USING ENVIRONMENTAL DNA AND OCCUPANCY MODELING TO ESTIMATE RANGEWIDE METAPOPULATION DYNAMICS OF THE ENDANGERED TIDEWATER GOBY *EUCYCLOGOBIUS Spp.*

By

Chad Martel

A Thesis Presented to
The Faculty of Humboldt State University
In Partial Fulfillment of the Requirements for the Degree
Masters of Science in Natural Resources: Fisheries

Committee Membership
Dr. Andrew P. Kinziger, Committee Chair
Dr. Andre Buchheister, Committee Member
Dr. Nicholas A. Som, Committee Member
Dr. Rick Zechman, Graduate Program Coordinator

July 2019

ABSTRACT

USING ENVIRONMENTAL DNA AND OCCUPANCY MODELING TO ESTIMATE RANGEWIDE METAPOPULATION DYNAMICS IN AN ENDANGERED SPECIES, TIDEWATER GOBY *EUCYCLOGOBIUS Spp.*

Chad Martel

Conservation of species is most effective when metapopulation dynamics are well understood and incorporated into management plans, allowing managers to target conservation efforts where they will be most effective. The development of environmental DNA (eDNA) methods provides an efficient and highly sensitive approach to generate presence and absence data needed to elucidate metapopulation dynamics. Combining sample detection histories from eDNA surveys with occupancy models that account for non-detection can offer unbiased estimates of rangewide metapopulation dynamics. However, traditional occupancy models do not allow direct evaluation of the drivers of site occupancy, extinction, and colonization. Herein, I utilize a *novel* dynamic multiscale occupancy model that accounts for non-detection to estimate rangewide metapopulation dynamics and their drivers in an endangered fish, tidewater goby *Eucyclogobius spp.*, a genus endemic to California estuarine habitats. I collected rangewide eDNA data from 190 sites (813 total water samples) surveyed from two years (2016 and 2017) and analyzed the data using a dynamic multiscale occupancy model. Rangewide, estimates of the proportion of sites that were occupied varied little between 2016 (0.524) and 2017 (0.517). Although I uncovered stability in the number of sites that were occupied across the two study years, there was evidence for extinction and colonization dynamics. Rangewide estimates of extinction probability of occupied sites (0.106) and colonization probability of unoccupied sites (0.085) were nearly equal. The consistent rangewide occupancy proportions combined with the presence of extinctions and colonizations suggests a dynamic equilibrium between the two years surveyed. There was no latitudinal gradient or regional differences in extinction and colonization dynamics across the tidewater goby geographic range. Assessment of covariate effects on metapopulation dynamics revealed that colonization probability increased as the number of occupied neighboring sites increased and as distance between occupied sites decreased. I show that eDNA surveys can rapidly provide a snapshot of a species distribution over a broad geographic range, and when these surveys are paired with dynamic multiscale occupancy modeling, they can uncover rangewide and regional scale metapopulation dynamics and their drivers.

ACKNOWLEDGMENTS

I would like to thank Dr. Andrew Kinziger for giving me the opportunity to participate in this project and his long running research into tidewater goby. This work would not have been possible without his support, encouragement, and mentorship. Thank you to my committee members, Dr. Andre Buchheister and Dr. Nicholas Som for their support and their time; both have been invaluable. A big thank you to Dr. Robert Dorazio for making this analysis possible and guiding it along the way as we made it as thorough as it needed to be. Thank you to those who provided financial support; without you this research would not have been possible: the US Fish and Wildlife Service, the Marin Rod and Gun Club, and the Humboldt Fishin' Lumberjack/Roelofs Humboldt Fisheries Fund. Thank you to Leslie Farrar and the California Cooperative Fish and Wildlife Research Unit at Humboldt State University for helping to keep this project moving forward. Thank you to David Baston in the HSU BioCore for assistance in and out of the lab. Thank you to my lab mates and those who accompanied me in the field: Michael Sutter, Ely Boone, Jaycee Gail-Owsley, and Ian Kelmartin. Thank you to those who assisted in site access and sample collections: Doreen Hansen at the Humboldt County Resource Conservation District, Alex Blessing with the Wildlands Conservancy at Eel River Estuary Preserve, Rhys Evans at Vandenberg Air Force Base, Doug Gibson at the San Elijo Lagoon Conservancy, Sally Brown with the USFWS, Keith Parker with the Yurok Tribe of California, HSU Boating Safety Officer Steve Monk, and Wayne and Susan Trivelpiece for lodging. Lastly, thank you to my friends and family that loved and

supported me throughout this process; from the supportive hugs to the couches I slept on and garages I filtered in, you deserve more thanks than I can offer.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

INTRODUCTION

Metapopulation dynamics, as originally conceived by Levin (1969), assumes that a set of isolated sites of suitable habitat can periodically experience local extirpations if occupied, while in the same time step some unoccupied sites may be recolonized by dispersing individuals. Conservation and management decisions regarding a species thought to exist as a metapopulation are most effective when extinction and colonization dynamics are understood and incorporated into management plans (Anthes et al. 2003, Armstrong 2005, Marsh and Trenham 2001, Semlitsch 2008, Ying et al. 2011). The understanding and effective incorporation of extinction and colonization dynamics requires accurate, unbiased, estimates of extinction and colonization across their range. Despite the need for accurate estimation of metapopulation dynamics, these dynamics are notoriously difficult to describe and subject to a number of potential errors, including incorrect estimation of site size, unknown sites existing in a study area, and the potential for non-detections, where non-detection of a target species is interpreted as a true absence when the target was actually present (Moilanen 2002, MacKenzie 2003).

Extinction and colonization rates of metapopulations are typically estimated using a record of site occupancy across repeated field surveys through time. These presence-absence data are then used to generate site occupancy histories for many habitat sites from which extinction and colonization of individual sites is inferred. A critical assumption of this approach is that non-detection is indicative of a true absence. However, it is well-established that this assumption is violated for most field survey methods, especially when abundance is low, field collections are difficult, or collection effort is limited (Gu & Swihart 2004, MacKenzie and Royle 2005). Confusing non-detection as an absence leads to biased estimates of site occupancy (Gu & Swihart 2004, Moilanen 2002), and for metapopulation studies non-detection errors can lead to biased estimates of extinction and colonization dynamics. Unlike occupancy, which is biased low by non-detections, bias in extinction and colonization dynamics can be over- or under- estimated depending on which time period the non-detection error occurred. In response to non-detection biases, analytic methods have been developed to account for non-detection (MacKenzie et al., 2003, Moilanen 2002). Non-detection can be accounted for by completing multiple surveys of each habitat site within a single season and analyzing the resulting occupancy data with models that use the within season sampling data to estimate detection probabilities (Mackenzie et al. 2002; MacKenzie et al. 2003, Moilanan 2002).

A key goal of metapopulation studies is to elucidate environmental drivers of extinction and colonization dynamics (Hanksi 1989, 1998, Hanski and Gilpin1991). Armstrong (2005) suggested species declines can be halted by the integration of two paradigms: (1) a metapopulation paradigm which focuses on factors that influence site connectivity such as site size, spatial structure, and site density, and (2) a habitat paradigm that emphasizes the link between local persistence and site-level environmental covariates. Metapopulation paradigm models are predicated on two widely accepted and generally supported assumptions: (1) isolated sites have lower colonization rates relative to more densely spaced sites, and (2) smaller sites are more vulnerable to local extirpation relative to larger habitat sites (Armstrong 2005). Research within the habitat paradigm has shown how habitat factors such as topography, vegetation, and available resources can affect site occupancy (Akçakaya and Atwood 2002, Fleishman et al. 2002, Thomas et al. 2001). Generally, high quality habitat within a site leads to colonization or persistent occupation while poor quality habitat can lead to extirpations or lower occupancy rates (Armstrong 2005, Fleishman et al. 2002, Franken and Hik 2004, Thomas 1994).

Few studies have elucidated rangewide metapopulation dynamics and their environmental drivers owing to time and money constraints. Repeated rangewide surveys require extensive planning, particularly if the area is large, the sites are numerous, and the target is small or cryptic (Moilanen 2002). Environmental DNA (eDNA) has gained popularity in the past decade as a method of rapidly and efficiently detecting species (Foote et al. 2012, Gingera et al. 2016, Goldberg at al. 2013, Pilliod et al. 2013). Environmental DNA surveys are a method of surveying the environment for genetic material that has been sloughed off an individual to use as an index of presence. Environmental DNA has been repeatedly shown to outperform traditional methods of species detection, often detecting a species when traditional methods fail (Boussarie et al. 2018, Dejean et al. 2012, Port et al. 2016, Schmelzle and Kinziger 2016, Thompsen and Willerslev 2015). Direct comparisons have shown that eDNA monitoring can have double the detection probability of traditional approaches (Schmelzle and Kinziger 2016).

Environmental DNA data collected using a hierarchical, multi-season sampling design can be analyzed via multiscale occupancy models (Dorazio and Erickson 2018, Sutter and Kinziger 2019). Data from this project were analyzed using a *novel* class of *dynamic* multiscale occupancy model developed in collaboration with Dr. Robert Dorazio of San Francisco State University (a description of the model, authored by Dr. Dorazio, can be found in Appendix A). Briefly, the occupancy model used here belongs to a class of multiscale occupancy models that uses a hierarchical structure to allow estimation of non-detection at different sampling levels. In the case of eDNA surveys, there are three sampling levels: sites, samples, and qPCR replicates (Dorazio and Erickson 2018). Multiscale models do not directly estimate extinction and colonization dynamics; detecting extinction and colonization dynamics using requires a supplementary analysis of the change in modeled site occupancy across seasons. The novel '*dynamic*' class of model used in this study builds on previous multiscale models by incorporating extinction and colonization as model parameters, thus allowing testing of potential covariates that may drive extinction and colonization dynamics. This model assumes that sites are closed to immigration during sampling periods but open between sampling periods. This assumption allows estimation of detection probabilities at site, sample, and qPCR replicate levels as well as extinction and colonization dynamics.

To demonstrate the power of estimating metapopulation dynamics using eDNA surveys combined with multiscale occupancy modeling, I analyzed tidewater goby *Eucyclogobius spp*. occupancy data from two years of repeated rangewide eDNA surveys. Tidewater goby are a benthic fish endemic to coastal California where they inhabit lagoons, bays, and the estuaries of rivers and streams (Swenson 1999). Originally thought of as one species across their range, recent genetic research has divided the genus into two distinct species (Swift et al. 2016). The southern tidewater goby *E. kristinae* is known from only nine site south of the Palos Verdes peninsula, Los Angeles County, to San Diego County, whereas the northern tidewater goby *E. newberryi* is known from 70+ sites between the Palos Verdes peninsula and the California-Oregon border (Swift et al. 2016). The use of ‘tidewater goby’ in this text refers to either species, or the *Eucyclogobius* genus as a whole. Tidewater goby were listed as endangered under the U.S. Endangered Species Act in 1994; the most recent recovery plan for tidewater goby divides the California coast into six “recovery units.” From north to south these recovery units are North Coast, Greater Bay Area, Central Coast, Conception, Los Angeles/Ventura, and South Coast (USFWS 2005, Figure 1A). Northern tidewater goby are found in the North Coast, Greater Bay Area, Central Coast, Conceptions, and Los Angeles/Ventura recovery units, while the southern tidewater goby range is encompassed by the South Coast recovery unit. Each of these recovery units are further divided into various numbers of sub-units, 26 in total, for recovery and management efforts.

Tidewater goby metapopulation population structure has been described as a core-satellite model where populations in large wetlands serve as core populations with relatively high persistence and high dispersal while populations in smaller wetlands may experience higher extinction rates, possibly driven by unfavorable dry conditions (Lafferty et al. 1999a, 1999b). Under this model, recolonization of extirpated sites would be driven by individuals originating from the larger, more stable, core populations (Lafferty et al. 1999a, 1999b). Observations by Lafferty et al. (1999b) suggest that colonization may be promoted by localized flooding of small, occupied, streams where the associated increase in longshore current could deliver individuals flushed from one site to another unoccupied site. In this way, flooding may be beneficial to tidewater goby persistence at the rangewide level by allowing colonization of previously extirpated sites. Metapopulation dynamics have been incorporated into management and recovery efforts across their range (USFWS 2005). The most recent recovery plan states that in order to consider tidewater goby for downlisting a metapopulation viability analysis must find a greater than 75% chance of all recovery units surviving out to 100 years. But this requirement does not appear to be the best fit across all tidewater goby recovery units; in northern California, geographically separated populations were found to lack any signal of extinction and colonization dynamics over time scales several decades (Kinziger et al. 2015).

The objective of this study was to estimate rangewide metapopulation dynamics for endangered tidewater goby by analyzing site occupancy histories generated from eDNA surveys with a *novel* class of *dynamic* multiscale occupancy model that explicitly accounts for non-detection at the three hierarchical levels inherent to eDNA surveys: sites, samples, and qPCR replicates while allowing estimation of extinction and colonization as model parameters with covariates. The use of eDNA survey methods enabled rangewide monitoring of 190 locations in two consecutive years along the entire 1350 km of the California coast (Figure 1A). The multiscale dynamic occupancy model was used to examine: (1) rangewide occupancy, extinction, and colonization and how the measured covariates (vegetation, salinity, temperature, dissolved oxygen, estuary size, and distance between neighboring sites) affected these dynamics, (2) the biases resulting from naïve estimates of metapopulation dynamics and their relationship to model estimates, (3) extinction and colonization within recovery units to better understand the spatial variability in tidewater goby metapopulation dynamics, and lastly (4) which covariates (tide, salinity, turbidity, temperature, and dissolved oxygen) affected detection of tidewater goby eDNA in water samples and in qPCR. This work illustrates that eDNA, when used in combination with multiscale occupancy modeling, has the ability to efficiently survey a large number of locations, detect metapopulation dynamics, and evaluate their drivers, at both broad and narrow geographic scales.

METHODS

Data Collection, Field, and Laboratory Methods

Site occupancy histories were available for two years, 2016 and 2017. The 2016 site occupancy histories are described by Sutter and Kinziger (2019), and the 2017 site occupancy data was generated by the author following Sutter and Kinziger (2019). This created a congruent dataset between the Sutter and Kinziger (2019) survey conducted in 2016 and this follow-up 2017 survey thereby allowing analysis of the combined occupancy and environmental data from both surveys. Between May and September of 2017, I visited 194 sites, from which I collected between one and six 2.0 L water samples (413 in total). I conducted between three and six qPCR replicates per water sample. These surveys ran along the California Coast from 2 km south of California-Oregon border, to 45 km north of the California-Mexico border (Figure 1A). I used Global Positioning System (GPS) data and photographs from the 2016 survey to collect samples as near as possible to where the previous samples were collected, usually within 100 m. If access or conditions had changed between the two seasons, samples were collected as close as possible to the 2016 locations while duplicating sampling protocol from Sutter and Kinziger (2019); i.e. maintaining distance between multiple samples at a site to avoid the transport of eDNA from disturbed sediment or water from a previous sampling location to another location. Three of the sites sampled in 2016 were not resampled in 2017 due to access restrictions. Environmental data collected with each water sample included water temperature (°C), dissolved oxygen (mg/L), salinity (ppt), presence or absence of tidal influence (open or closed to daily tidal flow at time of sampling), and presence or absence of aquatic vegetation. Additional data collected at each sample location included date, time, and latitude and longitude.

Water filtering, eDNA extraction, and qPCR protocols were conducted following Sutter and Kinziger (2019). Samples were filtered over a 47 mm diameter polycarbonate track-etched filter membrane with 3.0 µm pore size (Millipore TSTP 04700). Extraction of eDNA from filters was conducted using a DNeasy Blood and Tissue Kit (Qiagen, USA) with modifications for eDNA extraction (Schmelzle and Kinziger 2016, Sutter and Kinziger 2019). Detection of northern and southern tidewater goby was completed using two assays that are designed to target a 119 base pairs region on tidewater goby's mitochondrial cytochrome *b* gene (Schmelzle and Kinziger 2016, Sutter and Kinziger 2019). A northern tidewater goby assay was used to test for presence of northern tidewater goby eDNA in samples taken north of Palos Verdes (North Coast, Greater Bay Area, Central Coast, Conception, and Los Angeles/Ventura recovery units) while a southern tidewater goby eDNA assay was used to test for presence of southern tidewater goby eDNA in samples taken from south of Palos Verdes (South Coast recovery unit). Extraction of eDNA from filters and qPCR preparation was performed on separate benchtops in a dedicated laboratory space away from high copy number PCR product. Additionally, qPCR preparations were performed under a hood with high efficiency particulate air (HEPA) filter that was exposed to ultra-violet light for at least 30 minutes prior to use, along with pipettes, centrifuges, and consumables used in reaction preparation.

Spatial data for sites

Spatial data for sites were obtained from the California Department of Fish and Wildlife (CDFW) and Pacific States Marine Fisheries Council (PSMFC). A California shore type data, which classifies the California coast by habitat using NOAAs Environmental Sensitivity Index (ESI), was obtained from CDFW. This was used to measure the total distance between pairs of sites within sub-units and to measure the distance between pairs of sites within sub-units that are rocky. Distance between sites within sub-units was determined using program *R* and the package RIVERDIST. To generate matrices of shoreline distance the shoreline data was altered to maintain a continuous shoreline; embayments, estuaries, or rivers were removed from the original shoreline data and the data was reconnected across their openings in the shortest path. The distance of rocky coast between sites within subunits was obtained using ArcMap 10.5.1 (*Esri Inc*.) and the California shore type dataset. I recorded any length of coast between sites in each sub-unit whose primary ESI category was one of the following: 1A (exposed rocky shore/exposed), 1B (solid man-made structures), 1C (exposed rocky cliffs with boulder talus), 2A (exposed wave-cut platforms in bedrock, mud, or clay), or 2B (exposed scarps and steep slopes in clay). Lastly, estuary size was determined using the West Coast USA Current and Historical Estuary Extent data set from the Pacific Marine and Estuarine Fish Habitat Partnership (PSMFC GIS 2017). Sites were assigned to one of two bins based on estuary size, those sites equal to or less than 1.0 ha were deemed “small” while “large” sites were over 1.0 ha following Lafferty (1999).

