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.

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

1   INTRODUCTION

2   Metapopulation dynamics, as originally conceived by Levin (1969), assumes that

3   a set of isolated sites of suitable habitat can periodically experience local extirpations if

4   occupied, while in the same time step some unoccupied sites may be recolonized by

5   dispersing individuals. Conservation and management decisions regarding a species

6   thought to exist as a metapopulation are most effective when extinction and colonization

7   dynamics are understood and incorporated into management plans (Anthes et al. 2003,

8   Armstrong 2005, Marsh and Trenham 2001, Semlitsch 2008, Ying et al. 2011). The

9   understanding and effective incorporation of extinction and colonization dynamics

10   requires accurate, unbiased, estimates of extinction and colonization across their range.

11   Despite the need for accurate estimation of metapopulation dynamics, these dynamics are

12   notoriously difficult to describe and subject to a number of potential errors, including

13   incorrect estimation of site size, unknown sites existing in a study area, and the potential

14   for non-detections, where non-detection of a target species is interpreted as a true absence

15   when the target was actually present (Moilanen 2002, MacKenzie 2003).

16   Extinction and colonization rates of metapopulations are typically estimated

17   using a record of site occupancy across repeated field surveys through time. These

18   presence-absence data are then used to generate site occupancy histories for many habitat

19   sites from which extinction and colonization of individual sites is inferred. A critical

20   assumption of this approach is that non-detection is indicative of a true absence.

21   However, it is well-established that this assumption is violated for most field survey

22 methods, especially when abundance is low, field collections are difficult, or collection

23 effort is limited (Gu & Swihart 2004, MacKenzie and Royle 2005). Confusing non-

24 detection as an absence leads to biased estimates of site occupancy (Gu & Swihart 2004,

25 Moilanen 2002), and for metapopulation studies non-detection errors can lead to biased

26 estimates of extinction and colonization dynamics. Unlike occupancy, which is biased

27 low by non-detections, bias in extinction and colonization dynamics can be over- or

28 under- estimated depending on which time period the non-detection error occurred. In

29 response to non-detection biases, analytic methods have been developed to account for

30 non-detection (MacKenzie et al., 2003, Moilanen 2002). Non-detection can be accounted

31 for by completing multiple surveys of each habitat site within a single season and

32 analyzing the resulting occupancy data with models that use the within season sampling

33 data to estimate detection probabilities (Mackenzie et al. 2002; MacKenzie et al. 2003,

34 Moilanan 2002).

35   A key goal of metapopulation studies is to elucidate environmental drivers of

36 extinction and colonization dynamics (Hanksi 1989, 1998, Hanski and Gilpin1991).

37 Armstrong (2005) suggested species declines can be halted by the integration of two

38 paradigms: (1) a metapopulation paradigm which focuses on factors that influence site

39 connectivity such as site size, spatial structure, and site density, and (2) a habitat

40 paradigm that emphasizes the link between local persistence and site-level environmental

41 covariates. Metapopulation paradigm models are predicated on two widely accepted and

42 generally supported assumptions: (1) isolated sites have lower colonization rates relative

43 to more densely spaced sites, and (2) smaller sites are more vulnerable to local

44 extirpation relative to larger habitat sites (Armstrong 2005). Research within the habitat

45 paradigm has shown how habitat factors such as topography, vegetation, and available

46 resources can affect site occupancy (Akçakaya and Atwood 2002, Fleishman et al. 2002,

47 Thomas et al. 2001). Generally, high quality habitat within a site leads to colonization or

48 persistent occupation while poor quality habitat can lead to extirpations or lower

49 occupancy rates (Armstrong 2005, Fleishman et al. 2002, Franken and Hik 2004, Thomas

50 1994).

51   Few studies have elucidated rangewide metapopulation dynamics and their

52 environmental drivers owing to time and money constraints. Repeated rangewide surveys

53 require extensive planning, particularly if the area is large, the sites are numerous, and the

54 target is small or cryptic (Moilanen 2002). Environmental DNA (eDNA) has gained

55 popularity in the past decade as a method of rapidly and efficiently detecting species

56 (Foote et al. 2012, Gingera et al. 2016, Goldberg at al. 2013, Pilliod et al. 2013).

