




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

Nicolas Alcalá ^{1*}, Alan E. Launer,² Michael F. Westphal,³ Richard Seymour,⁴ Esther M. Cole,² † and Noah A. Rosenberg¹ †

¹Department of Biology, Stanford University, 371 Serra Mall, Stanford, CA, 94305-5020, U.S.A.

²Land Use and Environmental Planning, Stanford University, 3160 Porter Drive, Suite 200, Palo Alto, CA, 94304-8442, U.S.A.

³US Bureau of Land Management, Hollister Field Office, 20 Hamilton Court, Hollister, CA, 95023, U.S.A.

⁴Stanford Conservation Program, 3160 Porter Drive, Suite 200, Palo Alto, CA, 94304-8442, U.S.A.

Abstract: Assessing causes of population decline is critically important to management of threatened species. Stochastic patch occupancy models (SPOMs) are popular tools for examining spatial and temporal dynamics of populations when presence-absence data in multiple habitat patches are available. We developed a Bayesian Markov chain method that extends existing SPOMs by focusing on past environmental changes that may have altered occupancy patterns prior to the beginning of data collection. Using occupancy data from 3 creeks, we applied the method to assess 2 hypothesized causes of population decline—in situ die-off and residual impact of past source population loss—in the California red-legged frog. Despite having no data for the 20–30 years between the hypothetical event leading to population decline and the first data collected, we were able to discriminate among hypotheses, finding evidence that in situ die-off increased in 2 of the creeks. Although the creeks had comparable numbers of occupied segments, owing to different extinction-colonization dynamics, our model predicted an 8-fold difference in persistence probabilities of their populations to 2030. Adding a source population led to a greater predicted persistence probability than did decreasing the in situ die-off, emphasizing that reversing the deleterious impacts of a disturbance may not be the most efficient management strategy. We expect our method will be useful for studying dynamics and evaluating management strategies of many species.

Keywords: amphibians, connectivity, metapopulations, statistics, streams

Uso de Modelos Estocásticos de Ocupación de Fragmentos de la Rana Californiana de Patas Rojas para la Inferencia Bayesiana con Respecto a Eventos Pasados y su Persistencia en el Futuro

Resumen: La evaluación de las causas de la declinación poblacional es de importancia crítica para el manejo de especies amenazadas. Los modelos estocásticos de ocupación de parches (SPOMs, en inglés) son herramientas populares para examinar las dinámicas espaciales y temporales de las poblaciones cuando están disponibles los datos de presencia-ausencia para múltiples parches de hábitat. Desarrollamos un método bayesiano de cadena de Markov que extiende a los SPOMs existentes al enfocarse en los cambios ambientales pasados que podrían haber alterado los patrones de ocupación previos al inicio de la recolección de datos. Con los datos de ocupación de tres arroyos, aplicamos este método para evaluar dos causas hipotéticas de la declinación poblacional - muerte in situ e impacto residual de causas anteriores de pérdida de una población fuente - de la rana californiana de patas rojas. A pesar de no tener datos para 20 - 30 años entre el evento hipotético que derivó en la declinación poblacional y los primeros datos recolectados, pudimos discriminar entre las hipótesis, encontrando evidencia de que la muerte in situ incrementó en dos de los arroyos. Aunque los arroyos tuvieron un número comparable de segmentos ocupados, debido a diferentes

*email nalcala@stanford.edu

†These authors contributed equally to this work.

Article impact statement: Patch occupancy data can be used to detect disturbances and to predict the impact of management strategies. Paper submitted March 23, 2017; revised manuscript accepted June 21, 2018.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

dinámicas de colonización - extinción, nuestro modelo predijo una diferencia ocho veces mayor en las probabilidades de persistencia de las poblaciones hasta el 2030. La suma de una población fuente resultó en una mayor probabilidad de persistencia pronosticada que con la reducción de la muerte in situ, enfatizando que la reversión de los impactos dañinos de una perturbación puede no ser la mejor estrategia de manejo. Esperamos que nuestro método sea útil para el estudio de las dinámicas y para la evaluación de las estrategias de manejo de muchas especies.

Palabras Clave: anfibios, arroyos, conectividad, estadística, metapoblaciones

摘要: 在濒危物种管理中, 评估种群数量下降的原因十分重要。当拥有多个斑块的物种的“出现”和“缺失”数据时, 常用随机斑块占有模型来研究种群的时空动态。我们开发了一种贝叶斯马尔可夫链的方法, 对现有的随机斑块占有模型进行了拓展, 它聚焦于在开展数据收集之前, 可能已经改变了种群占有格局的历史环境变化。利用三条溪流的加利福尼亚红腿蛙的占域数据, 我们用这个方法评估了该物种种群数量下降的两个可能原因——原地灭绝以及历史上源种群数量下降的残余影响。虽然可能导致种群数量下降的事件与首次数据收集之间的二三十年内都没有数据, 但我们还是能够区分不同的假设, 并且发现在三条溪流中有两条溪流原地灭绝的情况一度增长。由于不同的种群灭绝-定殖动态, 这三条溪流被占有的河段数量相当, 尽管如此, 我们的模型预测, 到 2030 年不同溪流种群续存的可能性相差八倍之多。而相比于减少原地灭绝, 增加源种群更有助于种群续存, 这一结果表明, 扭转干扰的不利影响可能并非最有效的管理策略。最后, 我们期望这一方法能应用于更多物种的种群动态研究及管理策略评估。【翻译: 胡怡思; 审校: 聂永刚】

关键词: 两栖类, 连接度, 集合种群, 统计, 溪流

Introduction

Occupancy models, which consider observations of presence or absence of a species across habitat patches, have been used to test ecological hypotheses on metapopulations (Hanski 1994), species invasions (Yaculic et al. 2012), disease dynamics (Adams et al. 2010), species distributions (Gormley et al. 2011), population trends (With & King 1999), abiotic relationships (Cole & North 2014), and community-level interactions (Welsh et al. 2006). Stochastic patch occupancy models (SPOMs) (Gyllenberg & Silvestrov 1994; Hanski 1994) are a family of occupancy models that describe transitions between occupancy states in terms of extinction and colonization. This class of models has advanced to accommodate imperfect detection, demographic dynamics, sparse data sets, and spatially explicit data. Some early SPOMs considered incidence-function methods, in which the stationary probability of occupancy was used to infer extinction and colonization rates (Hanski 1994; ter Braak et al. 1998), based on the assumption they are time independent. To address this oversimplification, several researchers developed Markov chain models that permit time-varying extinction and colonization probabilities dependent on values from other patches (O'Hara et al. 2002; ter Braak & Etienne 2003; Moilanen 2004) and that accommodate imperfect detection (MacKenzie et al. 2003; Johnson et al. 2013). A recent method with demographic dynamics is spatially explicit and allows imperfect detection (Sutherland et al. 2014).

Our study was motivated by the population decline of the U.S. federally threatened California red-legged frog (*Rana draytonii*) and a desire to recover an *R. draytonii* metapopulation on Stanford University lands. The species

has declined by 70% in its range (Hayes & Jennings 1988; Fisher & Shaffer 1996). On Stanford lands, *R. draytonii* was reported in 23 stream segments in 1997 and in 12 in 2012. Factors at multiple spatial scales may influence the populations: habitat loss (Davidson et al. 2002), predation by exotic species (Lawler et al. 1999), disease (Fisher et al. 2012), and climate change (Davidson et al. 2002).

We found existing SPOM methods imperfectly suited to our study system. Our data set had missing data in the sampling period 1997–2016, and did not include any data from before 1997, when events relevant to the decline likely took place. Many current spatially implicit methods use a Bayesian framework and allow years of missing data (Royle & Dorazio 2008; Fiske et al. 2011; Kéry & Schaub 2011; Bailey et al. 2014), but do not estimate parameters for those years. Although Markov methods can be used to infer parameters during missing years, they have not previously been designed to infer changes in parameter values during missing-data periods (O'Hara et al. 2002; ter Braak & Etienne 2003; Risk et al. 2011).