Occupancy data

For the two species assays used in these analyses, Sutter and Kinziger (2019) previously determined their limit of detection and associated critical threshold values (Ct), this value is the cycle number at which a qPCR replicate reaches a threshold of fluorescence for a positive reaction. For the northern tidewater goby assay, a Ct value of 40.87 or lower was assumed to be a positive eDNA detection. For the southern tidewater goby assay a Ct of 40.04 or lower was assumed to be a positive detection. This data was passed to the multiscale occupancy model as either a positive (1) or negative (0) detection for each of the qPCR replicates preformed on a sample. I assumed that a single positive detection was indicative of tidewater goby presence at that location, and that there were not false positives.

Occupancy Model

Data collected by this research was analyzed using a dynamic multiscale occupancy model (see Appendix A for full details). Briefly, the model was used to estimate rangewide occupancy in 2016 (ψ_1) and 2017 (ψ_2), rangewide extinction (ε), rangewide colonization (η), and colonization within specified neighborhoods (ω) between 2016 and 2017. The model accounted for non-detection errors by estimating the probability of detecting eDNA in a water sample given its presence at a site (θ), and probability of detecting eDNA in a qPCR replicate given its presence in water sample (ρ). Occupancy, detection in a water sample, and detection in a qPCR replicate are estimated using the likelihood function $L(\psi, \theta, p \mid H) \propto \prod_{i=1}^{S} Pr(H_i)$, where H_i represents qPCR detection history at site *i*, across all possible sites *S*. Site occupancy state is then predicted for single sites in a single year as either an absence (0) or a presence (1). Occupancy states are represented in the form of a $Z_{i,t}$ value, where $Z_{i,t}$ is the occupancy state at site *i* at time *t*. The modeled occupancy state for all sites are compared across time *t* and *t+1* to estimate the probability of a site's occupancy state transitioning from state *k* in time *t* to state *l* in time *t+1*; or $\Phi_{k,l,i,t} = \Pr(Z_{i,t+1} = l \mid Z_{i,t} = k)$. The transition probabilities are contained in the transition matrix $\Phi_{i,t}$ and represent extinction (ε) and colonization (η) probabilities. If an unoccupied site has occupied neighbors, it is assumed that any colonization at that site stems from its occupied neighbors within its recovery sub-unit, and not from outside the sub-unit. The structure of this model allows separation of metapopulation parameter estimates (occupancy, extinction, and colonization) based on geography. This allows the estimation of occupancy, extinction, and colonization within the tidewater goby management delineated recovery units. Consequently, results are presented at both the rangewide geographic scale as well as the scale of the recovery unit. This model is Bayesian based; the models Markov Chain-Monte Carlo algorithm was run for 110,000 iterations, of which the first 10,000 were discarded.

Covariate data

Model parameters were modeled as functions of different environmental and spatial covariates. Instead employing a model selection approach where all possible covariate combinations were tested and ranked, each covariate was carefully and individually considered before being included as a possible covariate for a parameter in the analysis. Consideration was based on previously published literature regarding tidewater goby ecology and eDNA methods (Table 1). The environmental covariates examined included salinity (ppt), water temperature (°C), dissolved oxygen (mg/l), depth (cm), turbidity (seconds of filtering time), proportion of sample locations with aquatic vegetation, and presence or absence of tidal influence. The two spatial covariates investigated were the pairwise distance between sites within sub-units and the size of the estuary at each site. The distances between sites within a sub-unit were used as the basis for a neighborhood analysis that examined the probability of colonization from neighboring sites. Further explanation of which covariates were tested for each parameter can be found in Table 1. Model results and the strength of evidence for each parameter-covariate relationship were considered individually via the posterior distribution of each parameter-covariate relationship; evidence of parameter-covariate relationships is presented as the proportion or percent of the posterior distribution above or below zero.

RESULTS

Of the original 194 sites surveyed in both 2016 and 2017, a total of 190 were available for analysis (Figure 1A). One site was removed because samples were collected at a location where tidewater goby occupancy seemed highly improbable, the habitat was riverine in nature and not proximal to the river's estuary. Three other sites were removed due to missing covariate values. The analysis was conducted on a combined data set that included both the northern and southern tidewater goby species; this was done because southern tidewater goby are restricted too few locations to allow analysis using the model employed herein. The parameter-covariate relationships examined by the dynamic multiscale model can be found in Table 1. The posterior distributions for each relationship were considered individually, those distributions that were strongly shifted away from zero were considered to represent a parameter-covariate relationship. Overall, occupancy (ψ) exhibited a relationship to the proportion of sample locations at a site with vegetation, but not with salinity (Figure **2**). Colonization within sub-units (ω) decreased with increasing distance between sites. None of the tested covariates, including estuary size, salinity, temperature, and dissolved oxygen, had a significant effect on probability of extinction (ε, Figure **2**). Detection in a water sample (θ) was found to vary with salinity, dissolved oxygen, and the presence of tide (Figure 3). Turbidity and temperature had no effect on detection in water samples. Detection in a qPCR replicate (ρ) was found to vary with salinity (Figure 3).

Rangewide

The dynamic multiscale model estimated that the proportion of sites occupied by tidewater goby was the same in 2016 and 2017 (Figure 2). The mean posterior probability of site occupancy in 2016 (ψ_1) was 0.52, with 95% of the posterior observations in the credible interval (95% CRI) 0.51 to 0.56. In 2017, ψ_2 was estimated as 0.51 (95% CRI: 0.50 to 0.56). In terms of number of sites, the model estimates that in 2016 there were about 99 (95% CRI: 97 to 107) occupied sites and in 2017 there were about 97 (95% CRI: 94 to 106) occupied sites. Thus, the number of sites occupied by tidewater goby was essentially unchanged between the two years surveyed. The naïve rangewide occupancy rates were lower than the model estimates: 0.44 in 2016 and 0.41 in 2017, or 83 occupied sites in 2016 and 78 occupied sites is 2017 (Figure 2). The posterior distribution for the effect of vegetation shows strong support (99.9% above zero) that probability of occupancy at a site increases as the proportion of sample locations with vegetation increase (Figure 3); whereas there is no evidence (53.2% above zero) that salinity has an effect on site occupancy (Figure 4, Table 2). Probability of site occupancy increased from 0.42 (95% CRI: 0.27 to 0.58) at sites with no vegetation to a maximum of 0.68 (95% CRI: 0.55 to 0.79) at sites where vegetation was present at all sample locations (Figure 3).

The lack of change in proportion of occupied sites between years was not indicative of an absence of metapopulation dynamics. Rather, both extinction and colonization rates were predicted nearly equal rates (Figure 5). The mean of the posterior distribution of the probability of an extinction of an occupied site was 0.11 (95% CRI:

0.03 to 0.22). In terms of number of sites, this equates to an extinction at about 10 (95% CRI: 3 to 19) occupied sites. There was no evidence that extinction varied substantially with any of the tested covariates. The extinction-covariate relationships were examined and their posterior distributions were only moderately above or below zero: salinity (75.8% below zero), temperature (53.1% below zero), dissolved oxygen (53.1 below zero), and estuary size (54.8% above zero, Table 2). I examined two additional covariates, the change in salinity at a site from 2016 to 2017 and the absolute value of that change as potential covariates of extinction, but inclusion of these covariates prevented the model from converging, thus they were excluded. Results for the naïve, rangewide extinction rate between the two years was 0.29. This is the equivalent of 24 sites occupied in 2016 that were unoccupied 2017. This extinction rate and number of extirpated sites were higher than the model estimates (CRI: 3 to 19, Figure 5).

The mean of the posterior distribution of the probability of colonization of an unoccupied site was 0.09 (95% CRI: 0.02 to 0.16), or about 8 (95% CRI: 2 to 13) sites being colonized. There was strong posterior support (88.6% below zero) for the relationship between probability of colonization within a sub-unit (ω) and the distance between neighboring sites; as the distance between sites increases, colonization within sub-units decreases (Figure 6). Colonization rates for unoccupied sites increased with the number of occupied neighbors (Figure 7); the increase in probability per occupied neighbor is a function of the equation $1 - (1 - x)^n$ where x is the colonization rate per neighbor (0.01), and n is the number of occupied neighbors. I explored the possibility of using the distance between sites that was rocky coast as a covariate of colonization as rocky coast is believed to be a hindrance to tidewater goby colonization, but its inclusion prevented the model from converging properly, causing poor model fit. The naïve conditional rangewide colonization rate (0.18) was biased high when compared to the model estimate of colonization; this naïve rate equates to 19 of the unoccupied sites in 2016 being occupied by 2017, but this naïve estimate was outside of the model's credible interval of the predicted number of colonizations (Table 3).

Recovery Units

The probability of change in occupancy between years was relatively low in most recovery units: four out of the six units had no significant change in occupancy while one recovery unit (Los Angeles/Ventura Units) experienced a significant increase and one (Conception Unit) experienced a significant decrease (Appendix B Figure 12, Table 3). There was no evidence of change in occupancy in the North Coast, Central Coast, and South Coast Units (Appendix B Figure 11, Table 3). The posterior distribution of the probability of change in occupancy in Greater Bay Area was negatively skewed with 75.8% of the posterior distribution being less than zero, suggesting a possible decline in the number of occupied sites between the two years of sampling. The Conception unit saw a decrease in occupancy probability (95% CRI: -0.30 to 0.15). Los Angeles/Ventura recovery unit exhibited a significant increase in the probability of occupancy (95% CRI: -0.29to 0.44). Model estimates of change in occupancy within recovery units were not indicative of extinction and colonization dynamics within recovery units (Appendix B Figure 11, Table 3). All recovery units, regardless of the amount of occupancy change within the unit, experienced some level of extinction and colonization. There was no geographic pattern or gradient seen in the model estimates of metapopulation dynamics between the recovery units across the tidewater goby range. Naïve estimates of extinction and colonization within the recovery units were biased high in most recover units (Appendix B Figure 13 and Figure 14).

Environmental DNA

Model estimates of detection of eDNA in a water sample (θ) and in qPCR replicates (ρ) were informative of probability of detecting tidewater goby using eDNA methods, as well as which covariates affected those detection probabilities (Figure 8). Specifically, detection in a water sample was affected by salinity, dissolved oxygen, and if a site was open to the tide, whereas turbidity and temperature did not significantly influence eDNA detection in water samples (Figure 8). Given the presence of tidewater goby eDNA at a site, the probability of detecting tidewater goby eDNA in a water sample as 0.76 (95% CRI: 0.66-0.84). The posterior distribution of the covariate dissolved oxygen strongly supports the conclusion that detection in a water sample increases with dissolved oxygen (94.4% above zero; Figures 8 and 9). There was strong evidence, 100% of the posterior distribution below zero, that presence of tide at a site reduced the probability of detection in a water sample (Figures 8 and 9). Likewise, there was strong evidence (98.9% below zero) that detection in a water sample decreased significantly with increasing salinity; an increase in salinity of 20 ppt results in a decrease in the probability of detection in a water sample of between 0.11 and 0.12 depending on whether a site is open to the tide or not (Figure 10). There was no evidence of effect of turbidity (60.6% above zero) or temperature (67.8% below zero) on detection in water samples. Detection in a qPCR replicate given presence in a water sample (ρ) was strongly impacted by salinity. Given the presence of tidewater goby eDNA in a water sample, the probability of qPCR detection was 0.59 (95% CRI: 0.56-0.63); there was strong evidence, that this probability decreased with increasing salinity (100% of the posterior distribution being less than zero; Figure 11).

DISCUSSION

Occupancy and Dynamics

Model results indicated that rangewide occupancy was stable across the two survey years, and that the balanced rates of extinction and colonization led to no significant loss or gain in the number of occupied sites. This suggests that for the single transition analyzed here, tidewater goby were at a dynamic equilibrium. For a species of conservation concern, like tidewater goby, these results may be considered promising as they show no sign of decline between these two years, but these results must be considered in light of the fact that they are limited in temporal scope, representing only two years of data and capturing only a single transition; I do not expect that the occupancy, extinction, and colonization rates found here would continue beyond the time period in which they were found.

The two years in which field work for this project occurred coincided with a substantial climatic transition in California; 2016 and several years before it are considered dry or drought years, with 2011 to 2016 being the driest years on record for California. In contrast, the winter of 2016/2017 was exceptionally wet, breaking rainfall records across the state (Wang et al. 2017). It has been predicted that tidewater goby occupancy would be negatively impacted with drought conditions causing site to dry whereas increased rainfall and river flow might promote increased colonization (Lafferty 1999a, 1999b). Based on these observations, it might be hypothesized that over my study period I would find high colonization rates and low extinction rates rather than the more comparable rates found. Transient time, or the time it takes for population dynamics to return to equilibrium following a perturbation, is typically longer when perturbations are strong, species are close to their persistence threshold, species have slow turnover, or when a site network is composed of only a few important sites (Ovaskainen and Hanski 2002). If the prolonged drought conditions represent a significant perturbation to tidewater goby metapopulation equilibrium, it is unlikely that the relative stability found in my results represents rangewide stability that could be expected to continue into the future. It is more likely that rangewide, tidewater goby dynamics are recovering from a perturbation and not yet fully stabilized, and so their extinction and colonization rates are likely to vary in future years.

The rangewide spatial coverage of this research has allowed us to examine metapopulation dynamics within the recovery units delineated by the tidewater goby recovery plan (USFWS 2005). Occupancy, extinction, and colonization for the recovery units were variable and tended not to reflect the rangewide model estimates precisely, suggesting minor variation in dynamics by region or recovery unit. While occupancy, extinction, and colonization varied across recovery units, they did not appear to follow any geographic cline or environmental gradient along the California coast. However, my results emphasize the importance of the spatial structure of localized patches; the probability of colonization was higher within sub-units when occupied and unoccupied sites are closely spaced. Lafferty (1999b) suggested that tidewater goby might exist in a core-satellite structure and my results appear to support this hypothesis; colonizing individuals are more likely to originate from occupied neighbors, as opposed to distantly located occupied sites.

Covariates

The results from this analysis support the assertion of Armstrong (2005) that incorporation of both the metapopulation paradigm and the habitat paradigm can provide robust and detailed results that could be of more use to managers than studies conducted under only one of these paradigms. Within the metapopulation paradigm, the results of this study outline the importance of occupied site density; I show that probability of colonization for an unoccupied location decreases rapidly as the distance from an occupied neighbor increases (Figure 6), and that colonization probability can be higher within a recovery sub-unit where sites are closely spaced, as opposed to a rangewide colonization process. These results are indicative of the well-supported premise in metapopulation ecology that colonization is higher between more closely located habitat sites as opposed to isolated sites (Hanski 1999) and suggest metapopulation resiliency in areas where occupied sites are proximally located. The results showing a positive correlation between tidewater goby occupancy and increased vegetation at a site fall under Armstrong's (2005) habitat paradigm. Past research on tidewater goby has repeatedly found a positive association between tidewater goby occupancy and the presence of aquatic vegetation (McGourty 2006, Swenson 1999, Worcester 1992).

The failure to find a significant relationship between site level environmental covariates (salinity, dissolved oxygen, and temperature) and extinction might be attributed to the limited scope of the environmental data used in this analysis. The habitats surveyed in this study are dynamic systems where environmental conditions can be subject to hourly, daily, and seasonal fluctuations. The two surveys used here measured each of these covariates once per water sample collected at a site, and most sites had all water samples from a single year collected in a single day. Thus, the recorded environmental conditions did not capture data that might be informative of a metapopulation dynamic process that occurs at an annual scale. Use of covariates that summarized the annual trends in these conditions at each site might have yielded significant results. Unfortunately, most of the sites surveyed here are small, poorly studied systems where longer term or continuous environmental monitoring data is not available.

Imperfect detection

The results of the dynamic multiscale occupancy model reinforce the findings of previous research emphasizing that failure to account for non-detection results in biased estimates of metapopulation dynamics (MacKenzie et al. 2002, MacKenzie 2003, Moilanen 2002). Rangewide, naïve occupancy estimates were consistently low compared to model estimates. At both the rangewide and recovery unit scale, naïve estimates of extinction and colonization were consistently biased high compared to model estimates, indicating extinction and colonization would be overestimated unless non-detection errors are unaccounted for, consistent with the findings of Moilanen (2002).

In instances where the informed and naïve estimates of extinction or colonization were similar, naïve estimates were zero and model estimates were low. This occurred in two recovery units: Los Angeles/Ventura and Conception. The data from Los Angeles/Ventura show no extinctions between the two years surveyed, and there was no evidence of colonizations in the Conception recovery unit. Because of this, the naïve estimates for these recovery units cannot be higher than zero. But, by accounting for non-detections in the analysis, the model is able to detect a probability of extinction and colonization in both of these recovery units. Once accounted for, detection probability highlights the possibility of occupancy at sites where target species were not detected, as well as the possibility of extinctions or colonizations in sites.

