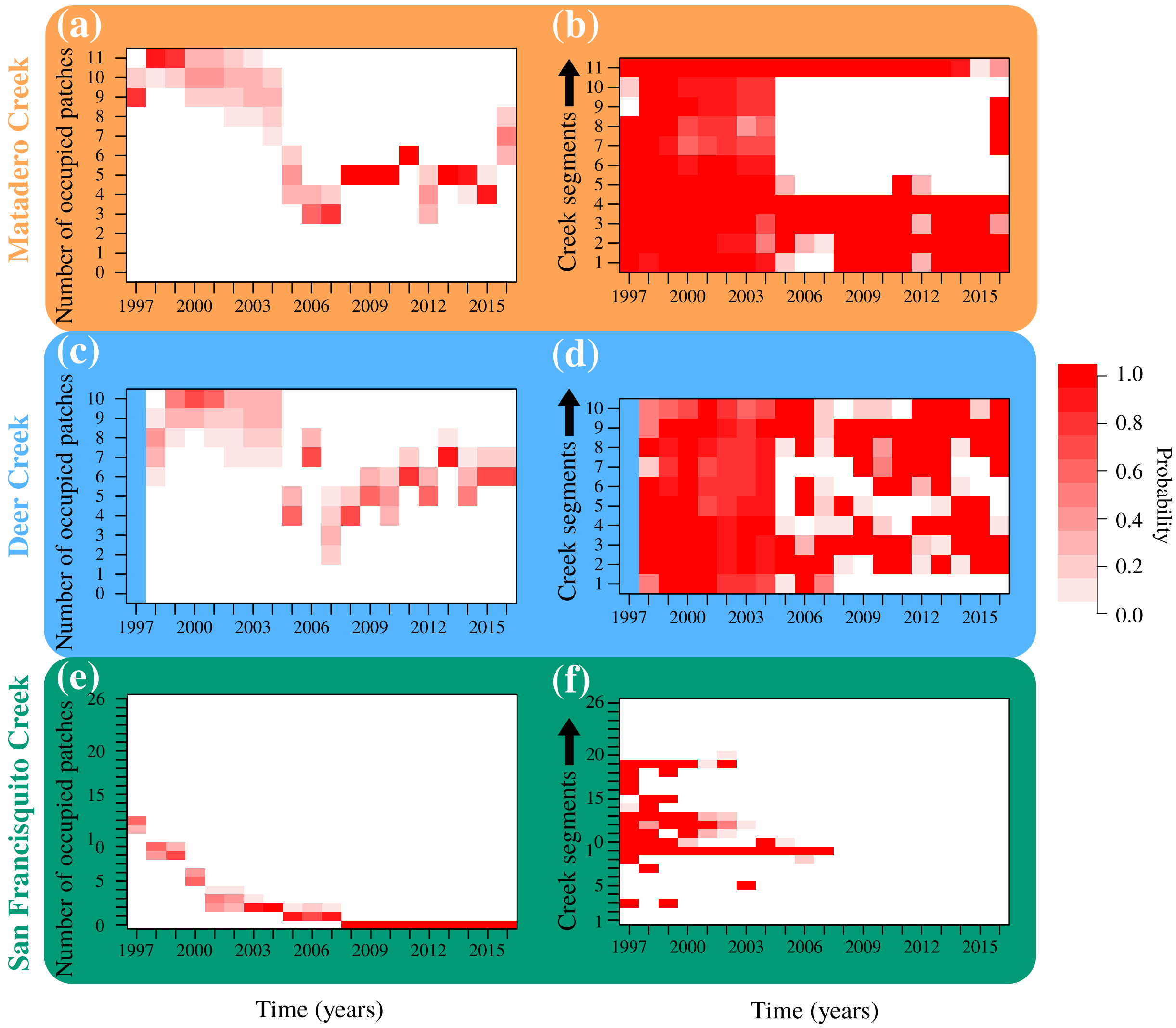

Figure Appendix S17: **Estimated trajectories of the occupancy and segment occupancy of three creeks.** (A) Number of occupied segments, Matadero Creek. (B) Segment occupancy, Matadero Creek. (C) Number of occupied segments, Matadero Creek. (D) Segment occupancy, Matadero Creek. (E) Number of occupied segments, Matadero Creek. (F) Segment occupancy, Matadero Creek. Shades of red represent probabilities (see legend). We assumed that extinction and colonization parameters correspond to their maximum *a posteriori* estimates (see Fig. 3).