We developed a Bayesian Markov chain method for inference under a novel SPOM that permits substantial missing data, handles imperfect detection, and explicitly models periods predating sampling, allowing temporal parameter changes owing to disturbance prior to sampling. The method enables inference of extinction and colonization rates and detection probabilities, hypothesis testing, and prediction of extinction risk. It entails a likelihood of model parameters given occupancy data and incorporates prior parameter information to obtain posterior probabilities and credible intervals. The framework uses inferred parameter distributions for probabilistic imputation of missing occupancy data and to estimate future probabilities of extinction. For the period before

sampling, we compared 2 causes of decline in *R. draytonii*: increased in situ die-off (e.g., due to disease or introduced predators) and source population loss (e.g., through habitat loss). We evaluated possible management actions and their impact on extinction risk. This example illustrates the potential of the new method for a variety of problems.

Methods

Data

From 1997 to 2016, up to 2 visual encounter surveys per year were completed in prespecified segments of 3 creeks known to contain California red-legged frog (Fig. 1). The creeks are often partially dry during the summer dry season and are continuously wet during the winter wet season. The proportions of a creek that have pools, riffles, and runs vary over time and shift with flood events; the locations of these features also vary.

Surveys focused on presence and absence of adult frogs, the life stage most closely linked to long-term population persistence in *r*-selected species such as *R. draytonii* (Biek et al. 2002; Vonesh & De la Cruz 2002). Summer surveys, when water levels are low, maximize detection probability of adult frogs and surveyor safety.

We designed our model for the annual life cycle of *R. draytonii*. Breeding and dispersal happen during the wet season. Breeding occurs in permanent or seasonal ponds or stream pools, then tadpoles metamorphose in subsequent months. Dispersal of metamorphosed frogs occurs both upstream and downstream. Mean dispersal distance is 50–500 m (Lannoo 2005), and dispersal can occur across interstitial habitats (Bulger et al. 2003). We assumed the population was closed during summer (Supporting Information); therefore, surveys that did not detect *R. draytonii* were regarded as detection failures if occupancy was detected in the same segment in other surveys of the same year.

Stream segments were numbered from upstream to downstream and had a mean length of 202 m (SD 41). There were 11 segments in Matadero Creek, 10 in Deer Creek, and 26 in San Francisquito Creek (Supporting Information). Missing data rates were 17% in Matadero Creek, 22% in Deer Creek, and 13% in San Francisquito Creek (Supporting Information).

General SPOM

In our model (Fig. 2a), we divided a linear habitat into N patches. The distance between patches i and j is denoted by d_{ij} . In a discrete time period t , each patch is in 1 of 2 states, either occupied or unoccupied. Because our system has an annual cycle, we considered years rather than a generic time unit. We defined variables

as follows: p_j is per-patch detection probability of the species during survey j ; $J_{i,t}$ is number of surveys in patch i and year t ; $Y_{i,j,t}$ is detected occupancy (1, presence; 0, absence) in the j th survey of patch i in year t , and $\mathbf{Y}_t = (Y_{1,1,t}, \dots, Y_{N,1,t}, Y_{1,2,t}, \dots, Y_{N,J_{N,t},t})$ is the vector of observations of all patches in all surveys of year t .

In our model of sequential extinction-colonization dynamics, we assumed occupancy measurement precedes extinction. First, patches are surveyed, providing \mathbf{Y}_t . Following MacKenzie et al. (2003), survey results were assumed to depend on the detection probability p_j and the occupancy state $\mathbf{z}_t = (z_{1,t}, \dots, z_{N,t})$. Conditional on presence of the species ($z_{i,t} = 1$), the probability of observing presence in survey j of patch i in year t is p_j , and the probability of observing absence is $1 - p_j$. The probability of $Y_{i,j,t} = 1$ given the species is absent ($z_{i,t} = 0$) is 0, and the probability of $Y_{i,j,t} = 0$ given species absence is 1. We assumed independence of surveys, so \mathbf{Y}_t has probability equal to the product of the probabilities of the $Y_{i,j,t}$ across surveys j and patches i . We henceforth assumed p_j was constant across all j and equal to p for all surveys. The framework can be extended easily to allow the detection probability to vary spatially and temporally.

In the cycle corresponding to each discrete time step, after occupancy measurement, extinction occurs. The extinction phase represents the dry season in which some patches become empty. We denote by E_i the extinction probability of patch i (i.e., the probability the patch converts from occupied to unoccupied in the extinction phase of a discrete time unit) (Fig. 2a). Following Hanski and Ovaskainen (2000), we assumed extinction rates are inversely proportional to population size and colonized patches instantly reach carrying capacity (K_i) for patch i . Thus, $E_i = e/K_i$, where e is a global extinction parameter that is constant across patches; e but not E_i may exceed 1.

After the extinction phase of a time step, the colonization phase occurs, corresponding to the wet season in which patches are recolonized. We denote by $C_{i,t}$ the colonization probability of patch i in year t (Fig. 2a). Following Hanski and Ovaskainen (2000), we assumed colonization rates are proportional to the total number of migrants entering from occupied patches. We also assumed an exponential dispersal kernel. Thus, the colonization rate of patch i at time t is $C_{i,t} = c \sum_{j \neq i} \exp(-\alpha d_{ij}) K_j z'_{j,t}$, where c is the global colonization parameter (a constant across patches independent of i and t), α is the inverse of the mean dispersal distance of the species, and $z'_{j,t}$ is the occupancy of patch j after the extinction phase of year t . Although K_j values are unaffected by the extinction phase, $C_{i,t}$ is indirectly affected because extinction leads to $z'_{j,t} = 0$ for some segments, j , and thus decreases $C_{i,t}$. The c but not $C_{i,t}$ may exceed 1.

Model likelihood computation is in Supporting Information. In the likelihood computation, the E_i and $C_{i,t}$ are used to compute Φ_t , the $2^N \times 2^N$ transition matrix

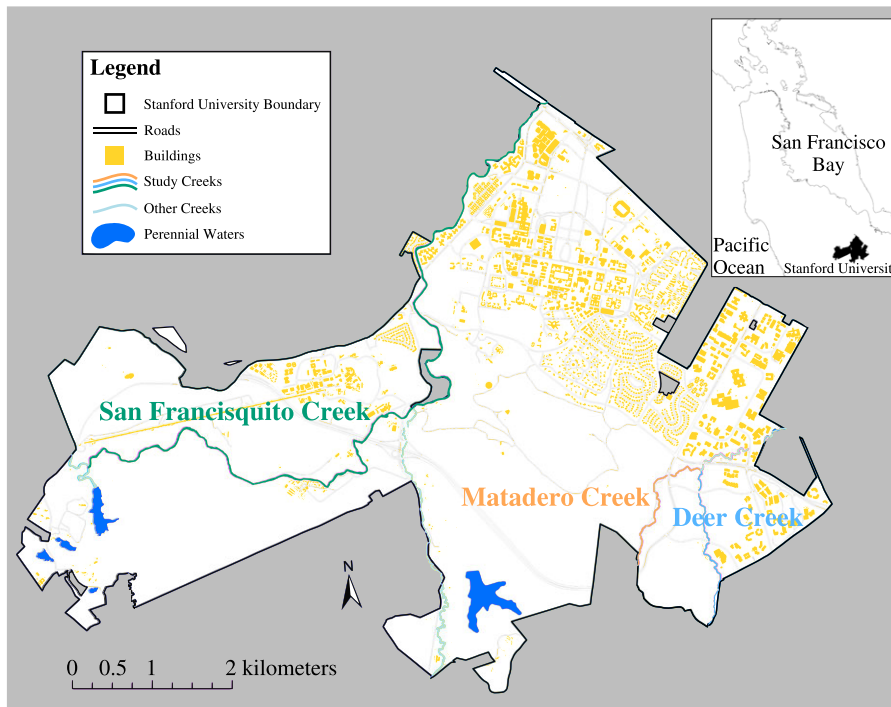


Figure 1. Locations of 3 creeks (Matadero, Deer, and San Francisquito) sampled for the California red-legged frog on Stanford University lands.