Three of the six recovery units (North Coast, Central Coast, and South Coast) had model estimates of occupancy change (from 2016 to 2017) similar to their naïve estimates. In these three recovery units, the number of occupied sites in the survey data was the same for both 2016 and 2017. While the change in proportion of occupied sites from one year to the next can be informative in an analysis of metapopulation dynamics, this particular metric may not convey the details in site occupancy changes from one season to the next. For instance, both the naïve estimate and the model results may similarly indicate relatively low changes in the proportion of occupied sites between the two seasons, but this may not be reflective of the extinction and colonization dynamics. This phenomenon is evident in these three units where both model and naïve estimates indicate that site occupancy was the same or similar between 2016 and 2017. Examination of the extinction and colonization estimates within these recovery units demonstrates that each of these recovery units experienced at least one extinction, colonization, or multiple instances of both. The loss of resolution when examining interannual occupancy on its own could lead to extinctions and colonization dynamics going unnoticed. This may be of particular concern for an endangered species, or any species managed under a metapopulation paradigm, where the density of occupied sites is of particular importance.

Implications for eDNA surveys

Model results of detection probabilities in water samples and qPCR replicates are useful in informing future surveys of how environmental conditions might affect their ability to detect eDNA. The results from this analysis highlight the need to incorporate some previous knowledge about a study system into sampling design; I show that detection probability in water samples decreases at sites that are open to the tide and at sites with higher salinity. Future surveys may need to account for tidal conditions and site level salinity when determining how many water samples need to be collected at a site in order to consider it thoroughly surveyed. Using the detection probability from these results ($\hat{\theta} = 0.76$), before accounting for salinity, based on the equation $1 - (1 - \hat{\theta})^n = 0.95$, three water samples ($n$) need to be collected to have a detection probability greater than 0.95. Factoring in salinity and presence of tide would increase the number of water samples needed to reach the same level of detection. Detection in qPCR replicates yields similar results. Using the same formula, $1 - (1 - \hat{\rho})^n = 0.95$, before accounting for salinity, given eDNA presence in a water sample and a $\hat{\rho}$ value of 0.593, there needs to be at least four qPCR replicates to have greater than a 95% chance of detection. Again, this number could increase as salinity increases due to increase in $\hat{\rho}$

Management Implications

The results from the research suggest that there is no latitudinal gradient in tidewater goby metapopulation dynamics. The lack of extinction and colonization signal in some northern population segments noted by Kinziger et al. (2015) may be a product of the core-satellite population structure of tidewater goby; genetic samples used Kinziger et al. (2015) may have originated from stable core populations, as opposed to less stable satellite populations, and so would offer no evidence of recent founder events or bottlenecks. Future genetic work may offer some insight into tidewater goby metapopulation dynamics by targeting recently re-colonized area for which historic samples exist.

This research demonstrates several tools that would be useful in future tidewater goby monitoring and conservation. Environmental DNA survey techniques and dynamic multiscale occupancy modeling could serve as a foundation for conduction a metapopulation viability analysis as required by the species recovery plan. The colonization and extinction rates found here represent metapopulation dynamics for two years of data and one transition and are unlikely to characterize future transitions; however, if the rates found here remained consistent these estimates could be extrapolated across longer time scales using the formula $1-(1-x)^n$ where x is the rate in question, extinction or colonization, and n is the number of years. Examining the dynamic rates found here across a decadal scale shows that individual occupied sites may have a high probability of extinction, 0.67, and that unoccupied sites have a high probability of colonization, 0.59. At the 100-year mark, there is 99.9% probability of both colonization of unoccupied sites and extinction of occupied sites. These simplistic estimates do not take into account the presumed core-satellite structure, assume that all sites have similar extinction or colonization probabilities, and come from data that is limited in scope; with further monitoring and the incorporation of additional years of data, these probabilities could be tailored to and predicted for sites within recovery units and offer a higher resolution view of the metapopulation viability within recovery units. Additionally, inclusion of data from more time periods with varying climate and environmental conditions might offer insight into the environmental drivers of extinction and colonization not found here, and could act as a basis for modeling tidewater goby metapopulation dynamics into the future as functions of climate change and larger scale weather patterns such as the El Ninõ/Southern Oscillation or the Pacific Decadal Oscillation.

The two relationships found here, between (1) occupancy and vegetation and (2) colonization and distance between sites, highlight the underpinnings of Armstrong’s (2005) call for incorporation of metapopulation paradigms and habitat paradigms into management and conservation. My results suggest that in order to best conserve tidewater goby, management may need to incorporate themes of both the metapopulation paradigm; (e.g., distance between sites) as well as the habitat paradigm (e.g., quality habitat with presence of aquatic vegetation). For example, restoration efforts that include consideration of aquatic vegetation would have little impact on tidewater goby populations if colonization rates for the restored habitat are low due to large gaps of unoccupied locations along the coast. Likewise, increasing the number of occupied sites in those gaps of unoccupied locations by planting tidewater goby may not lead to permanent occupancy if the habitat is not suitable.

In this study, I modeled two years of eDNA occupancy data using a dynamic multiscale occupancy model that explicitly accounts for non-detection at three levels: site, sample, and in qPCR detection. The model developed for this analysis was used to examine the probabilities of occupancy, extinction, and colonization of tidewater goby at two scales: rangewide and in distinct management units. Additionally, the model allowed testing of spatial and environmental covariate effects on occupancy, extinction, and colonization, as well as covariate effects on eDNA detection at both the water sample and qPCR replicate level. This study design and associated dynamic multiscale occupancy model are widely applicable and likely useful in situations where eDNA survey techniques or a hierarchical sampling design are an option. The increasing use of eDNA will likely provide opportunities to apply this model over increasing temporal scales in an increasing number of ecological settings to aid in monitoring and conservation of endangered species like the tidewater goby.

FIGURES

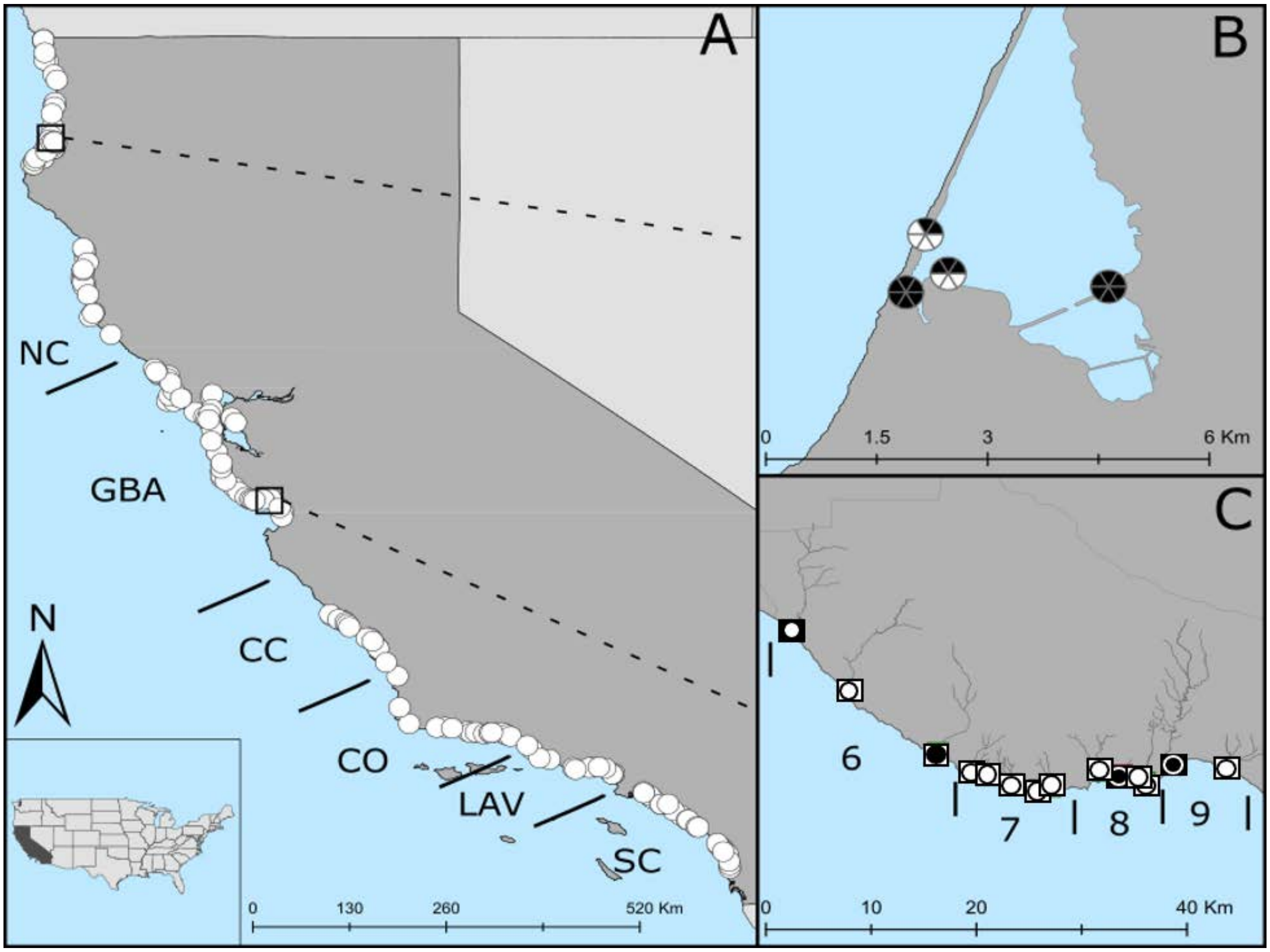

Figure 1 Tidewater goby sampling locations. The hierarchical nature of eDNA sampling is demonstrated by the transition from panel A, showing the 190 sites surveyed for tidewater goby along the 1350 km California coast, to panel B where each pie represents the location of individual water samples collected at a site, in this case Big Lagoon, Humboldt County, and pie slices that represent replicate qPCR reactions conducted for each water sample. Positive qPCR replicates are represented by filled in pie slices. Panel A displays the tidewater goby recovery units (NC = north coast, GBA = greater bay area, CC = central coast, CO = Conception, LAV = Los Angeles/Ventura, SC = south coast). The SC Recovery Unit is exclusively occupied by the southern tidewater goby *Eucyclogobius kristinae* and the other five Recovery Units are occupied by northern tidewater goby *E. newberryi*. The temporal nature of this sampling is represented in panel C, showing two years of occupancy for a section of the Central Coast (CC) Recovery Unit. The black lines separate four recovery sub-units (6 - 9) as an example of the neighborhood structure

used in the dynamic multiscale occupancy model. Square symbols represent sites surveyed in 2016 and the circles represent sites surveyed in 2017. White indicates a detection in at least one qPCR reaction at a site whereas indicates non-detection at a site at all water samples and qPCR replicates.

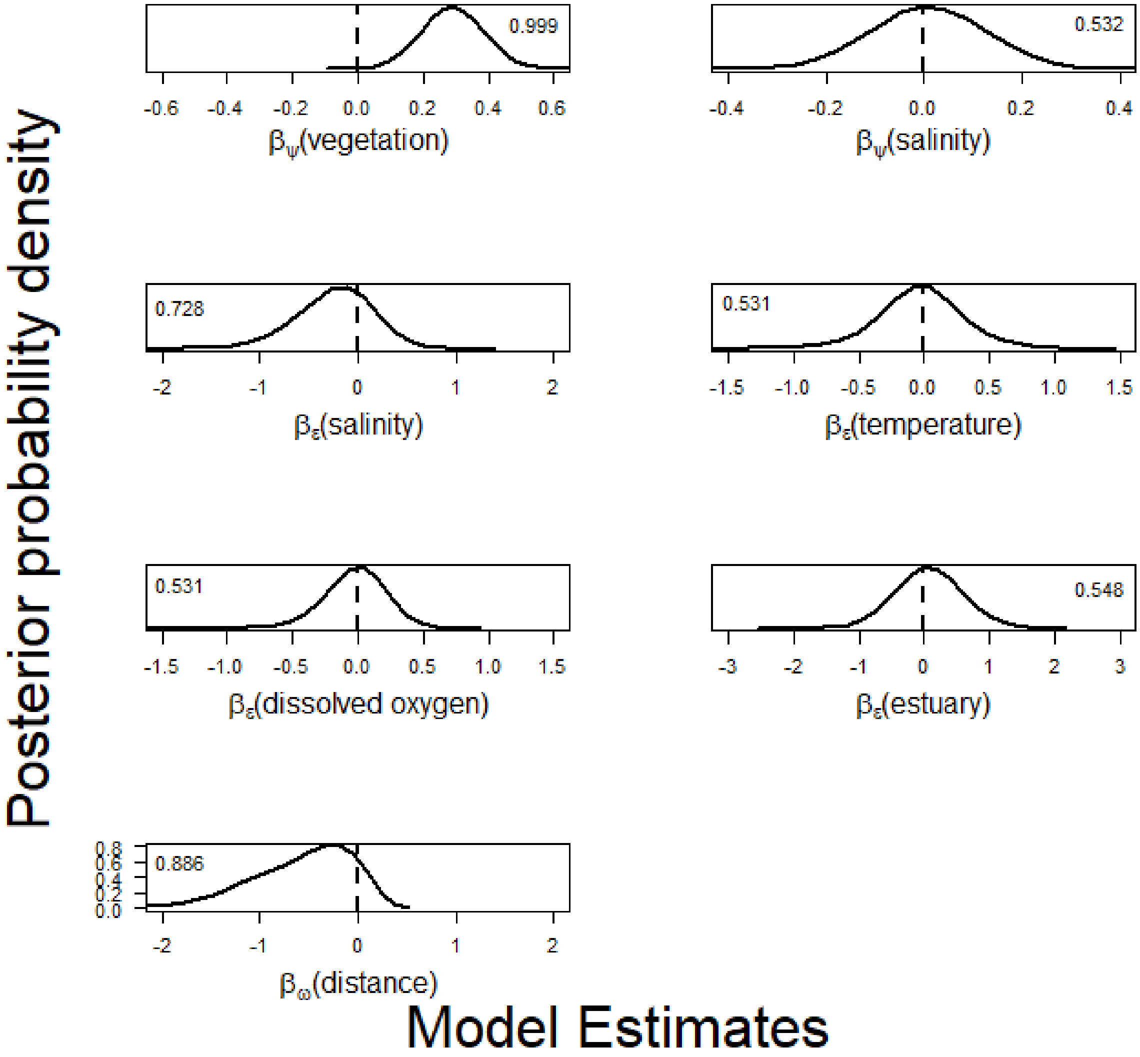

Figure 2 **Posterior probability densities for the estimated covariate effects (β) on occupancy (ψ), extinction (ε), and colonization by a neighbor (ω). Result are from an analysis of two years (2016 and 2017) of rangewide tidewater goby *Eucyclogobius spp.* eDNA occupancy data using a dynamic multiscale occupancy. Numerical values represent the proportion of the posterior distribution greater or less than zero.**

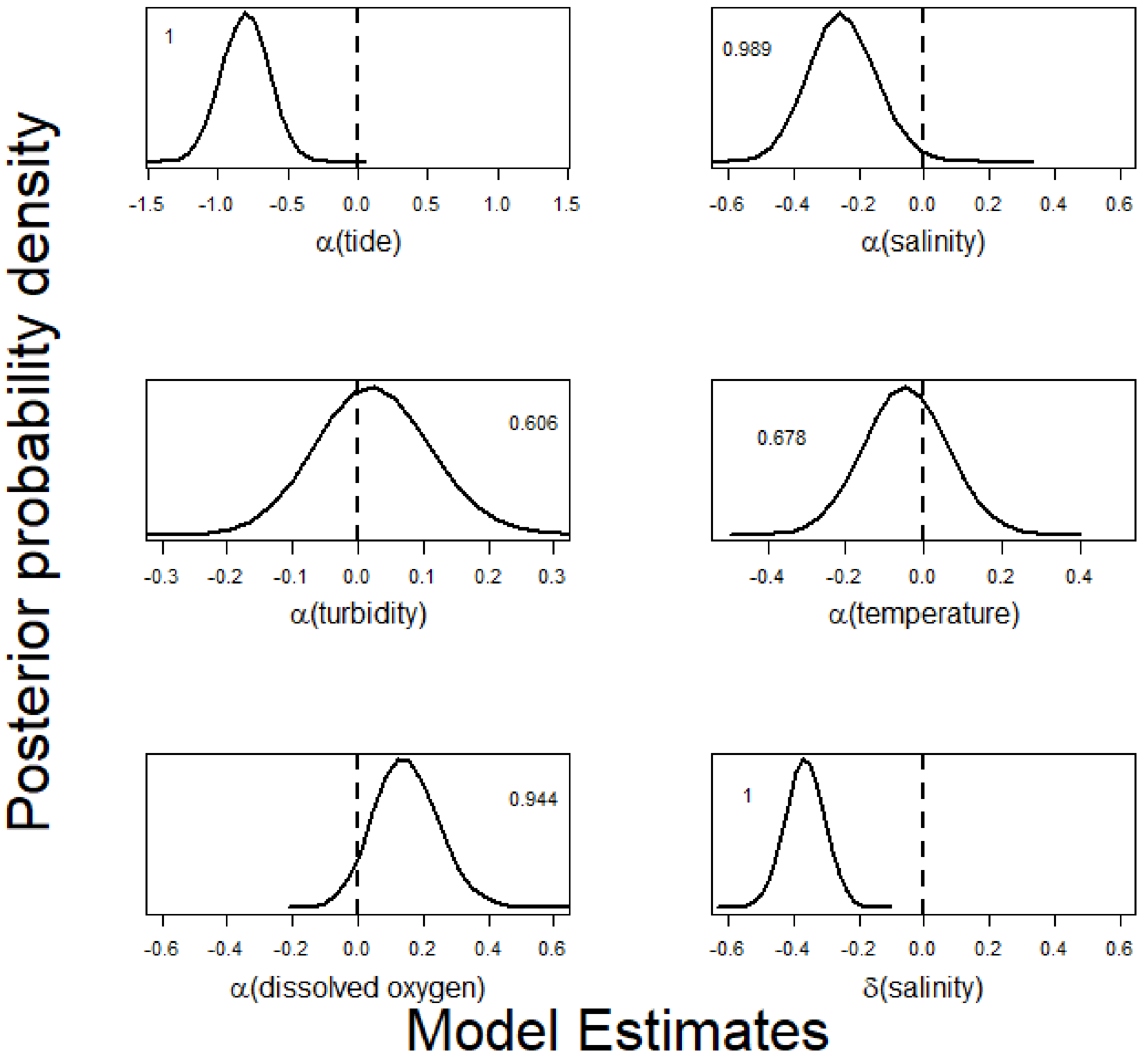

Figure 3 Posterior probability densities for covariates of detection in a water sample (α) and detection in qPCR replicates (δ) that were examined as part of an analysis of two years (2016 and 2017) of rangewide eDNA occupancy data of tidewater goby *Eucyclogobius spp*. Data was analyzed using a dynamic multiscale occupancy model that accounts for non-detection at the site, water sample, and qPCR replicate level. Numerical values represent the proportion of the posterior distribution greater or less than zero (the dashed line).