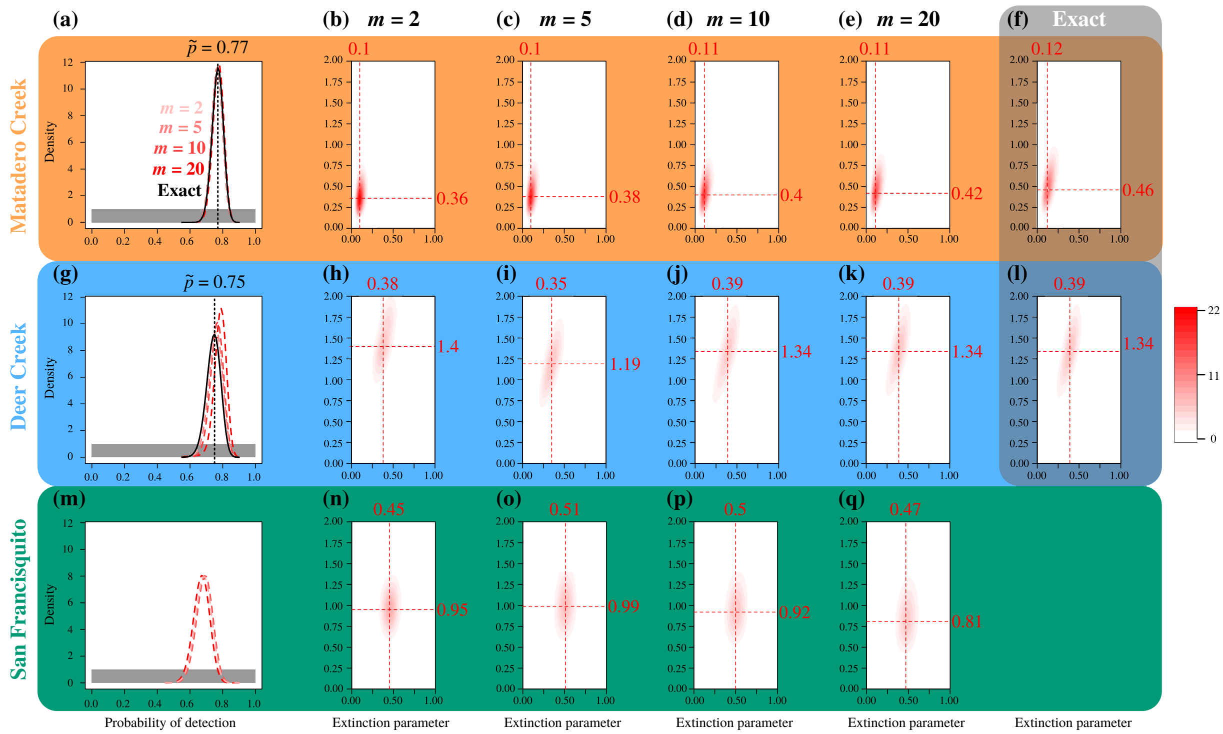

Figure Appendix S18: **Bayesian parameter estimation of the probability of detection (p), the extinction parameter (e), and the colonization parameter (c) of the California red-legged frog in the three creeks, using the many-patches approximate algorithm.** Panels (a), (g), and (m) are analogous to Fig. 3(a), (c), and (e). Panels (b)-(e), (h)-(k) , and (n)-(q) are analogous to Fig. 3(b), (d), (f). Panels (f) and (l) are copied from Fig. 3(b) and (d) for comparison. Note that the exact algorithm cannot be used in San Francisquito Creek due to the large number of states, so it is not reported here.