from all possible states in year t to states in year $t+1$, where entry ϕ_{tkl} represents the transition probability from state k to state l . The exact likelihood function of model parameters (Θ_0) given a series of detections measured from data $\mathbf{Y}_{1997}, \dots, \mathbf{Y}_{2016}$ is

$$\begin{aligned} \mathcal{L}(\Theta_0 | \mathbf{Y}_{1997}, \dots, \mathbf{Y}_{2016}) \\ = \Phi_{1997} \left[\prod_{t=1997}^{2015} D(\mathbf{q}_t) \Phi_t(\Theta_0) \right] \mathbf{q}_{2016}. \end{aligned} \quad (1)$$

In this equation, Φ_t denotes the 1×2^N probability vector of all 2^N possible states in year t ; \mathbf{q}_t denotes the $2^N \times 1$ column vector whose elements correspond to the probabilities of observing \mathbf{Y}_t given each state; and $D(\mathbf{q}_t)$ denotes the $2^N \times 2^N$ diagonal matrix for which diagonal entries are the elements of \mathbf{q}_t (Supporting Information). Thus, Eq. 1 matches Eq. 5 of MacKenzie et al. (2003), but our computation of Φ_t is related to that of O'Hara et al. (2002) and ter Braak & Etienne (2003) instead of that from MacKenzie et al. (2003).

To reduce computation time when the number of segments N is large, we built an approximate likelihood function, $\tilde{\mathcal{L}}(\Theta_0 | \mathbf{Y}_1, \dots, \mathbf{Y}_T)$ (Supporting Information). The approximate likelihood function considers only the most likely occupancy states.

Hypothetical Causes of Population Decline

We built models corresponding to 2 hypothetical causes of population decline for *R. draytonii*: in situ die-off and loss of a source population (Fig. 2b-c). For both models,

we assumed an event changed the model parameters in the past, before the first survey that produced occupancy data. We assumed parameters were constant for a long time before the event of interest; thus, the initial occupancy state was assumed to have limited impact on occupancy at the time of the event. We also assumed parameters were constant after the event.

Under hypothesis 1 (in situ die-off), population declines result from sudden mortality increases in all patches at time t_D , prior to the first sampling time. In situ die-off could result from disease, introduced predators, or reductions in habitat quality or availability. We assumed population sizes were equal in all patches. The per-patch population size before in situ die-offs is labeled K_D and the subsequent size is denoted K . Because extinction and colonization are treated as functions of population size, we assumed in situ die offs increase population extinction rates and decrease colonization rates. See Supporting Information for derivation of the likelihood under hypothesis 1 from the general likelihood (Eq. 1).

Under hypothesis 2 (source population loss), habitat destruction or local extinction leads to loss of a neighboring source population (e.g., a pond not subject to seasonal disappearance) at time t_L , before the first sampling time. Stanford lands are bordered on 3 sides by urban developments that have increased in density and spatial scope over the last 50 years. We hypothesized that a source population of size K_L was formerly near the current habitat, at distance d_L , and that it became extinct at time t_L . We assumed this source population was simply present and not subject

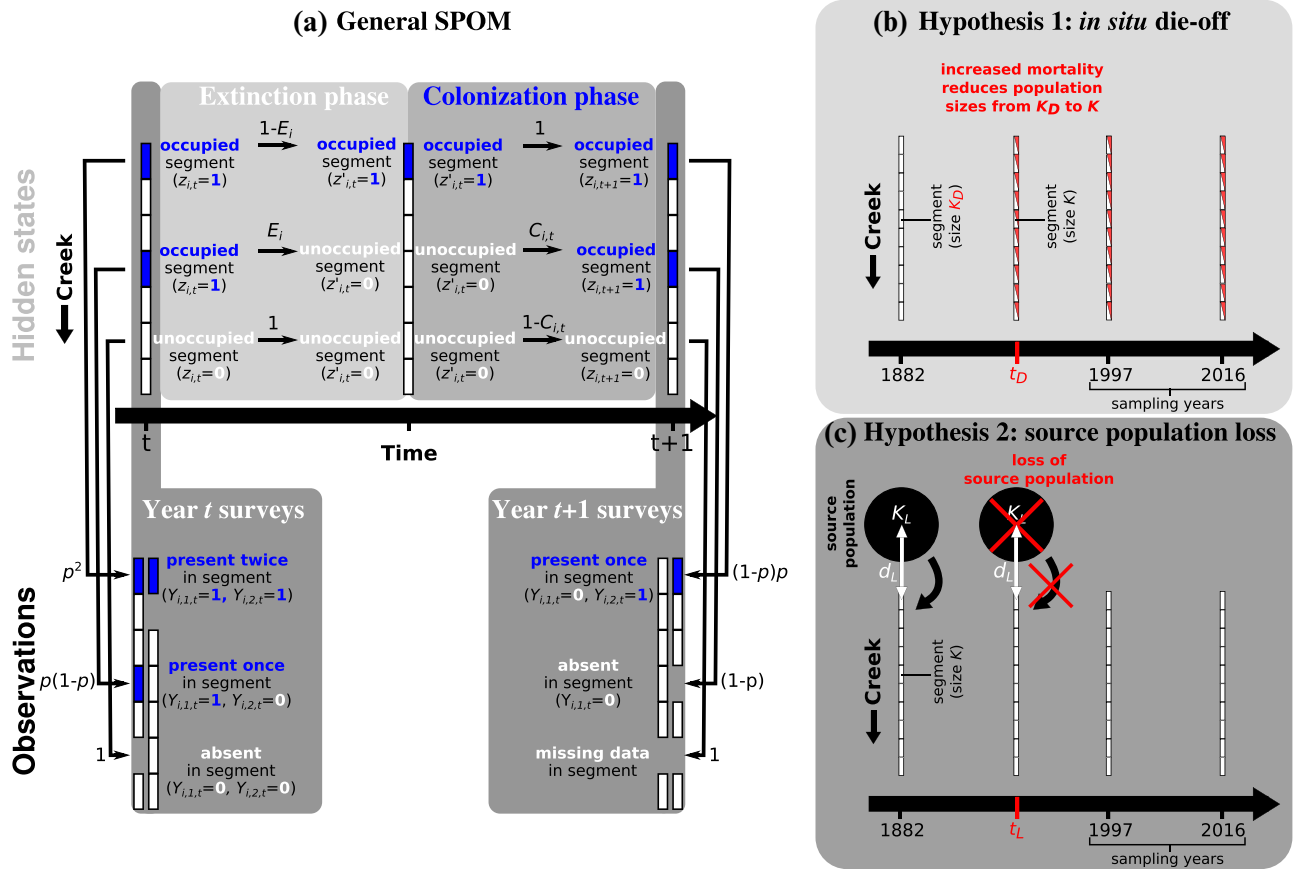


Figure 2. (a) General stochastic patch occupancy model (SPOM) (t , year; $z'_{i,t}$ and $z_{i,t+1}$, hidden random variables describing occupancy states at the end of extinction and colonization phases, respectively; $Y_{i,j,b}$, observed occupancy in survey j of patch i in year t ; p , detection probability; quantities on arrows, probabilities). (b) Hypothesis 1: In year t_D , in situ die-off (e.g., due to infection) reduces the population size of each patch from K_D to K . (c) Hypothesis 2: In year t_L , a source population of size K_L at distance d_L from the creek is lost (e.g., due to anthropogenic disturbance). In (b) and (c), year considered for the initial condition is 1882. In (c), curved arrows indicate direction of flow.

to extinction-recolonization dynamics before the loss (e.g., a pond). We assumed equal population size for all patches other than the source population. The difference in hypothesis 2 compared with hypothesis 1 is that the past event is localized rather than occurring as a global event affecting all patches similarly. Under hypothesis 2, global parameters e and c do not change following source population loss. Within-patch population sizes or extinction rates also do not change; these depend only on e and the population size. However, change occurs for the patch colonization rates—which depend on dispersal from the source population. The source-population-loss hypothesis encodes reduced colonization by preventing population rescue from the source population.