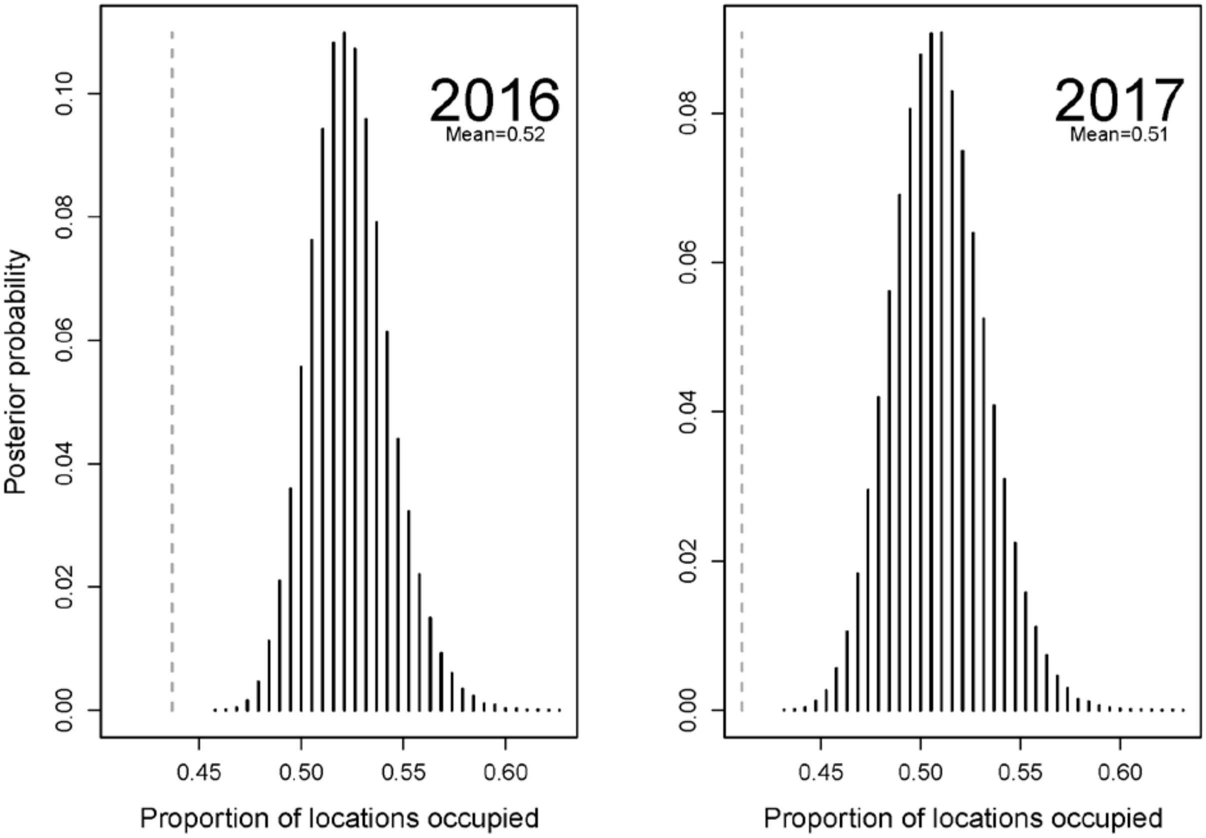

Figure 4 Posterior distribution of tidewater goby *Eucyclogobius spp.* rangewide occupancy. The proportion of occupied sites were estimated using dynamic multiscale occupancy model to analyze eDNA survey data for 2016 (left) and 2017 (right). The dashed line represents the naïve occupancy rate.

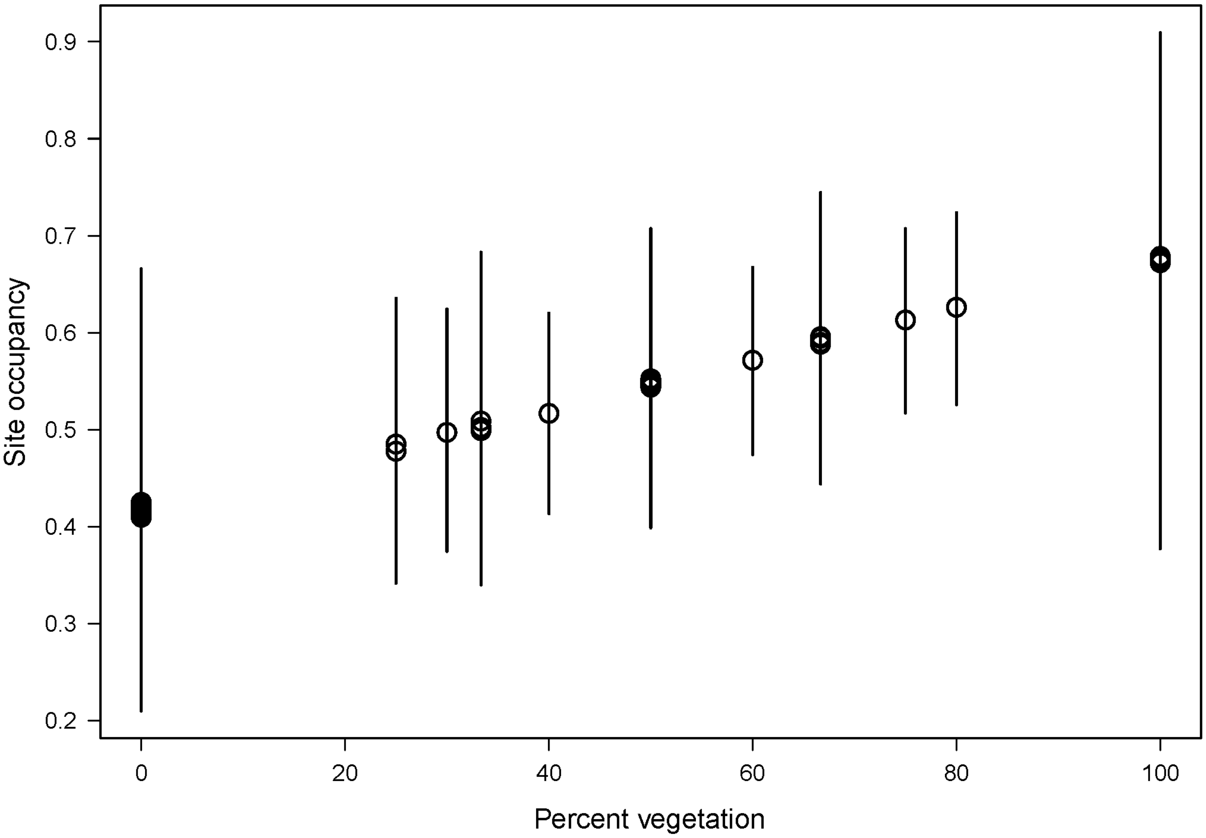

Figure 5 Effects of vegetation levels on site occupancy. The presence of aquatic vegetation was tested as a covariate of tidewater goby *Eucyclogobius spp.* occupancy at a site (ψ) in an analysis of two years, 2016 and 2017, of eDNA survey data using a multiscale occupancy model. The *x*-axis is the proportion of water sample locations at a site where aquatic vegetation was present. The bars represent the 95% credible interval of occupancy for the sites displayed.

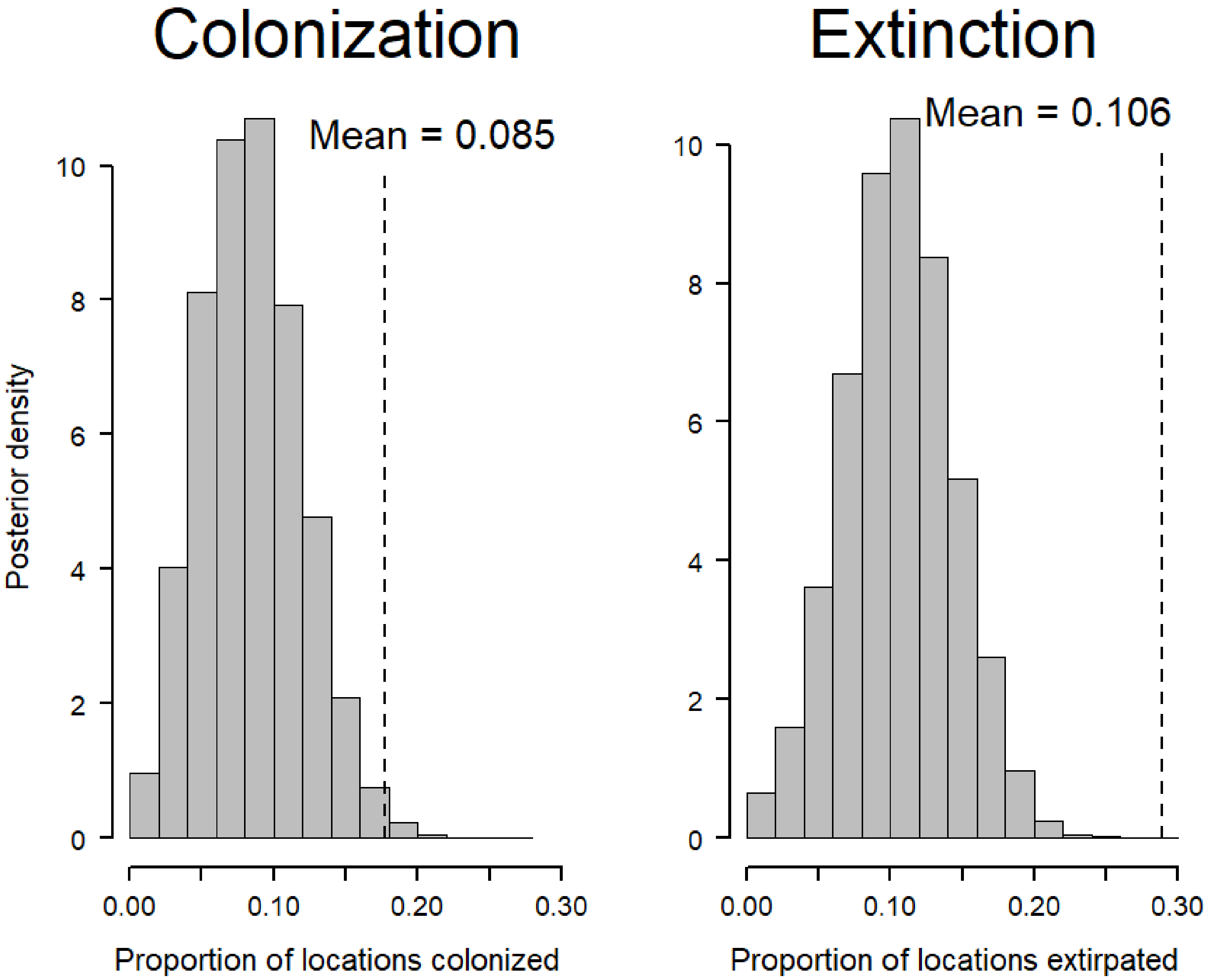

Figure 6 Posterior density distribution of rangewide colonization of unoccupied sites (left) extinction of occupied sites (right) for tidewater goby *Eucyclogobius spp.* between 2016 and 2017. Rangewide eDNA survey data was analyzed using a dynamic multiscale occupancy model. Super-imposed mean values are for their respective distributions; the dashed lines represent the naïve estimates that do not account for non-detection.

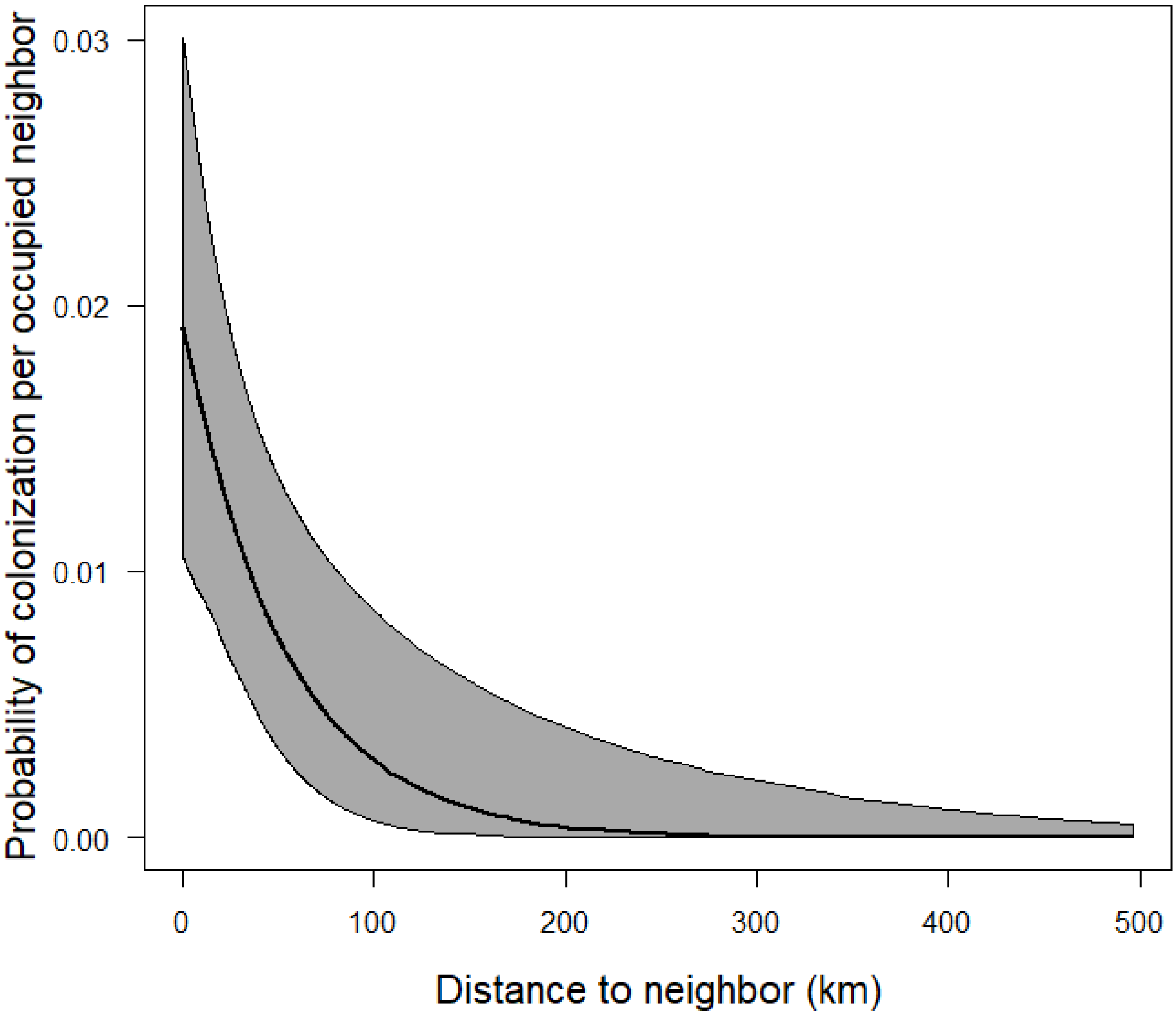

Figure 7 Plot of the shoreline distance between neighbors versus the probability of colonization of an unoccupied site from an occupied neighbor (ω). Analysis was done using a dynamic multiscale occupancy model to analyze two years, 2016 and 2017, of tidewater goby *Eucyclogobius spp.* eDNA occupancy data. The shaded area represents the 95% credible interval.