We computed the likelihood of the 2 hypotheses (see Supporting Information). With Θ_b denoting the parameters under hypothesis b , under hypothesis 1, $\Theta_b = (K_D, t_D)$, and, under hypothesis 2, $\Theta_b = (K_L, d_L, t_L)$. With t_0 denoting the initial year, the probability

vector of all possible states in t_0 is Φ_{t_0} . Because transition probabilities after the past event of interest do not depend on Θ_b , we denote the event date t_D or t_L by t_e . Then

$$\mathcal{L}(\Theta_b, \mathbf{z}_{t_0} | \mathbf{Y}_{1997}) = \Phi_{t_0} \left[\prod_{t=t_0}^{t_e-1} D(\mathbf{q}_t) \Phi_t(\Theta_0, \Theta_b) \right] \times \left[\prod_{t=t_e}^{1996} D(\mathbf{q}_t) \Phi_t(\Theta_0) \right] \mathbf{q}_{1997}. \quad (2)$$

Special cases of Eq. 2 corresponding to hypotheses 1 and 2 are in Supporting Information. The null hypothesis is a special case of hypothesis 1, with $K_D = 1$.

Bayesian Parameter Estimation

Estimation proceeded in 2 steps. First, using the whole occupancy data set, we inferred the detection probability p and the unknown parameters shared in

both hypotheses: the mean dispersal distance $1/\alpha$ and the global extinction and colonization parameters, e and c . The Bayesian estimation yielded posterior distributions of the parameters given the data set by multiplying the model likelihood (Eq. 1) by the prior distributions of model parameters (Supporting Information). We then used the mode of the joint posterior distribution to obtain maximum a posteriori estimates of α^{-1} , \tilde{z} , and \tilde{c} in each creek. We identified 95% credible intervals (CI) as the 2.5% and 97.5% quantiles of marginal posterior distributions. We used the posterior to perform missing data imputation (details in Supporting Information).

Second, we inferred the unknown parameters distinct to the 2 hypotheses: population size before infection K_D and infection timing t_D (hypothesis 1) and size of the lost habitat K_L , its distance to the creek d_L , and loss timing t_L (hypothesis 2). We multiplied the prior distribution of the parameters by the likelihood under the hypothesis (Eq. 2) to obtain the posterior distribution of Θ_b (Supporting Information). Maximum a posteriori parameter estimates and CI were computed as in the analysis used for both hypotheses. Our method provided a posterior distribution for the initial occupancy \mathbf{z}_{1882} ; however, this value is not of interest, and we integrated the joint posterior for the other parameters over all possible values of \mathbf{z}_{1882} .

The numerical implementation of the inference method and an approximate method for the case of a large number of patches, which we used for San Francisquito Creek, are in Supporting Information. We implemented the method in software MIDASPOM (Supporting Information).

Parametrization

Model parameters and prior distributions are in Table 1 (details in Supporting Information). The mean dispersal distance $1/\alpha$ was unknown and assumed equal for all creeks. The detection probability p and the global extinction and colonization parameters e and c were also unknown and differed by creek. The size of the populations under hypothesis 1 (K_D), source population size under hypothesis 2 (K_L), and timing of the hypothetical increase of in situ die-off (t_D) and of the hypothetical source population loss (t_L) were also unknown, and we assumed events occurred between 1902 and 1982 (Padgett-Flohr & Hopkins 2009; Fofonoff et al. 2017). Occupancies before t_D and t_L were unknown, and we assumed the species was present in 1882 (Hayes & Jennings 1988; Fellers 2005), so likelihood computation incorporated at least 20 years before t_D and t_L . The position of the hypothetical source population under hypothesis 2 was unknown. We assumed its distance d_L to the first creek segment was between 200 m and 4 km. Elements of the vector Φ_{1997} represent possible states \mathbf{z}_{1997} , and their values provide prior probabilities of each state in 1997; we considered a Bernoulli prior.

We assumed all other parameters were known. Patches corresponded to creek segments, and distances d_{ij} between segment pairs were assumed known from geographic data. Because the habitat is linear and distances between midpoints of consecutive segments were similar (Supporting Information), we assumed consecutive segments had fixed distance d . The relative carrying capacities K_i were treated as known from geographic data, and we assumed proportionality between patch area and population size (Hanski & Ovaskainen 2000). As a first approximation, we assumed $K_1 = K_2 = \dots = K_N = 1$ (Supporting Information).

Hypothesis Testing

To quantify fit of hypotheses to data, we performed Bayesian model selection. For hypotheses i and j , we computed Bayes factors by numerically integrating under each hypothesis the product of the model likelihood and the prior probabilities of parameters, considering all possible parameter values:

$$B_{i,j} = \frac{\Pr(\mathbf{Y}_{1997}|H_i)}{\Pr(\mathbf{Y}_{1997}|H_j)} = \frac{\int_{\Theta_i} \sum_{\mathbf{z}_{t_0}} \mathcal{L}(\Theta_i, \mathbf{z}_{t_0}|\mathbf{Y}_{1997}) \Pr(\Theta_i) \Pr(\mathbf{z}_{t_0}) d\Theta_i}{\int_{\Theta_j} \sum_{\mathbf{z}_{t_0}} \mathcal{L}(\Theta_j, \mathbf{z}_{t_0}|\mathbf{Y}_{1997}) \Pr(\Theta_j) \Pr(\mathbf{z}_{t_0}) d\Theta_j}, \quad (3)$$

where i and j equal 0 under the null hypothesis, 1 under hypothesis 1, and 2 under hypothesis 2, the likelihood was computed from Eq. 2, and the prior probabilities were computed from equations in Supporting Information. We interpreted B as follows (Jeffreys 1998): $|\log_{10}(B)| < 0.5$, little support for either hypothesis; $0.5 < |\log_{10}(B)| < 1$, substantial evidence for one hypothesis; and $1 < |\log_{10}(B)|$, strong evidence.

Extinction Probability Under Management Scenarios

We predicted future extinction probabilities in each creek under each of 4 management scenarios: no management, mortality reduction, habitat creation, and a combined scenario with both mortality reduction and habitat creation. In the mortality reduction scenario, local causes of mortality were controlled within every stream segment (e.g., non-native predator removal) to increase carrying capacity K . This scenario is equivalent to reversing the deleterious effects of the population decline that occur under hypothesis 1. In the habitat creation scenario, a population of size K_S was restored at distance d_S from the creek. This scenario is equivalent to reversing the deleterious effects of the decline that occur with source population loss in hypothesis 2. The combined strategy both reduced local causes of mortality and created new habitat.

Table 1. Summary of parameters of the stochastic patch occupancy model.

Hypothesis and variable	Interpretation	Range ^a	Value used ^b	Reference
Parameters shared across hypotheses ^b				
P	probability of detection	[0,1]	uniform prior	none
E	global extinction parameter	[0,1]	uniform prior	none
C	global colonization parameter	[0,1.5]	uniform prior	none
D	distance between consecutive segments (m)	[164,256] in MC ^c ; [147,249] in DC; [152,446] in SFC	200 m	Supporting Information
$1/\alpha$	mean annual dispersal distance	[50,500]	uniform prior	Bulger et al. 2003; Lannoo 2005; Fellers & Kleeman 2007
$z_{1997,i}$	occupancy state in patch i in 1997	[0,1]	Bernoulli prior	none
Hypothesis 1 In situ die-off				
t_D	timing of infection or increased predation	[1902,1982]	uniform prior	Padgett-Flohr & Hopkins 2009; Fofonoff et al. 2017
K_D	per-patch population size prior to infection	[0.1,100]	log-uniform prior	none
Hypothesis 2 Source population loss				
t_L	timing of source population loss	[1902,1982]	uniform prior	Google Historical Imagery ^d
K_L	population size of extinct population	[0.1,100]	log-uniform prior	none
d_L	distance to source population (m)	[200,4000]	uniform prior	none

^aRanges correspond to the maximal and minimal values of the prior distributions detailed in Supporting Information.

^bValues of parameters with priors are inferred using our Bayesian method, and distributions are detailed in Supporting Information.

^cAbbreviations: MC, Matadero Creek; DC, Deer Creek; SFC, San Francisquito Creek.

^dUsed to identify years when rapid expansion of human habitation in San Mateo and Santa Clara Counties around MC, DC, and SFC occurred.

To implement these scenarios, we performed Monte Carlo simulations from the current state \mathbf{p}_{2016} with patch extinction (E_i) and colonization (C_i) probabilities computed from e and c values sampled from their joint posterior distribution (Supporting Information). For each creek, we recorded the proportion of simulations with all segments unoccupied (total extinction of the population within a creek) from 2016 to 2065. This quantity is the probability that extinction has occurred by a given year when applying the transition probabilities from the observed state in 2016 and assuming that occupancy data are missing in all future years (Supporting Information).

To account for the difficulty of implementing management plans, for each management scenario, we chose parameter values that corresponded to reversing some but not all effects of disturbance events (hypotheses 1 and 2; parameter estimates in Supporting Information). For the mortality-reduction scenario above, this meant considering an increased local population size smaller than the estimated size before the event ($\tilde{K}_D \geq 1.45$ for hypothesis 1, from Fig. 4). We thus tested the impact of a small increase ($K = 1$ to 1.05) and a larger increase in population size ($K = 1$ to 1.25). For the habitat-creation scenario, this meant considering an increased source population size smaller than the estimated size before the event ($\tilde{K}_L = 66.1$ for hypothesis 2, from Supporting Information). We thus tested the impact of the addition of a small population ($K = 1$) either near (200 m) or farther from (600 m) the creek.

Results

Estimation of Shared Model Parameters

We estimated the value of the mean dispersal distance at $\alpha^{-1} = 175$ m (95% CI, 125–425) (Supporting Information). In addition, we estimated similar probabilities of detection in Matadero and Deer Creeks (Fig. 3a,c): $\tilde{p} = 0.77$ in Matadero (95% CI, 0.69–0.82) and 0.75 in Deer (95% CI, 0.64–0.81). Values were lower in San Francisquito (Fig. 3e), $\tilde{p} = 0.69$ (95% CI, 0.57–0.77).

The extinction and colonization parameter estimates differed among creeks (Fig. 3b,d,f). In Matadero Creek, point estimates of both parameters were small, $\tilde{e} = 0.12$ and $\tilde{c} = 0.46$. The marginal posterior distributions were narrow, indicating the data set was informative about extinction and colonization dynamics (95% CI, 0.06–0.24 for e and 0.22–0.96 for c).

By contrast, point estimates in Deer Creek were large: $\tilde{e} = 0.39$ and $\tilde{c} = 1.34$. The marginal posterior distributions of e (95% CI, 0.21–0.52) and c were also wider (95% CI, 0.75–1.87). The choice of either an informative or an uninformative prior for the missing data in the initial year (1998) had little effect on point estimates and credible intervals (Supporting Information).

In San Francisquito Creek, the extinction point estimate was the largest among all creeks, $\tilde{e} = 0.47$, and the colonization estimate was between those of Matadero and Deer Creeks, $\tilde{c} = 0.81$. The marginal posterior distributions of e and c were both moderately

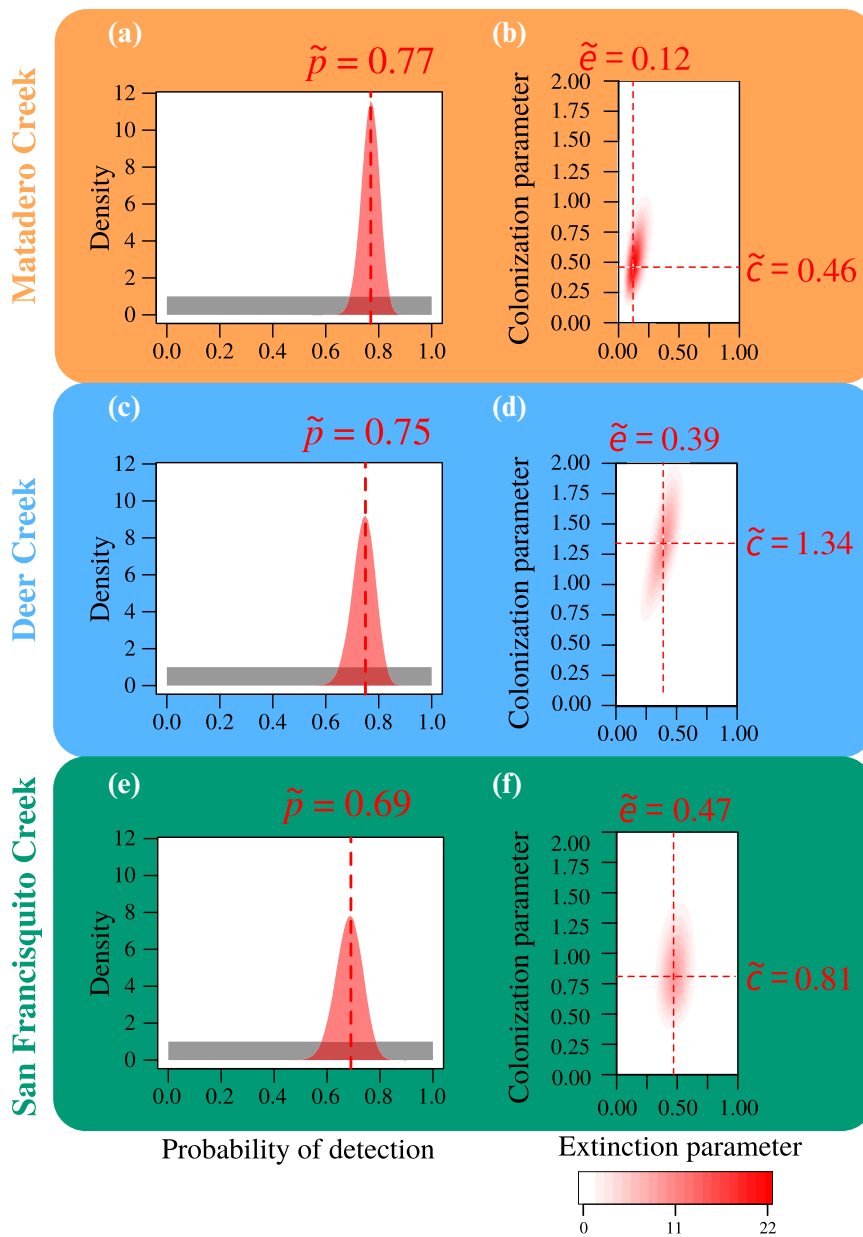


Figure 3. In Matadero, Deer, and San Francisco Creeks, Bayesian parameter estimation of (a, c, e) detection probability parameter and (b, d, f) joint distribution of extinction and colonization parameters. In (a, c, e), gray is prior distribution (Supporting Information), red is posterior distribution given observed data (from Eq. 1 for Matadero and Deer Creeks and from Supporting Information for San Francisco Creek), and red dashed line is maximum a posteriori estimate. In (b, d, f), red is joint posterior distribution (posterior density scale 0–22) and red dashed lines are maximum a posteriori estimates (\tilde{e} , \tilde{c} , respectively, extinction and colonization). An informative prior (Bernoulli, Supporting Information) was used for the missing data in the initial state of Deer Creek. Supporting Information contains the posterior distribution under an uninformative prior.

large (95% CI, 0.33–0.62 for e , 0.43–1.30 for c). Supporting Information provides an assessment of the accuracy of the estimation in Fig. 3.