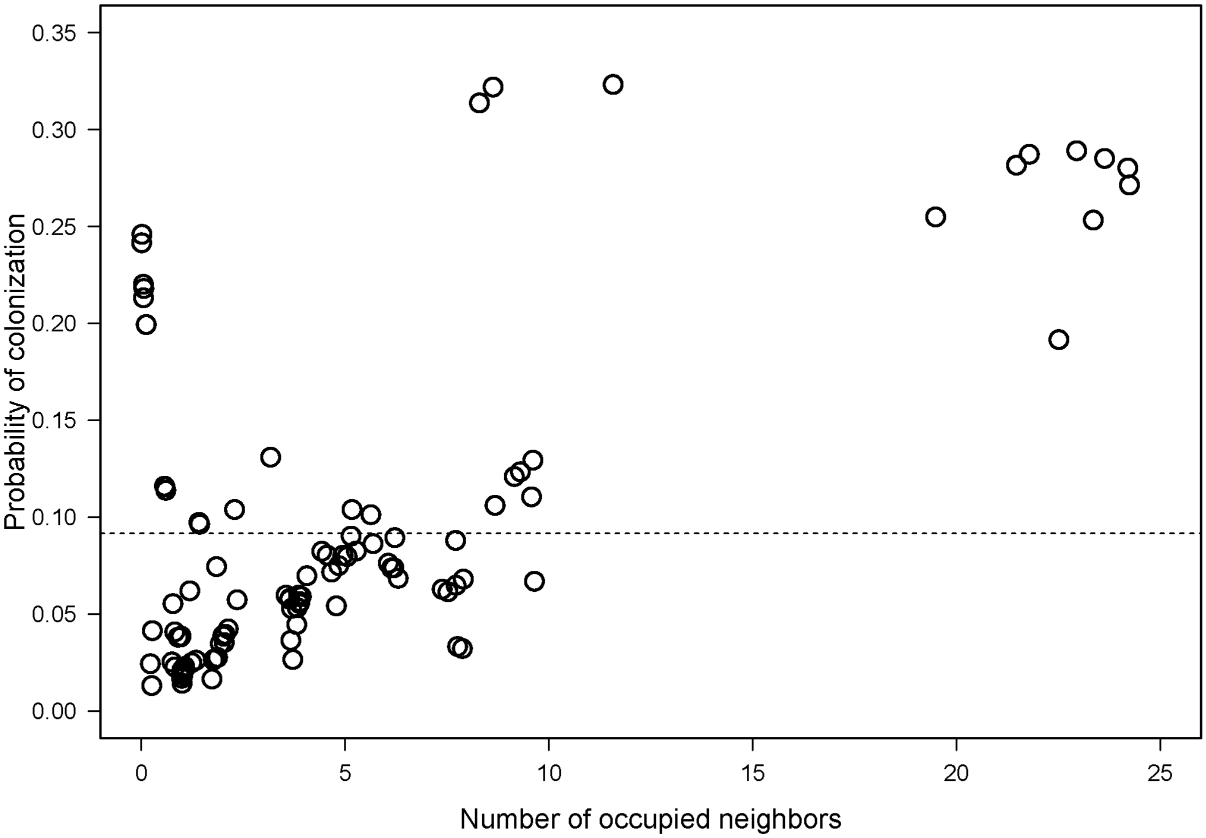

Figure 8 Probability of colonization of unoccupied sites versus the number of occupied neighbors. Colonization at sites with occupied neighbors is only a process of localized, neighborhood colonization (ω), while sites with no occupied neighbors are subject to the rangewide colonization rate (η). Probability of colonization was calculated from parameter estimates of a dynamic multiscale occupancy model based on eDNA occupancy data of tidewater goby *Eucyclogobius spp.* collected in 2016 and 2017. The dashed grey line represents the model estimate of rangewide extinction.

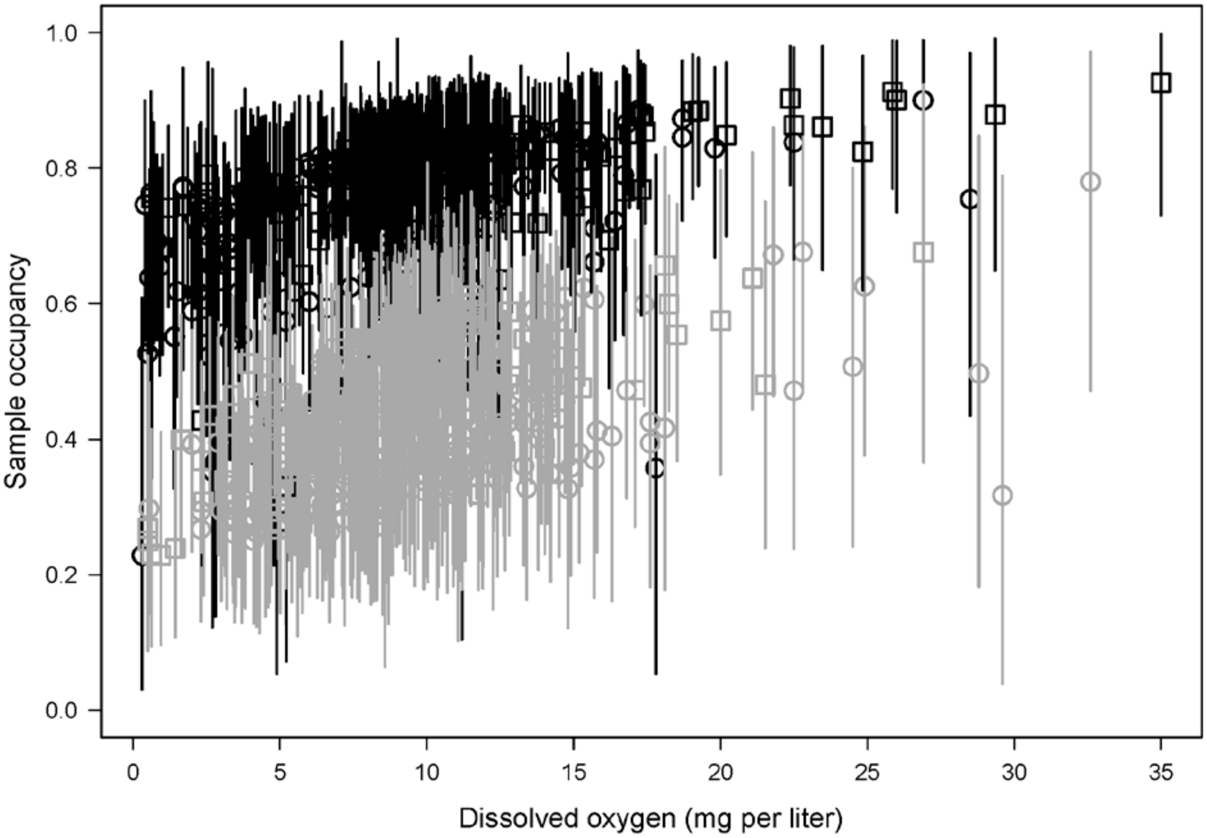

Figure 9 Effect of dissolved oxygen and tidal presence on eDNA presence (with 95% credible intervals) in a water sample. Water samples with tidal influence are in grey, water samples without tidal influence are in black. Water samples were collected in 2016 (squares) and 2017 (circles) as part of a rangewide eDNA survey of tidewater goby *Eucyclogobius spp.* and analyzed using a dynamic multiscale occupancy model.

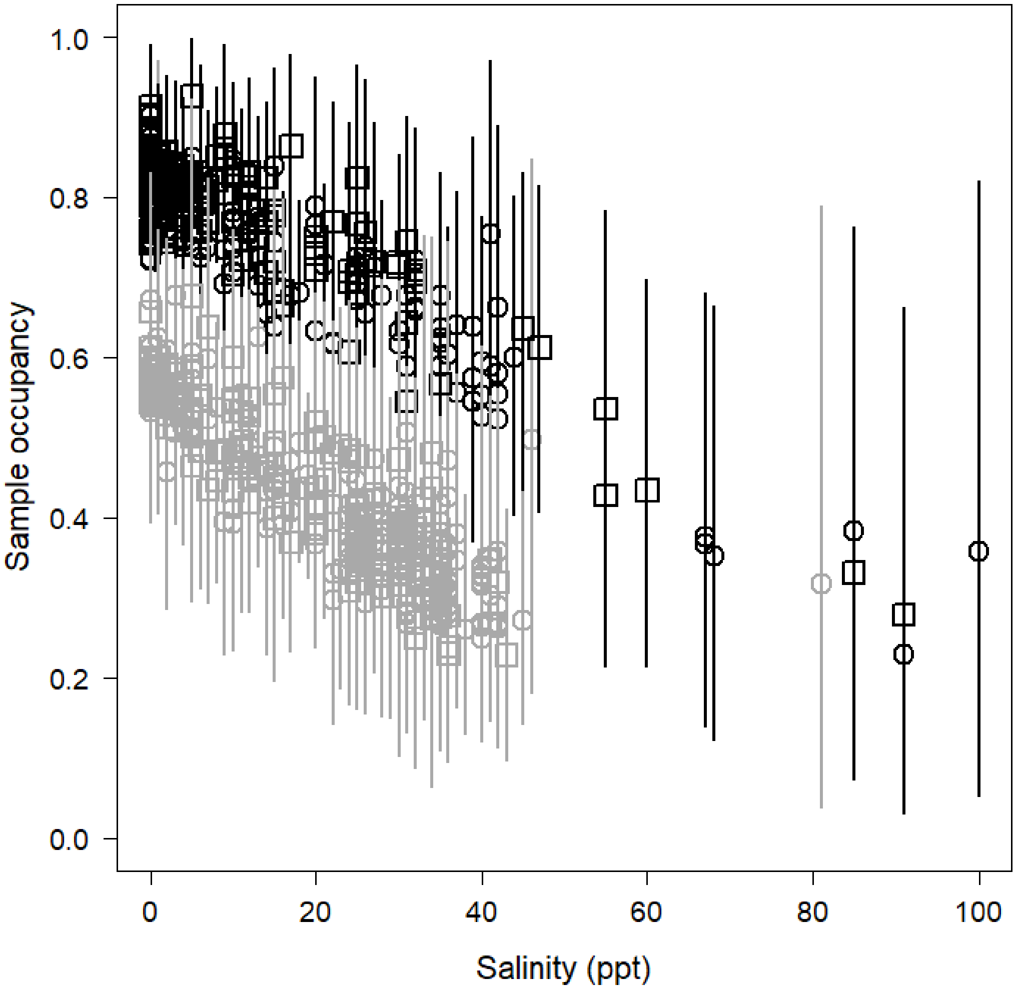

Figure 10 Effect of salinity and tidal presence on eDNA presence (with 95% credible intervals) in a water sample. Water samples with tidal presence are in grey, water samples without tidal presence are in black. Water samples were collected in 2016 (squares) and 2017 (circles) as part of a rangewide eDNA survey of tidewater goby Eucyclogobius spp. and analyzed using a dynamic multiscale occupancy model.

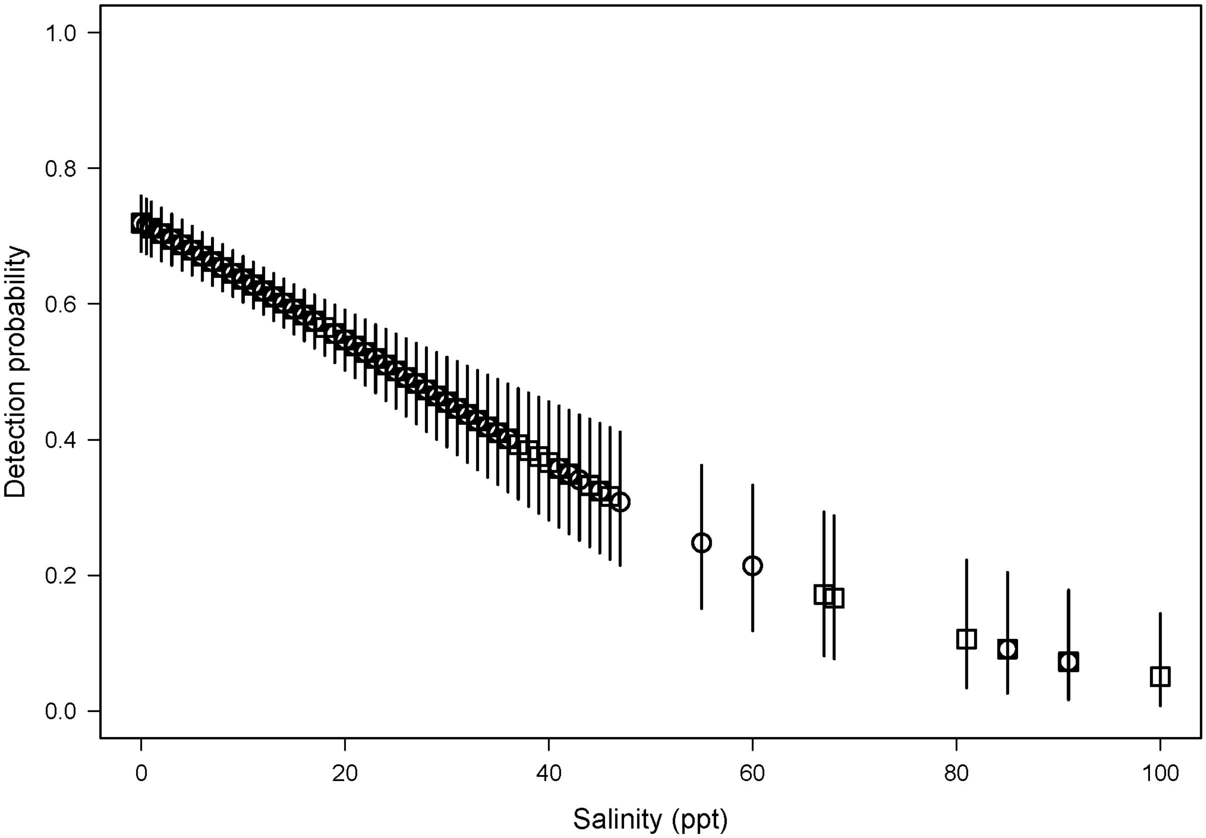

Figure 11 Estimates of probability of detection of eDNA (with 95% credible intervals) in qPCR replicates from water samples of with varying salinities. Water samples were collected in 2016 (squares) and 2017 (circles) as part of a rangewide eDNA survey of tidewater goby *Eucyclogobius spp.* Data were analyzed using a dynamic multiscale occupancy model that accounts for non-detection at site, water sample, and qPCR replicate levels.

Table 1 The dynamic multiscale model allows for the incorporation of covariates that may drive patterns in tested parameters. Here I present the covariates tested while analyzing tidewater goby *Eucyclogobius spp.* occupancy data from two consecutive years (2016-2017) of rangewide eDNA surveys, and their respective literature-based hypothesis.

Parameter	Covariate	Hypothesis
Presence (ψ)	Salinity	Frequent occurrence of tidewater goby in habitats with less than full strength sea water suggests habitats with lower salinity would have increased occupancy (Swenson 1997, Swift et al. 1989, Sutter 2017)
	Vegetation	Habitats with increased vegetation provide cover resulting in higher occupancy (McGourty 2008).
Extinction (ε)	Estuary Size (binned)	Larger (>1.0 ha) estuaries buffer against extinction by providing more suitable habitat and reduced chance of desiccation (Lafferty1999).
	Salinity	Occupied sites with salinity outside of the tolerance range are more likely to undergo extinction.
	Temperature	Occupied sites with temperature outside of the tolerance range are more likely to undergo extinction.
	Dissolved Oxygen	Occupied sites with dissolved oxygen outside of the tolerance range are more likely to undergo extinction.
Colonization by neighbor (ω)	Pairwise distance between neighbor	Probability of colonization by a neighbor should decrease as the distance between neighbors increases (Earl et al. 2010).
Sample occupancy (α)	Salinity	Increased salinity would result in lower eDNA availability in a water sample due to degradation (Sutter 2017).
	Turbidity	Suspended sediment can increase eDNA persistence and therefore may increase availability of eDNA in a water sample (Barnes et al. 2014).
	Tidal Presence	Tidal presence decreases eDNA availability at a site (Schmelzle 2015).
	Temperature	Increased degradation due to warmer temperatures would result in lower availability of eDNA for water samples (Barnes et al. 2014).
	Dissolved Oxygen	Increased degradation of genetic material under higher dissolved oxygen will result in lower availability in a water sample (Weltz et al. 2017).
qPCR detectability (δ)	Salinity	Increased salinity would result in lower qPCR detection due to inhibition of amplification in qPCR(Sutter 2017).

Table 2 Results, presented as beta values (β_ψ – posterior distribution occupancy and its covariates, β_ε - posterior distribution extinction and its covariates, β_η - posterior distribution colonization and its covariates, β_ω - posterior distribution neighborhood colonization and its covariates, α - posterior distribution of eDNA detection in a water sample and its covariates, δ - posterior distribution eDNA detection in a qPCR replicate and its covariates) from the dynamic multiscale model of rangewide tidewater goby occupancy and extinction and colonization dynamics from two consecutive years (2016-2017) eDNA surveys. The bounds of the 95% credible interval of these posterior distributions is represented in the columns labeled "2.5%" and "97.5%." The proportion of these distributions below or above zero are located in the "<0" or ">0" columns, respectively. Significant results are denoted by an asterisk (*).

Parameter	Mean	2.5%	97.5%	<0	> 0
β_ψ (intercept)	0.165	-0.143	0.852	-	-
β_ψ (vegetation)*	0.292	0.086	0.487	0.001	0.999
β_ψ (salinity)	0.072	-0.233	0.728	0.469	0.531
β_ε (intercept)	-1.538	-4.596	-0.590	-	-
β_ε (salinity)	-0.236	-1.427	0.458	0.728	0.272
β_ε (temperature)	0.001	-0.680	0.864	0.530	0.470
β_ε (dissolved oxygen)	-0.031	-0.536	0.406	0.530	0.470
β_ε (estuary size)	0.137	-0.895	1.193	0.452	0.548
β_η (intercept)	0.089	-1.855	2.202	-	-
β_ω (intercept)	-2.302	-3.865	-1.733	-	-
β_ω (distance)*	-0.348	-1.493	0.237	0.888	0.112
α (intercept)	0.703	0.420	0.988	-	-
α (tide)*	-0.826	-1.179	-0.482	1.000	0.000
α (salinity)*	-0.261	-0.468	-0.039	0.989	0.011
α (turbidity)	-0.001	-0.164	0.169	0.395	0.605
α (temperature)	-0.058	-0.266	0.150	0.678	0.322
α (dissolved oxygen)*	0.159	-0.023	0.361	0.056	0.944
δ (intercept)	0.236	0.141	0.330	-	-
δ (salinity)*	-0.365	-0.483	-0.248	0.000	1.000

Table 3 Comparison of naïve and model based estimates of colonization and extinction rates for tidewater goby *Eucyclogobius spp.* across their range, separated by Recovery Units (NC = north coast, GBA = greater bay area, CC = central coast, CO = Conception, LAV = Los Angeles/Ventura, SC = south coast) as outline in the Tidewater goby Recovery Plan (2005 USFWS). Occupancy change represents the change in poropotion of occupied sites between 2016 and 2017.

Recovery unit	No. survey locations	Model-based estimates			Naïve estimates		
		Occupancy change	Colonize	Extinct	Occupancy change	Colonize	Extinct
NC	66	-0.004	0.078	0.116	0.000	0.159	0.32
GBA	53	-0.040	0.094	0.138	-0.075	0.214	0.400
CC	17	-0.008	0.108	0.035	0.000	0.250	0.077
CO	19	-0.088	0.013	0.114	-0.211	0.000	0.267
LAV	13	0.088	0.183	0.016	0.231	0.333	0.000
SC	22	0.010	0.065	0.132	0.000	0.111	0.500

LITERATURE CITED

Akçakaya, H. R., and J. L. Atwood. 1997. A habitat-based metapopulation model of the California Gnatcatcher. Conservation Biology 11(2):422–434.

Anthes, N., T. Fartmann, G. Hermann, and G. Kaule. 2003. Combining larval habitat quality and metapopulation structure – the key for successful management of pre-alpine *Euphydryas aurinia* colonies. Journal of Insect Conservation 7(3):175–185.

Armstrong, D. P. 2005. Integrating the metapopulation and habitat paradigms for understanding broad-scale declines of species. Conservation Biology 19(5):1402–1410.

Barbraud, C., J. D. Nichols, J. E. Hines, and H. Hafner. 2003. Estimating rates of local extinction and colonization in colonial species and an extension to the metapopulation and community levels. Oikos 101(1):113–126.

Bohmann, K., A. Evans, M. T. P. Gilbert, G. R. Carvalho, S. Creer, M. Knapp, D. W. Yu, and M. de Bruyn. 2014. Environmental DNA for wildlife biology and biodiversity monitoring. Trends in Ecology & Evolution 29(6):358–367.

Boussarie, G., J. Bakker, O. S. Wangensteen, S. Mariani, L. Bonnin, J.-B. Juhel, J. J. Kiszka, M. Kulbicki, S. Manel, W. D. Robbins, L. Vigliola, and D. Mouillot. 2018. Environmental DNA illuminates the dark diversity of sharks. Science Advances 4(5):eaap9661.

Dejean, T., A. Valentini, C. Miquel, P. Taberlet, E. Bellemain, and C. Miaud. 2012. Improved detection of an alien invasive species through environmental DNA barcoding: the example of the American bullfrog *Lithobates catesbeianus*: *Alien invasive species detection using eDNA*. Journal of Applied Ecology 49(4):953–959.

Dennis, R. L. H., and H. T. Eales. 1997. Patch occupancy in *Coenonympha tullia* (Muller, 1764) (Lepidoptera: *Satyrinae*): habitat quality matters as much as patch size and isolation. Journal of Insect Conservation 1(3):167–176.

Dorazio, R. M., and R. A. Erickson. 2018. ednaoccupancy: An r package for multiscale occupancy modelling of environmental DNA data. Molecular Ecology Resources 18(2):368–380.

Beissinger, S. R., M. I. Westphal. 1998. On the use of demographic models of population viability in endangered species management. The Journal of Wildlife Management 62(3):821–841.

Betts, M. G., N. L. Rodenhouse, T. S. Sillett, P. J. Doran, R. T. Holmes. 2008. Dynamic occupancy models reveal within-breeding season movement up a habitat quality gradient by a migratory songbird. Ecography 31(5):592-600.

Broms, K. M., M. B. Hooten, D. S. Johnson, R. Altwegg, L. L. Conquest. 2016. Dynamic occupancy models for explicit colonization processes. Ecology 97(1):194-204.

Earl, D. A., K. D. Louie, C. Bardeleben, C. C. Swift, and D. K. Jacobs. 2010. Rangewide microsatellite phylogeography of the endangered tidewater goby, *Eucyclogobius newberryi* (Teleostei: *Gobiidae*), a genetically subdivided coastal fish with limited marine dispersal. Conservation Genetics 11(1):103–114.

Eiler, A., A. Löfgren, O. Hjerne, S. Nordén, and P. Saetre. 2018. Environmental DNA (eDNA) detects the pool frog (*Pelophylax lessonae*) at times when traditional monitoring methods are insensitive. Scientific Reports 8(1).

Evans, N. T., P. D. Shirey, J. G. Wieringa, A. R. Mahon, and G. A. Lamberti. 2017. Comparative cost and effort of fish distribution detection via environmental DNA analysis and electrofishing. Fisheries 42(2):90–99.

Fernández Néstor, Román Jacinto, and Delibes Miguel. 2016. Variability in primary productivity determines metapopulation dynamics. Proceedings of the Royal Society B: Biological Sciences 283(1828):20152998.

Fleishman, E., C. Ray, P. Sjögren-Gulve, C. L. Boggs, and D. D. Murphy. 2002. Assessing the roles of patch quality, area, and isolation in predicting metapopulation dynamics. Conservation Biology 16(3):706–716.

Foote, A. D., P. F. Thomsen, S. Sveegaard, M. Wahlberg, J. Kielgast, L. A. Kyhn, A. B. Salling, A. Galatius, L. Orlando, and M. T. P. Gilbert. 2012. Investigating the potential use of environmental DNA (eDNA) for genetic monitoring of marine mammals. PLOS ONE 7(8):e41781.

Franken, R. J., and D. S. Hik. 2004. Influence of habitat quality, patch size and connectivity on colonization and extinction dynamics of collared pikas *Ochotona collaris*. Journal of Animal Ecology 73(5):889–896.

Gingera, T. D., T. B. Steeves, D. A. Boguski, S. Whyard, W. Li, and M. F. Docker. 2016. Detection and identification of lampreys in Great Lakes streams using environmental DNA. Journal of Great Lakes Research 42(3):649–659.

Goldberg, C. S. 2016. Critical considerations for the application of environmental DNA methods to detect aquatic species. Ecology and Evolution 7(11):1299–1307.

Goldberg, C. S., A. Sepulveda, A. Ray, J. Baumgardt, and L. P. Waits. 2013. Environmental DNA as a new method for early detection of New Zealand mudsnails (*Potamopyrgus antipodarum*). Freshwater Science 32(3):792–800.

Gu, W., and R. K. Swihart. 2004. Absent or undetected? Effects of non-detection of species occurrence on wildlife–habitat models. Biological Conservation 116(2):195–203.

Hanski, I. 1989. Metapopulation dynamics: Does it help to have more of the same? Trends in Ecology & Evolution 4(4):113–114.

Hanski, I. 1994. A practical model of metapopulation dynamics. The Journal of Animal Ecology 63(1):151.

Hanski, I. 1998. Metapopulation dynamics. Nature 396(6706):41–49.

Hanski, I. 1999. Habitat connectivity, habitat continuity, and metapopulations in dynamic landscapes. Oikos 87(2):209–219.

Hanski, I., and M. Gilpin. 1991. Metapopulation dynamics: brief history and conceptual domain. Biological Journal of the Linnean Society 42(1–2):3–16.

Hanski, I., and D. Simberloff. 1997. 1 - The metapopulation approach, its history, conceptual domain, and dpplication to conservation. Pages 5–26 *in* I. Hanski and M. E. Gilpin, editors. Metapopulation Biology. Academic Press, San Diego.

Hastings, A., and S. Harrison. 1994. Metapopulation dynamics and genetics. Annual Review of Ecology and Systematics 25(1):167–188.

Hill, J. K., C. D. Thomas, and O. T. Lewis. 1996. Effects of habitat patch size and isolation on dispersal by hesperia comma butterflies: implications for metapopulation structure. The Journal of Animal Ecology 65(6):725.

Hunter, M. E., S. J. Oyler-McCance, R. M. Dorazio, J. A. Fike, B. J. Smith, C. T. Hunter, R. N. Reed, and K. M. Hart. 2015. Environmental DNA (eDNA) sampling improves occurrence and detection estimates of invasive Burmese pythons. PLOS ONE 10(4):e0121655.

Jerde, C. L., W. L. Chadderton, A. R. Mahon, M. A. Renshaw, J. Corush, M. L. Budny, S. Mysorekar, and D. M. Lodge. 2013. Detection of Asian carp DNA as part of a Great Lakes basin-wide surveillance program. Canadian Journal of Fisheries and Aquatic Sciences 70(4):522–526.

Jerde, C. L., A. R. Mahon, W. L. Chadderton, and D. M. Lodge. 2011. “Sight-unseen” detection of rare aquatic species using environmental DNA. Conservation Letters 4(2):150–157.

McGourty K. R,. 2006. Spawning time, fecundity, habitat utilization, and parasites of *Eucyclogobius newberryi* in Big Lagoon, Humboldt County, Ca. Master’s Thesis. Humboldt State University, Arcata, California.

Kinziger, A. P., M. Hellmair, W. T. McCraney, D. K. Jacobs, and G. Goldsmith. 2015. Temporal genetic analysis of the endangered tidewater goby: extinction–colonization dynamics or drift in isolation? Molecular Ecology 24(22):5544–5560.

Kristine Bohmann, Alice Evans, M. Thomas P. Gilbert, Gary R. Carvalho, Simon Creer, Michael Knapp, Douglas W. Yu, and Mark de Buyn. 2014. Environmental DNA for wildlife biology and biodiversity monitoring. Trends in Ecology & Evolution 29(6):358–367.

Lafferty, K. D., C. C. Swift, and R. F. Ambrose. 1999a. Postflood persistence and recolonization of endangered tidewater goby populations. North American Journal of Fisheries Management 19(2):618–622.

Lafferty, K. D., C. C. Swift, and R. F. Ambrose. 1999b. Extirpation and recolonization in a metapopulation of an endangered fish, the tidewater goby. Conservation Biology 13(6):1447–1453.

Levins, R. 1969. Some demographic and genetic consequences of environmental heterogeneity for biological control. Bulletin of the Entomological Society of America 15(3):237–240.

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(8):2200–2207.

MacKenzie, D. I., J. D. Nichols, G. B. Lachman, S. Droege, J. Andrew Royle, and C. A. Langtimm. 2002. Estimating site occupancy rates when detection probabilities are less than one. Ecology 83(8):2248–2255.

Mackenzie, D. I., and J. A. Royle. 2005. Designing occupancy studies: general advice and allocating survey effort. Journal of Applied Ecology 42(6):1105–1114.

Marsh, D. M., and P. C. Trenham. 2001. Metapopulation dynamics and amphibian conservation. Conservation Biology 15(1):10.

Moilanen, A. 2002. Implications of empirical data quality to metapopulation model parameter estimation and application. Oikos 96(3):516–530.

Mordecai, R. S., B. J. Mattsson, C. J. Tzilkowski, R. J. Cooper. 2011. Addressing challenges when studying mobile episodic species: hierarchical Bayes estimation of occupancy and use. Journal of Applied Ecology 48:56-66.

Nichols, J. D., L. L. Bailey, A. F. O. Jr, N. W. Talancy, E. H. C. Grant, A. T. Gilbert, E. M. Annand, T. P. Husband, and J. E. Hines. 2008. Multi-scale occupancy estimation and modelling using multiple detection methods. Journal of Applied Ecology 45(5):1321–1329.

Nielsen, S. E., G. McDermid, G. B. Stenhouse, and M. S. Boyce. 2010. Dynamic wildlife habitat models: Seasonal foods and mortality risk predict occupancy-abundance and habitat selection in grizzly bears. Biological Conservation 143(7):1623–1634.

Ovaskainen, O., and I. Hanski. 2002. Transient dynamics in metapopulation response to perturbation. Theoretical Population Biology 61(3):285–295.

Pilliod, D. S., C. S. Goldberg, R. S. Arkle, and L. P. Waits. 2013. Estimating occupancy and abundance of stream amphibians using environmental DNA from filtered water samples. Canadian Journal of Fisheries and Aquatic Sciences 70(8):1123–1130.

Port, J. A., J. L. O’Donnell, O. C. Romero-Maraccini, P. R. Leary, S. Y. Litvin, K. J. Nickols, K. M. Yamahara, and R. P. Kelly. 2016. Assessing vertebrate biodiversity in a kelp forest ecosystem using environmental DNA. Molecular Ecology 25(2):527–541.

Schmelzle, M. C., and A. P. Kinziger. 2016. Using occupancy modelling to compare environmental DNA to traditional field methods for regional-scale monitoring of an endangered aquatic species. Molecular Ecology Resources 16(4):895–908.

Schmidt, B. R., M. Kéry, S. Ursenbacher, O. J. Hyman, and J. P. Collins. 2013. Site occupancy models in the analysis of environmental DNA presence/absence surveys: a case study of an emerging amphibian pathogen. Methods in Ecology and Evolution 4(7):646–653.

Semlitsch, R. D. 2002. Critical elements for biologically based recovery plans of aquatic-breeding amphibians. Conservation Biology 16(3):619–629.

Semlitsch, R. D. 2008. Differentiating migration and dispersal processes for pond-breeding amphibians. The Journal of Wildlife Management 72(1):260–267.

Strickler, K. M., A. K. Fremier, and C. S. Goldberg. 2015. Quantifying effects of UV-B, temperature, and pH on eDNA degradation in aquatic microcosms. Biological Conservation 183:85–92.

Sutter, M. 2018. Rangewide tidewater goby occupancy survey using environmental DNA. Master’s Thesis, Humboldt State University, Arcata, California.

Sutter, M., and A. P. Kinziger. 2019. Rangewide tidewater goby occupancy survey using environmental DNA. Conservation Genetics. 20(3):597-613.

Swenson, R. O. 1999. The ecology, behavior, and conservation of the tidewater goby, *Eucyclogobius newberryi*. Environmental Biology of Fishes 55(1):99–114.

Swenson, R. O., and A. T. McCray. 1996. Feeding ecology of the tidewater goby. Transactions of the American Fisheries Society 125(6):956–970.

Swift, C. C., B. Spies, R. A. Ellingson, and D. K. Jacobs. 2016. A new species of the bay goby genus *Eucyclogobius*, endemic to southern California: evolution, conservation, and decline. PLOS ONE 11(7):e0158543.

Thomas, C. D. 1994. Extinction, colonization, and metapopulations: environmental tracking by rare species. Conservation Biology 8(2):373-378.

Thomas J. A., Bourn N. A. D., Clarke R. T., Stewart K. E., Simcox D. J., Pearman G. S., Curtis R., and Goodger B. 2001. The quality and isolation of habitat patches both determine where butterflies persist in fragmented landscapes. Proceedings of the Royal Society of London. Series B: Biological Sciences 268(1478):1791–1796.

United States Fish and Wildlife Service (USFWS) (2005) Recovery plan for the tidewater goby (*Eucyclogobius newberryi*). U.S. Fish and Wildlife Service Portland, Oregon.

Vörös, J., O. Márton, B. R. Schmidt, J. T. Gál, and D. Jelić. 2017. Surveying Europe’s Oonly cave-dwelling chordate species (*Proteus anguinus*) Using Environmental DNA. PLOS ONE 12(1):e0170945.

Wang, S., A. Anichowski, M. K. Tippett, and A. H. Sobel. 2017. Seasonal noise versus subseasonal signal: forecasts of California precipitation during the unusual winters of 2015-2016 and 2016-2017: S2S forecast of CA precipitation. Geophysical Research Letters 44(18):9513–9520.

Worcester, K.R. 1992. Habitat utilization in a central California coastal lagoon by the tidewater goby (*Eucyclogobius newberryi*). Master’s Thesis, California Polytechnic State University, San Luis Obispo, California.

Ying, Y., Y. Chen, L. Lin, and T. Gao. 2011. Risks of ignoring fish population spatial structure in fisheries management. Canadian Journal of Fisheries and Aquatic Sciences 68(12):2101–2120.

APPENDICES

APPENDIX A

Occupancy Analysis – Authored by Dr. Robert Dorazio

Multiscale occupancy models (Nichols et al., 2008; Mordecai et al., 2011) provide a useful framework for the analysis of data collected in eDNA surveys. In fact, the hierarchical relationships assumed between parameters of these models are exactly those induced by the three-level, nested sampling design of eDNA surveys (Dorazio and Erickson, 2018). These models have been used to analyze eDNA surveys of several species, including a fungal pathogen (Schmidt et al., 2013), the Burmese python (Hunter et al., 2015), the tidewater goby (Schmelzle and Kinziger, 2016, Sutter and Kinziger 2019), and a cave-dwelling aquatic salamander (Vörös et al., 2017).