Hypothesis Testing

The Bayes factor strongly supported the in situ die-off hypothesis H_1 in Deer Creek ($\log_{10} B_{0,1} = -0.944$, $\log_{10} B_{0,2} = -0.387$, $\log_{10} B_{1,2} = 0.558$) and San Francisco Creek ($\log_{10} B_{0,1} = -44.616$, $\log_{10} B_{0,2} = 0.087$, $\log_{10} B_{1,2} = 44.680$) (Supporting Information). In Matadero Creek, it did not reject the null hypothesis H_1 ($\log_{10} B_{0,1} = 0.058$, $\log_{10} B_{0,2} = -0.011$, $\log_{10} B_{1,2} = -0.069$). Assessment of the accuracy of the hypothesis testing is in Supporting Information.

Parameter Estimation

For Deer Creek, the posterior probability of K_D plateaued above a threshold \tilde{K}_D of 1.45. For San Francisco Creek, it was highest for $K \geq \tilde{K}_D = 95.5$ (Fig. 4). Current declines in Deer and San Francisco Creeks, respectively, could therefore be explained by decreases of at least 31% and 98% of the population size in each segment due to increased in situ die-off. The posterior distribution of the event timing was relatively flat for Deer Creek, indicating the data contained little information about this quantity (Supporting Information), and was largest in 1982 for San Francisco Creek (Supporting Information).

For Matadero Creek, because H_0 was not rejected, parameter estimation under H_1 was not relevant

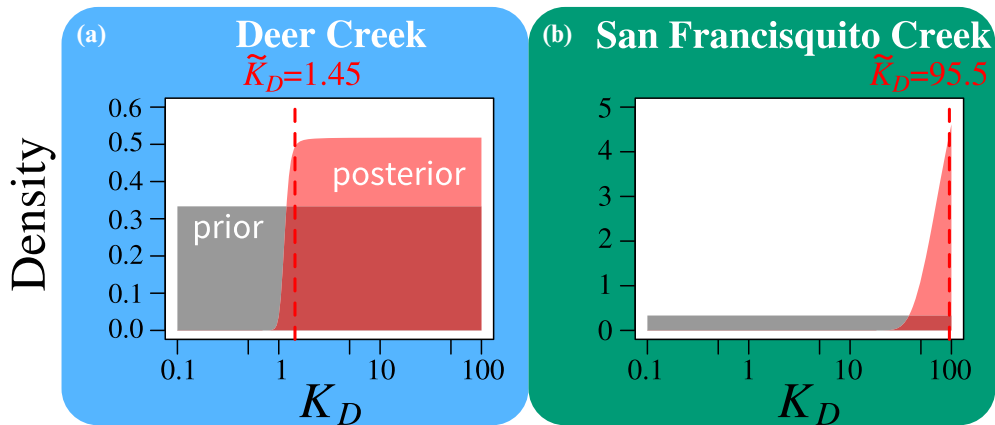


Figure 4. Bayesian estimation of model parameters under the hypothesis of in situ die-off of California red-legged frogs (H_1) in (a) Deer Creek and (b) San Francisquito Creek (K_D , population size before the event that increases in situ die-off; gray, prior distribution of K_D [Supporting Information]; red, posterior conditional on the observed data [Supporting Information]; red dashed lines, mode \tilde{K}_D of the posterior distribution used as a point estimate under hypothesis H_1). Model parameters are defined in Table 1. Only parameter estimates under the hypotheses with the greatest support (based on the Bayes factor, see Supporting Information) are shown. Parameter estimates under the (less strongly supported) source population loss hypothesis (H_2) and under hypothesis H_1 for Matadero Creek are in Supporting Information.

(Supporting Information). Parameter estimation under H_2 in the creeks is in Supporting Information.

Extinction Probability Under Management Scenarios

The Deer Creek population was more likely to go extinct than the Matadero Creek population (8-fold higher extinction probability by 2030, black line in Fig. 5a–b). The San Francisquito Creek population was probably already extirpated by 2008 (Supporting Information).

All management interventions decreased the extinction probability in Deer and Matadero Creeks. In San Francisquito Creek, only source-population creation restored the population and lowered the future extinction probability below 1. In the 3 creeks, source-population creation 200 m from the creek generated the lowest extinction probability.

Discussion

We developed a method to use temporal patch occupancy data to infer extinction and colonization rates and to test hypotheses under a Bayesian framework and applied the method to 3 populations of *R. draytonii*. With the method, we were able to test for changes in parameter values that occurred before the first occupancy survey. It also enabled tests of hypotheses of past disturbance and estimates of future population trajectories.

A key innovation is that our method detects disturbances that predate data collection. Many spatially implicit SPOM methods (Royle & Dorazio 2008; Bailey et al. 2014) use a Bayesian framework, allow arbitrary

missing data, and detect temporal parameter changes; however, they neither detect changes before the first survey nor test explicit hypotheses concerning such changes. Although the spatially explicit SPOMSIM (Moilanen 2004) and O'Hara et al.'s (2002) method permit years of missing data, they do not allow temporal parameter changes or missing data years before the first survey. The spatially explicit Bayesian method from ter Braak and Etienne (2003) allows years of missing data prior to data collection by setting a prior for the initial occupancy state, but it does not implement parameter changes prior to data collection. We extended their framework by using the whole data set to estimate current parameter values and the first year of data to estimate past values.

Past Influences on California Red-Legged Frogs

Our method identified local variation in *R. draytonii* populations. Despite the proximity of 3 creeks, we found differences in extinction and colonization (Fig. 3). Local variation in parameter estimates might reflect differences in community structure, habitat quality, or disease dynamics. For example, non-native predators are most abundant in San Francisquito Creek (A.E.L. and E.M.C. data).

We estimated the mean dispersal distance of the species at 175 m (Supporting Information). This estimate accords with movement studies in Point Reyes, California, which report a median of 185 m (Fellers & Kleeman 2007). However, this value may not be appropriate here because Point Reyes frogs breed in freshwater ponds rather than creeks.

We compared 2 past population-decline scenarios: in situ die-off and source-population loss. An in situ die-off

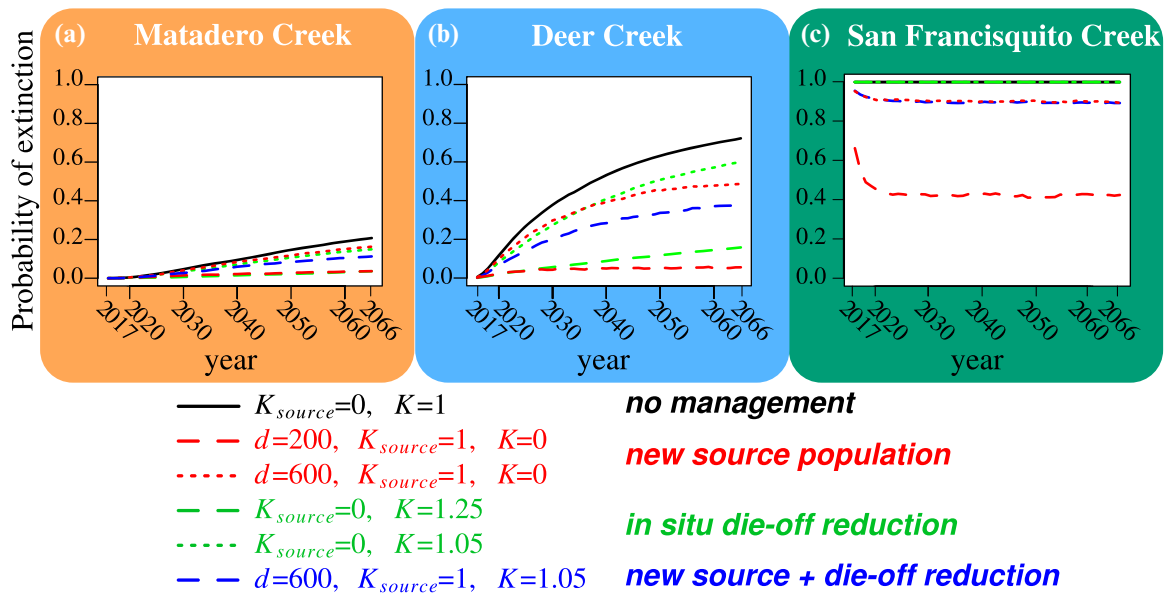


Figure 5. Cumulative probability of complete within-creek extinction of California red-legged frogs in 50 years computed using the maximum a posteriori estimates of extinction and colonization parameters: (a) Matadero, (b) Deer, and (c) San Francisquito Creeks. Maximum a posteriori extinction and colonization parameters are in Fig. 3.