Multiscale occupancy models can be used to estimate the spatial distribution of a species during a relatively short period of sampling. If surveys are repeated at sample locations (say, annually or seasonally), changes in occupancy state, such as colonization of a previously unoccupied location or extinction of an occupied location, can be inferred by analyzing each sampling period’s data separately. However, this approach does not allow potential drivers of the processes responsible for changes in occupancy state to be identified. To solve this problem, we propose a class of *dynamic* multiscale occupancy models in which temporal changes in occupancy states are specified as a function of explicit colonization and extinction processes. In the following section we first describe a versatile class of multistate occupancy models. We follow this section by describing models of sampling and observation processes induced by the sampling designs used in eDNA surveys.

Multistate models of occupancy dynamics

Suppose each of M survey locations is sampled during T disjoint (non-overlapping) periods and that we wish to model changes in occupancy state at these locations between the periods of sampling. Let $Z_{i,t}$ denote a random variable for the occupancy state of the ith survey location ($i = 1, \ldots, M$) during the tth sampling period ($t = 1, \ldots T$). We assume that $Z_{i,t}$ can have one of three values:

1 if the survey location is occupied,

2 if the survey location is unoccupied but has been occupied during the previous sampling period, or

3 if the survey location is unoccupied and has not been occupied previously.

State 3 distinguishes locations that have never been occupied from those that have previously been colonized but are temporarily unoccupied.

Our models of occupancy state dynamics include two distributional assumptions: one for the initial occupancy state during sampling period 1

$$Z_{i,1} \sim \mathrm{Cat}(\psi_i) \qquad (1)$$

and another for the change in occupancy state between sampling periods t and $t + 1$

$$Z_{i,t+1} | Z_{i,t} = k \sim \mathrm{Cat}(\phi_{k,i,t}) \qquad (2)$$

where Cat denotes the categorical distribution and where $t = 1, \ldots, (T - 1)$. In Eq. 1 $\psi_i = (\psi_i,,0,1-\psi_i)'$ denotes a vector containing the probabilities of each occupancy state during period 1. (We use the prime superscript to denote the transpose of a vector.) The parameter ψ_i is the probability that the ith location is occupied during period 1 ($Pr(Z_{i,1} = 1)$). If the ith location is not occupied during this period, $Z_{i,1}$ must equal 3 with probability $1-\psi_i$ because no previous sampling has occurred.

We assume that the initial occupancy probability may be formulated as a function of covariates whose values can vary spatially as follows:

$$\psi_i = F(\boldsymbol{x}^{\#}_{\boldsymbol{\psi,i}}\boldsymbol{\beta}_{\boldsymbol{\psi}} x'_{\psi,i} x'_{\psi,i} \beta_\psi) \tag{3 }$$

where $\boldsymbol{x}_{\psi,i}$ is a vector of regressors that codify the covariate measurements taken at the *i*th survey location during sampling period 1, and where $\boldsymbol{\beta}_\psi$ is a vector of parameters that specify the effects of the covariates on occupancy during sampling period 1. The function F provides a one-to-one mapping of a real-valued argument to a real number that lies on the closed interval (0, 1). We use the cumulative distribution function (cdf) of a Gaussian distribution for F but others, such as the cdf of a logistic distribution, also could be used without loss of generality.

In Eq. 2 the vector of probabilities $\boldsymbol{\phi}_{k,i,t}$ corresponds to the *k*th row of a matrix of transition probabilities $\boldsymbol{\phi}_{i,t}$ that we now define. Transitions among occupancy states correspond to *colonizations* of unoccupied locations and to *extinctions* of occupied locations. These transitions are assumed to be time- and state-dependent, as indicated by the following matrix of transition probabilities:

$$\Phi_{i,t} = \begin{pmatrix} 1-\varepsilon_{i,t} & \varepsilon_{i,t} & 0 \\ \gamma_{i,t} & 1-\gamma_{i,t} & 0 \\ \eta_{i,t} & 0 & 1-\eta_{i,t} \end{pmatrix}$$

where $\boldsymbol{\phi}_{k,l,i,t} = \Pr(Z_{i,t+1} = l \mid Z_{i,t} = k)$ is the probability of a transition from state k to state l during the interval between sampling periods t and $t + 1$; therefore, each row of $\boldsymbol{\phi}_{i,t}$ sums to one by construction. For example, the first row includes probabilities of extinction $\varepsilon_{i,t}$ (a change from state 1 to state 2) and persistence 1 - $\varepsilon_{i,t}$ (remaining in state 1) when the *i*th location is occupied during period t. Once a survey location is occupied it cannot change to state 3 (by definition), so the third element of row 1 is zero. Similarly, the second and third rows of $\boldsymbol{\Phi}_{i,t}$ include probabilities of colonization when the *i*th survey location is unoccupied during sampling period t. The third row contains the probability that this location is *first* colonized $\eta_{i,t}$ (a change from state 3 to state 1), whereas the second row contains the probability of recolonization $\gamma_{i,t}$ (a change from state 2 to state 1) of the *i*th survey location.

This model of occupancy-state dynamics is extremely versatile. Each of the probabilities of colonization and extinction may be formulated as functions of covariates whose values can vary spatially and/or temporally. In addition, the effects of dispersal of individuals from neighboring locations of each surveyed site can be used to specify colonization and extinction probabilities.

In the model that we developed for tidewater gobies, the probability of extinction at a survey location is specified as a function of location- and time-specific covariate values as follows:

$$\varepsilon_{i,t} = F(\boldsymbol{x}'_{\boldsymbol{\varepsilon},\boldsymbol{i},\boldsymbol{t}}\boldsymbol{\beta}_{\boldsymbol{\varepsilon}} x'_{\varepsilon,i,t} x'_{\varepsilon,i,t} \beta_{\varepsilon}) \qquad (4)$$

where $\boldsymbol{x}_{\varepsilon,i,t}$ is a vector of regressors that codify the covariate measurements taken at the *i*th survey location during sampling period t, and where $\boldsymbol{\beta}_{\varepsilon}$ is a vector of parameters that specify the effects of the covariates on the probability of extinction. The probabilities of colonization are formulated similar to that of extinction except that we also specify the effects of dispersal of individuals from neighboring locations. To be specific, we assume that colonization of an unoccupied survey location may occur from one of two processes depending on the occupancy states of the location's neighbors. If none of the neighboring locations are occupied, we assume that colonization rates are functions of location- and time-specific covariate values (that is, we adopt the approach used to specify extinction probability). However, if one or more neighboring locations are occupied, we assume that colonization occurs by the movements of individuals from occupied neighbors, that is, colonization is a localized process . Let $\omega_{i,n}$ denote the probability that the *i*th survey location is colonized by movements of individuals from a neighboring location (indexed by *n*) during the period between sampling intervals *t* and *t* + 1. The probability that the *i*th survey location is colonized by individuals from at least one of its occupied neighbors during this time interval is

$$1 - \prod_{n \epsilon \mathcal{N}_i} [1 - \omega_{i,n} I(Z_{n,t} = 1)] \tag{5}$$

where $\mathcal{N}_i$ denotes the set of location indices corresponding to the neighbors of the *i*th survey location and where *I* is an indicator function whose value equals one for a true argument and zero otherwise. Note that the expression in Eq. 5 equals zero if none of the *i*th survey location's neighbors are occupied (i.e., if $Z_{n,t} \neq 1$ for all $n \in \mathrm{N}_i$). In other words, if a survey location's neighbors are all unoccupied, we assume that local colonization's cannot occur and that any colonization must stem from the other (non-local) colonization process. Figure 16 contains a diagrammatic illustration of our model's specification of the effects of dispersal from occupied neighboring locations to an unoccupied survey location. In this figure only two of four neighboring locations are occupied, so the probability of being colonized by at least one of the four neighbors depends only on the colonization probabilities of the two occupied neighbors, as specified in Eq. 5.

We specify the colonization probabilities $\gamma_{i,t}$ and $\eta_{i,t}$ in terms of the two distinct colonization processes as follows:

$$\gamma_{i,t} = F\left(x'_{\gamma,i,t}\beta_\gamma\right)I\left(q_{i,t} = 0\right) + \left(1 - \prod_{n\epsilon\mathcal{N}_i}\left[1 - \omega_{i,n}\, I(Z_{n,t} = 1)\right]\right) \qquad \textbf{(6)}$$

$$\eta_{i,t} = F\left(x'_{\eta,i,t}\beta_\gamma\right)I\left(q_{i,t} = 0\right) + \left(1 - \prod_{n\epsilon\mathcal{N}_i}\left[1 - \omega_{i,n}\, I(Z_{n,t} = 1)\right]\right) \qquad \textbf{(7)}$$

where $q_{i,t} = \sum_{n\in \mathrm{N}_i} I(Zn, t = 1)$ denotes the number of neighbors of the *i*th survey location that are occupied during period *t*. As noted earlier, the second parenthesized, $I\left(q_{i,t} = 0\right)$, term in these equations equals zero whenever $q_{i,t}$ equals zero.

Our formulation of the local colonization process parameterized by $\omega_{i,n}$ is similar to the approach described by Broms et al. (2016). This formulation provides considerable flexibility by allowing different models of $\omega_{i,n}$ to be constructed. For example, in the simplest model we assume $\omega_{i,n}$ is a constant (say, ω) that does not differ among survey locations. In this case the expression in Eq. 5 simplifies to

$$1 - (1 - \omega)^{q_{i,t}}$$

where $q_{i,t}$ is the number of occupied neighbors of the *i*th survey location during sampling period *t*. In this case the probability of colonization of the *i*th survey location is a mono- tone increasing function of the number of occupied neighbors of that location. In a more complicated model, we might assume that colonization between locations depends on physical features of the path between locations (e.g., measures of connectedness) or on habitat gradients. In this case we could specify $\omega_{i,n}$ as a function of these covariates as follows:

$$\omega_{i,n} = F(x'_{\omega,i,n}\beta_{\omega})$$

where $\boldsymbol{x}_{\omega,i,n}$ is a vector of regressors that codify the covariate measurements assumed to influence colonization between the *i*th and *n*th survey locations and where $\boldsymbol{\beta}_{\omega}$ is a vector of parameters that specify the effects of these covariates on $\omega_{i,n}$.

Models of eDNA occurrence and detection

Thus far, we have only modeled occupancy states in the first level of sampling in eDNA surveys (i.e., at the location level). In most eDNA surveys multiple samples are collected at each location, and any eDNA that may be present in each of these samples is detected using multiple PCR replicates. Therefore, it necessary to model the presence of eDNA in samples and the detection of eDNA in PCR replicates. Following Dorazio and Erickson (2018), we model the occurrence and detection of eDNA in samples using nested conditional distributions.

Suppose $J_{i,t}$ samples are collected independently from the *i*th survey location during period *t*. Let $A_{i,j,t}$ denote a random variable whose values indicate the presence ($A_{i,j,t}$ = 1) or absence ($A_{i,j,t}$ = 0) of eDNA in the *j*th sample ($j = 1, \ldots, J_{i,t}$). We assume

$$A_{i,j,t} | Z_{i,t} \sim \text{Bernoulli}(\theta_{i,j,t} I(Z_{i,t} = 1)) \qquad \textbf{(8)}$$

where the parameter $\theta_{i,j,t}$ is the conditional probability that eDNA is present in the *j*th sample of location *i* during period *t* given that this location is occupied (more correctly, given that eDNA of the target species is present) during period *t*. Note that $A_{i,j,t}$ equals zero with probability one if the *i*th survey location is unoccupied during period *t*.

Suppose $K_{i,j,t}$ independent replicates are extracted from the *j*th sample of survey location *i* during period *t* and are amplified using PCR chemistry. Let $Y_{i,j,t}$ denote a random variable for the number of these replicates in which eDNA is detected. We assume

$$Y_{i,j,t}|A_{i,j,t} \sim \text{Binomial}(K_{i,j,t}p_{i,j,t}I(A_{i,j,t} = 1)) \tag{9}$$

where the parameter $p_{i,j,t}$ is the conditional probability that eDNA is detected in a single replicate of the *j*th sample given that eDNA is present in this sample. Eq. 9 implies that $Y_{i,j,t}$ equals zero with probability one if eDNA is absent from the sample.

As we have done previously, the parameters in Eqs. 8 and 9 are specified as functions of covariates that are thought to be informative of the occurrence or detection of eDNA in samples. For example, adopting the notation used by Dorazio and Erickson (2018), we assume

$$\theta_{i,j,t} = F(\omega'_{i,j,t}\alpha) \tag{10}$$

for sample occurrence probability and

$$p_{i,j,t} = F(v'_{i,j,t}\delta) \tag{11}$$

for the probability of detecting eDNA in a PCR replicate. Covariates of $\theta_{i,j,t}$ and $p_{i,j,t}$ are codified in the vectors of regressors ($\boldsymbol{w}_{i,j,t}$ and $\boldsymbol{v}_{i,j,t}$, respectively) and offer ample opportunity to identify factors that are thought to influence eDNA occurrence and detection at the sample level.

Table 4 Definitions of model parameters and covariates used in the dynamic multiscale occupancy model.

Parameter	Definition
ψ_1	Probability of initial occupancy of location during sampling period 1
β_{ψ_1}	Vector of parameters that specify covariate effects on ψ_1
$\varepsilon_{i,t}$	Probability of extinction of a location *i* between *t* and *t*+1
β_{ε}	Vector of parameters that specify covariate effects on ε
$\gamma_{i,t}$	Probability of colonization of a previously occupied location *i* between *t* and *t*+1 (recolonization)
β_{γ}	Vector of parameters that specify covariate effects on γ
$\eta_{i,t}$	Probability of colonization of a location *i* that has not been previously occupied between *t* and *t*+1 (colonization)
β_{η}	Vector of parameters that specify covariate effects on η
$\omega_{i,n}$	Probability that survey location *i* is colonized by neighboring location *n* between *t* and *t*+1
$\beta_{\omega_{i,n}}$	Vector of parameters that specify covariate effects on $\omega_{i,n}$
$\theta_{i,j,t}$	Probability of presence of eDNA in water sample *j* given presence of eDNA at a location *i* during time *t*
α	Vector of parameters that specify covariate effects on $\theta_{i,j,t}$
$\rho_{i,j,t}$	Probability of detection in a qPCR given presence in a water sample *j* and location *i* during time *t*
δ	Vector of parameters that specify covariate effects on $\rho_{i,j,t}$

APPENDIX B

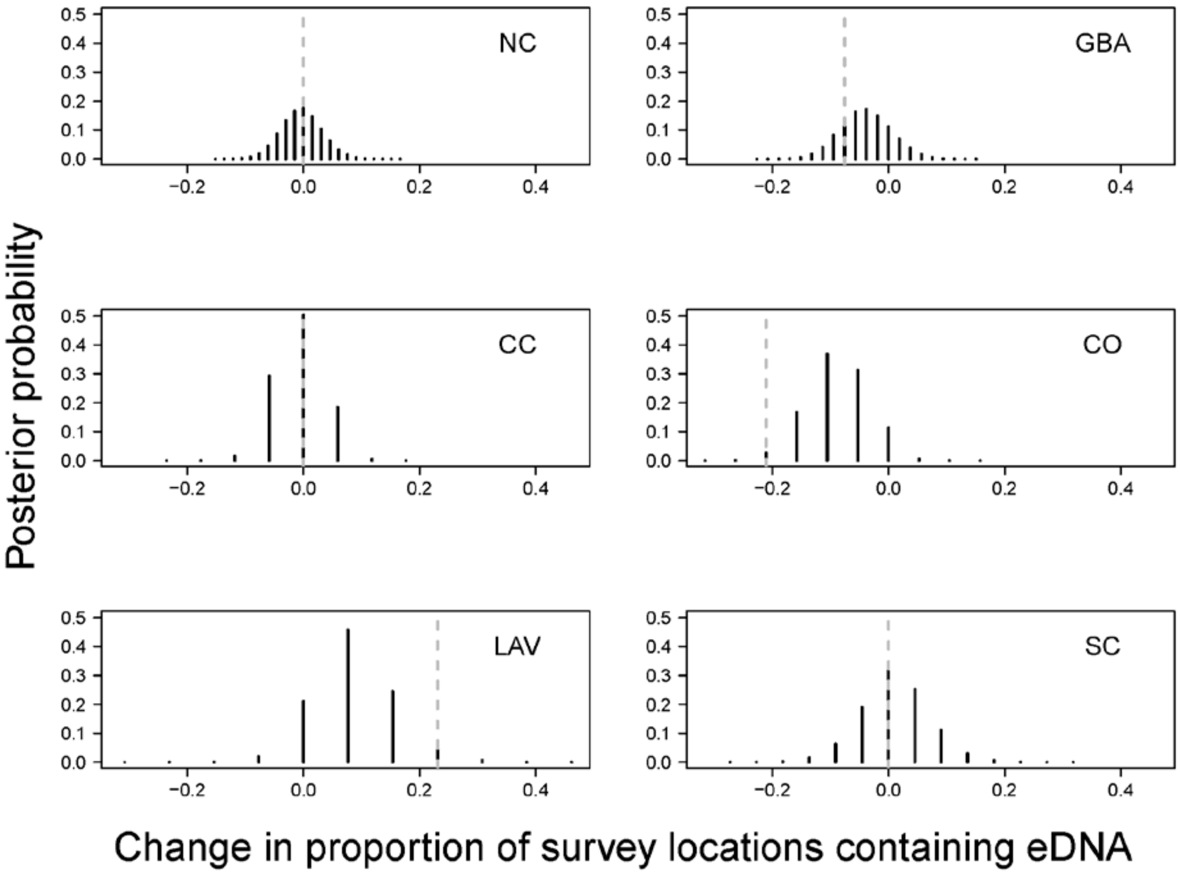

Figure 12 Posterior probability of change in proportion of sites containing eDNA from 2016 to 2017 within the Recovery Units (NC = North Coast, GBA = Greater Bay Area, CC = Central Coast, CO = Conception, LAV = Los Angeles/Ventura, SC = South Coast) delineated by the Tidewater goby Recovery Plan (USFWS 2005) along the California coast. A shift in the positive direction on the x-axis represents in increase in the number of occupied sites. Bar density is representative of the site density in the recovery units. Naïve rates of change in proportion of sites containing eDNA are represented by the dashed grey line. Data were analyzed using a dynamic multiscale occupancy model.