starting between the 1960s and the early 1980s was the most likely explanation for decline in Deer and San Francisquito Creeks (Supporting Information). Several factors could contribute to in situ die-off. Unfortunately, our model does not distinguish the effects of disease and introduced predators, phenomena that could have had similar impacts on in situ die-off. Infection by Bd (*Batrachochytrium dendrobatidis*), a fungal pathogen, is widespread in amphibians in California (Padgett-Flohr & Hopkins 2009) and worldwide (Fisher et al. 2012). Infections with Bd were observed in 1961 on the Stanford campus (Padgett-Flohr & Hopkins 2009), but might have begun in California as early as the late 1800s (Huss et al. 2013). Bullfrogs (*Lithobates catesbeianus*), a non-native predator of *R. draytonii*, were introduced into California around 1900 (Stebbins & McGinnis 2012), and signal crayfish (*Pacifastacus leniusculus*) were introduced in San Francisco watersheds as early as 1898 (Fofonoff et al. 2017). Non-native predators present in Stanford creeks, including mosquitofish (*Gambusia affinis*) (Lawler et al. 1999), bullfrogs (Moyle 1973), and signal crayfish (Allan and Tennant 2000), have contributed to similar *R. draytonii* declines elsewhere in California. The trajectory of population decline suggests extinction debt, where local extinction occurs after substantial delay following habitat degradation (Kuussaari et al. 2009).

Predicted Dynamics and Conservation

Our method predicted future persistence probabilities for the *R. draytonii* populations. Interestingly, despite

similar ratios of extinction and colonization rates in Matadero and Deer Creeks, stochastic dynamics produced a greater 50-year persistence probability in Matadero Creek (Fig. 5). In Deer Creek, large estimated extinction and colonization rates magnified variability in occupancy. Thus, successive years of substantial occupancy reduction could lead to chance extinction. In 1 year, only 2 segments were occupied in Deer Creek. With extinction parameter $e = 0.39$, because extinction precedes colonization, the probability of complete extinction in the next year depended only on the extinction rate ($0.39^2 = 0.1521$). In Matadero Creek, however, smaller extinction and colonization parameters and consequent lower occupancy variation made successive years of substantial occupancy reduction unlikely. For example, in 3 recent instances, 3 segments were occupied ($e = 0.12$), and the extinction probability was a relatively safe $0.12^3 = 0.001728$. Sensitivity to years with substantial occupancy reduction raises concerns for persistence under unpredictable climate change: a warming trend in California, where precipitation deficits and increased incidence of extreme drought (Diffenbaugh et al. 2015) have occurred, potentially threatens aquatic species (Meyer et al. 1999).

Our results suggest a preferred approach to maintenance of *R. draytonii* populations, predicting that establishment of a source population <200 m from Matadero and Deer Creeks would reduce extinction risk (Fig. 5). The greatest threat is stochasticity of extinction and colonization; adding a source would facilitate survival by population rescue. The Stanford

Conservation Program has pursued this management approach in a similar project, creating ponds that support California tiger salamander (*Ambystoma californiense*) reproduction. Recovery results have also been seen in other similar projects (Petranka et al. 2007; Rannap et al. 2009). Perhaps surprisingly, among possible conservation actions, the preferred action (adding a source) did not most directly reverse the most likely past disturbance (increased in situ die-off). Our results suggest prioritizing management actions for Deer Creek is likely to provide the most risk reduction for a given level of effort.

Extensions

Incorporating life-stage or demographic data by modeling local population sizes as temporally varying rather than constant could enable modeling of disturbances affecting specific life stages. It may also enable modeling of phenological shifts in life-history dynamics, such as differing dispersal by life stage. However, when data on population size are unavailable, as in our frog example, models integrating local population dynamics are not applicable (Sutherland et al. 2014). Another extension would allow patches to have different carrying capacities K_i , if estimates of these quantities are known. This change would simply rescale expressions for extinction and colonization rates E and C .

Our model permits nonlinear habitats and unequal patch sizes. Computations require only pairwise distances between populations and could therefore be applied to many geometries. Scenarios beyond increased in situ die-off and source population loss can be examined. For example, reductions affecting only certain patches would reduce some population sizes K_i , cyclic droughts would periodically push certain population size values K_i to 0, and so on. A covariate x that influences extinction or colonization can be accommodated by making the associated parameters in patch i at time t , $e_{i,t}$ and $c_{i,t}$, functions of $x_{i,t}$, for example with a logit link $e_{i,t}(x_{i,t}) = [\exp(a_0 + a_1 x_{i,t})] / [1 + \exp(a_0 + a_1 x_{i,t})]$. Posterior distributions of parameters a_0 and a_1 can then be inferred.

In summary, our method enables assessments of influences on population persistence of complex scenarios, including local (e.g., adding new patches) or global actions (e.g., reducing predation or disease). It can both assess precise scenarios (e.g., specifying new habitat locations) and perform exploratory analyses (e.g., predictions across a range of possible locations). In our system, it showed that the best management action for reversing population decline was not the action that most directly reversed the original cause of decline. This observation is particularly important for management of species for which threats are hard to address and suggests that alternative strategies can improve persistence. We encourage use of our framework for advancement of ecological theory and management.

Acknowledgments

We thank the Stanford Creek Monkeys for their tireless field work and data collection over many years. This work was supported by Swiss National Science Foundation Early Postdoc.Mobility fellowship P2LAP3_161869.

Supporting Information

Supplementary methods (Appendices S1 to S5), parameter estimation (Appendix S6), hypothesis tests (Appendix S7), model selection based on AIC (Appendix S8), parameter estimation under exact and approximate algorithms (Appendix S9), parameter estimation of dispersal distance (Appendix S10), parameter estimation in Deer Creek with an uninformative prior for missing occupancy (Appendix S11), accuracy of parameter estimation from Fig. 3 (Appendix S12), accuracy of model comparisons (Appendix S13), estimation of timing of population reduction event (Appendix S14), estimation of model parameters under hypothesis 1 in Matadero Creek (Appendix S15) and hypothesis 2 (Appendix S16), estimated occupancy trajectories (Appendix S17), parameter estimation based on many-patches approximate algorithm (Appendix S18), creek segment lengths (Appendix S19), segment occupancies (Appendix S20), and software MIDASPOM (Appendix S21) are available online.

Literature Cited

- Adams MJ, et al. 2010. Using occupancy models to understand the distribution of an amphibian pathogen, *Batrachochytrium dendrobatidis*. *Ecological Applications* 20:289–302.
- Allan MF, Tennant T. 2000. Evaluation of critical habitat for the California red-legged frog (*Rana aurora draytonii*). UC Riverside: Center for Conservation Biology, Riverside, California. Available from <https://escholarship.org/uc/item/7027275m> (accessed February 2017).
- Bailey LL, MacKenzie DI, Nichols JD. 2014. Advances and applications of occupancy models. *Methods in Ecology and Evolution* 5:1269–1279.
- Biek R, Funk WC, Maxell BA, Mills LS. 2002. What is missing in amphibian decline research: insights from ecological sensitivity analysis. *Conservation Biology* 16:728–734.
- Bulger JB, Scott NJ, Seymour RB. 2003. Terrestrial activity and conservation of adult California red-legged frogs *Rana aurora draytonii* in coastal forests and grasslands. *Biological Conservation* 110:85–95.
- Cole EM, North MP. 2014. Environmental influences on amphibian assemblages across subalpine wet meadows in the Klamath mountains, California. *Herpetologica* 70:135–148.
- Davidson C, Shaffer HB, Jennings MR. 2002. Spatial tests of the pesticide drift, habitat destruction, UV-B, and climate-change hypotheses for California amphibian declines. *Conservation Biology* 16:1588–1601.
- Diffenbaugh NS, Swain DL, Touma D. 2015. Anthropogenic warming has increased drought risk in California. *Proceedings of the National Academy of Sciences* 112:3931–3936.
- Fellers GM. 2005. *Rana draytonii* Baird and Girard 1852, California red-legged frog. In Lannoo MJ, editor. *Amphibian declines: the conservation status of United States species*. Volume 2. University of California Press, Berkeley.

- Fellers GM, Kleeman PM. 2007. California red-legged frog (*Rana draytonii*) movement and habitat use: implications for conservation. *Journal of Herpetology* **41**:276–286.
- Fisher MC, et al. 2012. Emerging fungal threats to animal, plant and ecosystem health. *Nature* **484**:186–194.
- Fisher RN, Shaffer HB. 1996. The decline of amphibians in California's Great Central Valley. *Conservation Biology* **10**:1387–1397.
- Fiske I, et al. 2011. Unmarked: An R package for fitting hierarchical models of wildlife occurrence and abundance. *Journal of Statistical Software* **43**:1–23.
- Fofonoff PW, Ruiz GM, Steves B, Simkanin C, Carlton JT. 2017. National exotic marine and estuarine species information system. Smithsonian Environmental Research Center, Edgewater, Maryland. Available from <https://invasions.si.edu/nemesis/calnemo/SpeciesSummary.jsp?TSN=97326> (accessed January 2018).
- Gormley AM, et al. 2011. Using presence-only and presence-absence data to estimate the current and potential distributions of established invasive species. *Journal of Applied Ecology* **48**: 25–34.
- Gyllenberg M, Silvestrov DS. 1994. Quasi-stationary distributions of a stochastic metapopulation model. *Journal of Mathematical Biology* **33**:35–70.
- Hanski I. 1994. A practical model of metapopulation dynamics. *Journal of Animal Ecology* **63**:151–162.
- Hanski I, Ovaskainen O. 2000. The metapopulation capacity of a fragmented landscape. *Nature* **404**:755–758.
- Hayes MP, Jennings MR. 1988. Habitat correlates of distribution of the California red-legged frog (*Rana aurora draytonii*) and the foothill yellow-legged frog (*Rana boylei*): implications for management. Pages 144–158 in Szaro RC, Severson KE, Patton DR, editors. *Proceedings of the symposium on management of amphibians, reptiles and small mammals*. U.S. Department of Agriculture Forest Service, Flagstaff, Arizona.
- Huss M, Huntley L, Vredenburg V, Johns J, Green S. 2013. Prevalence of *Batrachochytrium dendrobatidis* in 120 archived specimens of *Lithobates catesbeianus* (American bullfrog) collected in California, 1924–2007. *EcoHealth* **10**:339–343.
- Jeffreys H. 1998. *The theory of probability*. Oxford University Press, Oxford, United Kingdom.
- Johnson DS, Conn PB, Hooten MB, Ray JC, Pond BA. 2013. Spatial occupancy models for large data sets. *Ecology* **94**:801–808.
- Kéry M, Schaub M. 2011. *Bayesian population analysis using WinBUGS: a hierarchical perspective*. Academic Press, Waltham, Massachusetts.
- Kuussaari M, et al. 2009. Extinction debt: a challenge for biodiversity conservation. *Trends in Ecology & Evolution* **24**:564–571.
- Lannoo MJ. 2005. *Amphibian declines: the conservation status of United States species*. University of California Press, Oakland, California.
- Lawler SP, Drits D, Strange T, Holyoak M. 1999. Effects of introduced mosquitofish and bullfrogs on the threatened California red-legged frog. *Conservation Biology* **13**:613–622.
- MacKenzie DI, Nichols JD, Hines JE, Knutson MG, Franklin AB. 2003. Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. *Ecology* **84**:2200–2207.
- Meyer JL, Sale MJ, Mulholland PJ, Poff NL. 1999. Impacts of climate change on aquatic ecosystem functioning and health. *Journal of the American Water Resources Association* **35**:1373–1386.
- Moilanen A. 2004. SPOMSIM: software for stochastic patch occupancy models of metapopulation dynamics. *Ecological Modelling* **179**:533–550.
- Moyle PB. 1973. Effects of introduced bullfrogs, *Rana catesbeiana*, on the native frogs of the San Joaquin Valley, California. *Copeia* **1973**:18–22.
- O'Hara RB, Arjas E, Toivonen H, Hanski I. 2002. Bayesian analysis of metapopulation data. *Ecology* **83**:2408–2415.
- Padgett-Flohr GE, Hopkins RL. 2009. *Batrachochytrium dendrobatidis*, a novel pathogen approaching endemism in central California. *Diseases of Aquatic Organisms* **83**:1–9.
- Petranka JW, Harp EM, Holbrook CT, Hamel JA. 2007. Long-term persistence of amphibian populations in a restored wetland complex. *Biological Conservation* **138**:371–380.
- Rannap R, Lohmus A, Briggs L. 2009. Restoring ponds for amphibians: a success story. *Hydrobiologia* **634**:87–95.
- Risk BB, De Valpine P, Beissinger SR. 2011. A robust-design formulation of the incidence function model of metapopulation dynamics applied to two species of rails. *Ecology* **92**:462–474.
- Royle JA, Dorazio RM. 2008. *Hierarchical modeling and inference in ecology: the analysis of data from populations, metapopulations and communities*. Academic Press, London.
- Stebbins RC, McGinnis SM. 2012. *Field guide to amphibians and reptiles of California: revised edition*. Volume **103**. University of California Press, Berkeley, California.
- Sutherland CS, Elston DA, Lambin X. 2014. A demographic, spatially explicit patch occupancy model of metapopulation dynamics and persistence. *Ecology* **95**:3149–3160.
- ter Braak CJ, Etienne RS. 2003. Improved Bayesian analysis of metapopulation data with an application to a tree frog metapopulation. *Ecology* **84**:231–241.
- ter Braak CJ, Hanski I, Verboom J. 1998. The incidence function approach to modeling of metapopulation dynamics. Pages 167–188 in Bascompte J, Solé RV, editors. *Modeling spatio-temporal dynamics in ecology*. Landes Biosciences, Austin, Texas.
- Vonesh JR, De la Cruz O. 2002. Complex life cycles and density dependence: assessing the contribution of egg mortality to amphibian declines. *Oecologia* **133**:325–333.
- Welsh HH, Pope KL, Boiano D. 2006. Sub-alpine amphibian distributions related to species palatability to non-native salmonids in the Klamath mountains of northern California. *Diversity and Distributions* **12**:298–309.
- With KA, King AW. 1999. Extinction thresholds for species in fractal landscapes. *Conservation Biology* **13**:314–326.
- Yackulic CB, et al. 2012. Neighborhood and habitat effects on vital rates: expansion of the barred owl in the Oregon Coast Ranges. *Ecology* **93**:1953–1966.