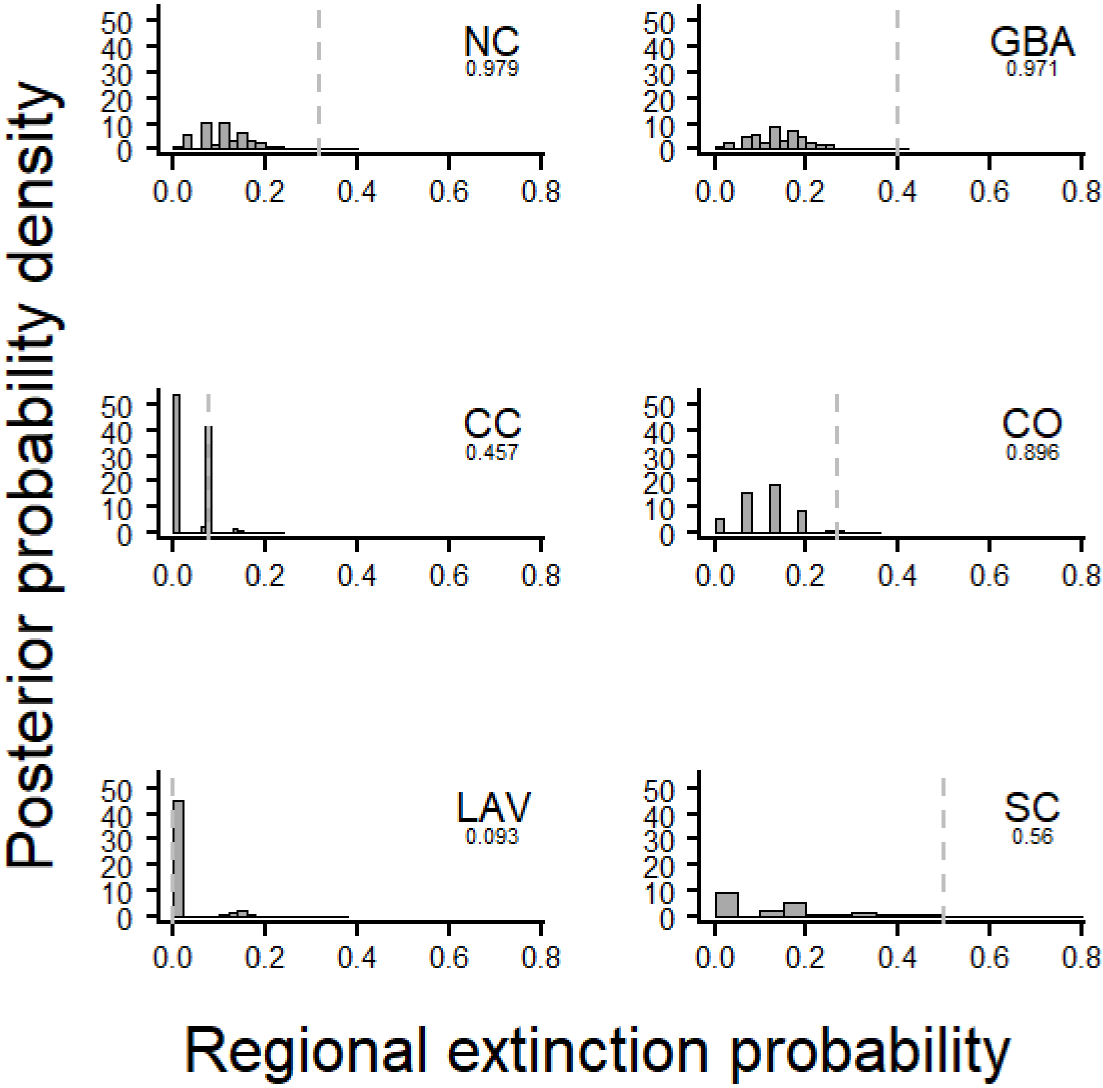

Figure 13 Estimates of probability of extinction within Recovery Units (NC = North Coast, GBA = Greater Bay Area, CC = Central Coast, CO = Conception, LAV = Los Angeles/Ventura, SC = South Coast) delineated by the Tidewater Goby Recovery Plan from a rangewide eDNA survey conducted in 2016 and 2017. Conditional naïve estimates that do not account for non-detection are represented by dashed grey lines. Grey bars represent the estimates produced from a multiscale model that explicitly accounts for non-detection at three levels of sampling. Number under the Recovery Unit designation are the proportion of model estimates that fall below the naïve estimate of extinction.

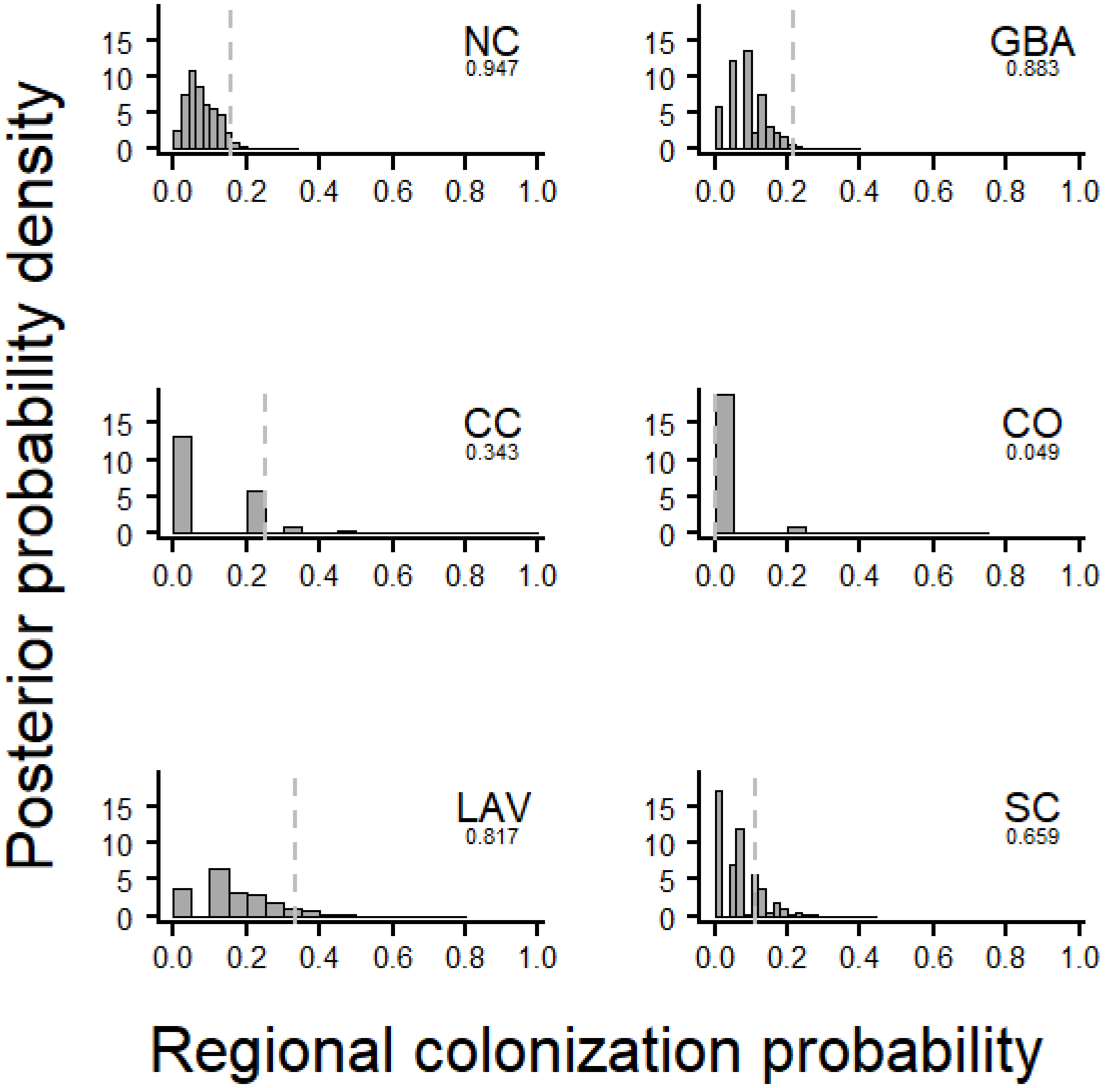

Figure 14 Estimates of probability of colonization within Recovery Units (NC = north coast, GBA = greater bay area, CC = central coast, CO = Conception, LAV = Los Angeles/Ventura, SC = south coast) delineated by the Tidewater goby Recovery Plan from a rangewide eDNA survey conducted in 2016 and 2017. Conditional naïve estimates that do not account for non-detection are represented by the dashed grey lines. Grey bars represent the estimates produced from a multiscale model that explicitly accounts for non-detection at three levels of sampling. Number under the Recovery Unit designation are the proportion of model estimates that fall below the naïve estimate of extinction.

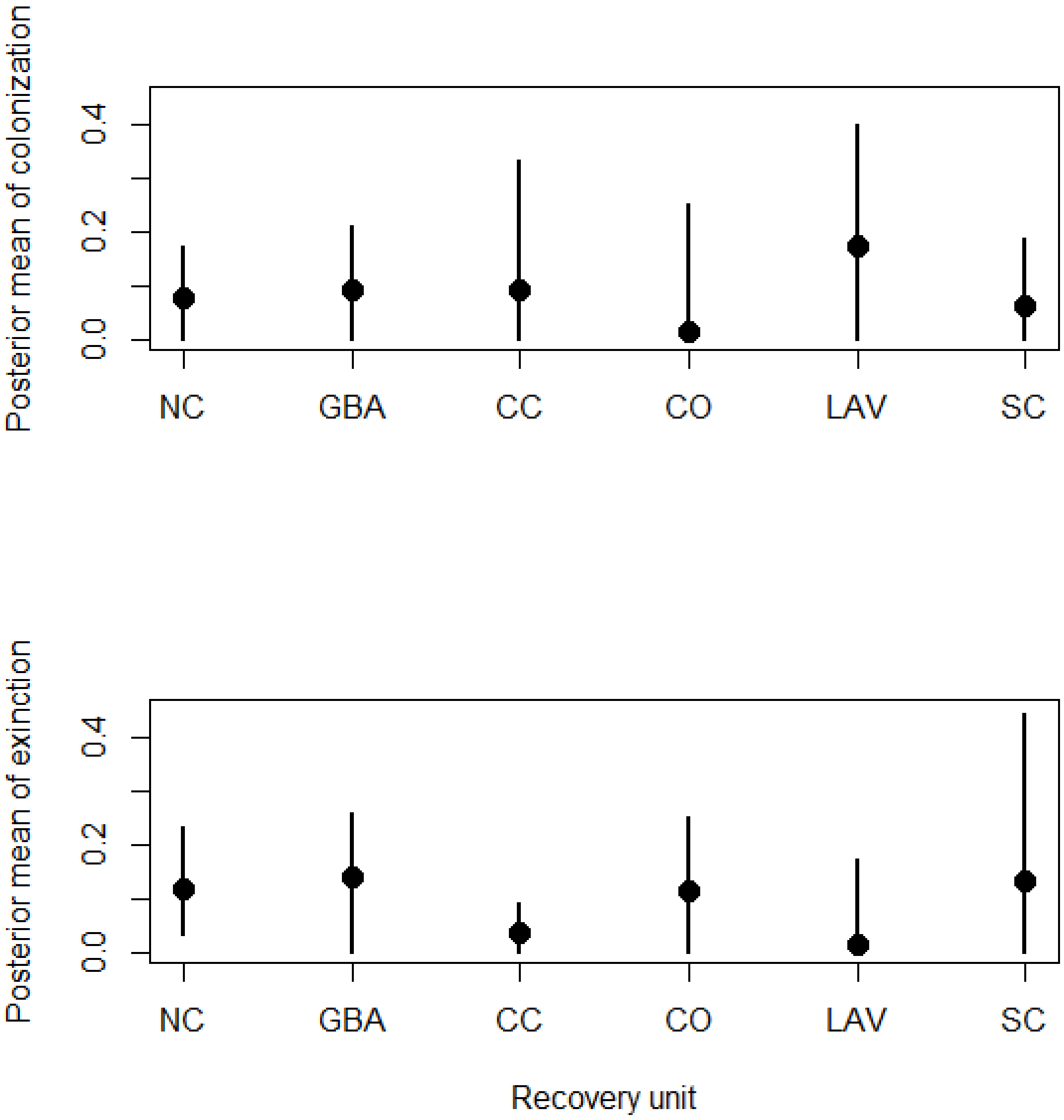

Figure 15 Estimates of extinction and colonization dynamics and their 95% credible interval for tidewater goby *Eucyclogobius spp.* populations in recovery units outlined by the 2005 USFWS recovery plan between 2016 and 2017 (NC = north coast, GBA = greater bay area, CC = central coast, CO = Conception, LAV = Los Angeles/Ventura, SC = south coast). Data were collected as part of rangewide eDNA survey and analyzed using a dynamic multiscale occupancy model.

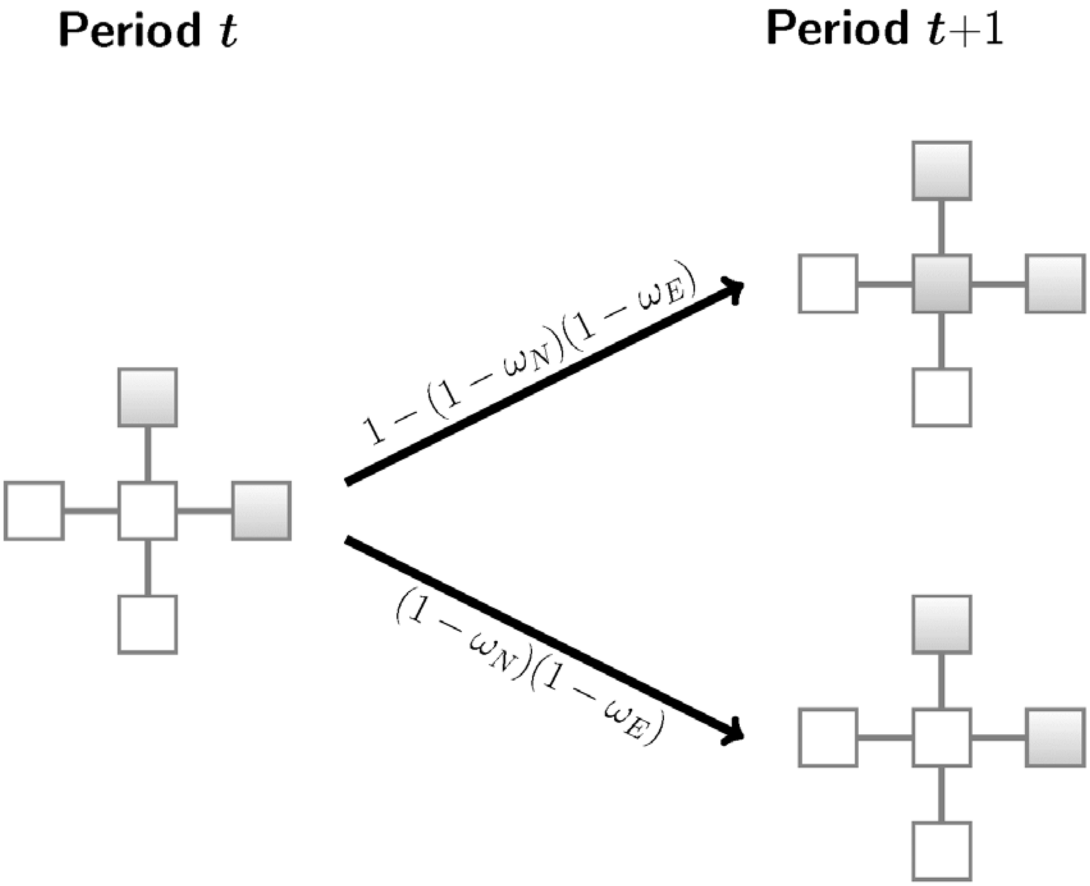

Figure 16 Diagrammatic illustration of our model's specification of the effects of dispersal from occupied neighboring locations to an unoccupied survey location. In this figure only two of four neighboring locations are occupied, so the probability of being colonized by at least one of the four neighbors depends only on the colonization probabilities of the two occupied neighbors, as specified by the equation: $\mathbf{1-\prod_{n\epsilon\mathcal{N}_i}[1-\omega_{i,n}\,I(Z_{n,t}=1)]}$**.**