57 Environmental DNA surveys are a method of surveying the environment for genetic

58 material that has been sloughed off an individual to use as an index of presence.

59 Environmental DNA has been repeatedly shown to outperform traditional methods of

60 species detection, often detecting a species when traditional methods fail (Boussarie et al.

61 2018, Dejean et al. 2012, Port et al. 2016, Schmelzle and Kinziger 2016, Thompsen and

62 Willerslev 2015). Direct comparisons have shown that eDNA monitoring can have

63 double the detection probability of traditional approaches (Schmelzle and Kinziger 2016).

64   Environmental DNA data collected using a hierarchical, multi-season sampling

65 design can be analyzed via multiscale occupancy models (Dorazio and Erickson 2018,

66   Sutter and Kinziger 2019). Data from this project were analyzed using a *novel* class of

67   *dynamic* multiscale occupancy model developed in collaboration with Dr. Robert Dorazio

68   of San Francisco State University (a description of the model, authored by Dr. Dorazio,

69   can be found in Appendix A). Briefly, the occupancy model used here belongs to a class

70   of multiscale occupancy models that uses a hierarchical structure to allow estimation of

71   non-detection at different sampling levels. In the case of eDNA surveys, there are three

72   sampling levels: sites, samples, and qPCR replicates (Dorazio and Erickson 2018).

73   Multiscale models do not directly estimate extinction and colonization dynamics;

74   detecting extinction and colonization dynamics using requires a supplementary analysis

75   of the change in modeled site occupancy across seasons. The novel '*dynamic*' class of

76   model used in this study builds on previous multiscale models by incorporating extinction

77   and colonization as model parameters, thus allowing testing of potential covariates that

78   may drive extinction and colonization dynamics. This model assumes that sites are closed

79   to immigration during sampling periods but open between sampling periods. This

80   assumption allows estimation of detection probabilities at site, sample, and qPCR

81   replicate levels as well as extinction and colonization dynamics.

82   To demonstrate the power of estimating metapopulation dynamics using eDNA

83   surveys combined with multiscale occupancy modeling, I analyzed tidewater goby

84   *Eucyclogobius spp*. occupancy data from two years of repeated rangewide eDNA

85   surveys. Tidewater goby are a benthic fish endemic to coastal California where they

86   inhabit lagoons, bays, and the estuaries of rivers and streams (Swenson 1999). Originally

87   thought of as one species across their range, recent genetic research has divided the genus

88   into two distinct species (Swift et al. 2016). The southern tidewater goby *E. kristinae* is

89   known from only nine site south of the Palos Verdes peninsula, Los Angeles County, to

90   San Diego County, whereas the northern tidewater goby *E. newberryi* is known from 70+

91   sites between the Palos Verdes peninsula and the California-Oregon border (Swift et al.

92   2016). The use of 'tidewater goby' in this text refers to either species, or the

93   *Eucyclogobius* genus as a whole. Tidewater goby were listed as endangered under the

94   U.S. Endangered Species Act in 1994; the most recent recovery plan for tidewater goby

95   divides the California coast into six "recovery units." From north to south these recovery

96   units are North Coast, Greater Bay Area, Central Coast, Conception, Los

97   Angeles/Ventura, and South Coast (USFWS 2005, Figure 1A). Northern tidewater goby

98   are found in the North Coast, Greater Bay Area, Central Coast, Conceptions, and Los

99   Angeles/Ventura recovery units, while the southern tidewater goby range is encompassed

100   by the South Coast recovery unit. Each of these recovery units are further divided into

101   various numbers of sub-units, 26 in total, for recovery and management efforts.

102   Tidewater goby metapopulation population structure has been described as a core-

103   satellite model where populations in large wetlands serve as core populations with

104   relatively high persistence and high dispersal while populations in smaller wetlands may

105   experience higher extinction rates, possibly driven by unfavorable dry conditions

106   (Lafferty et al. 1999a, 1999b). Under this model, recolonization of extirpated sites would

107   be driven by individuals originating from the larger, more stable, core populations

108   (Lafferty et al. 1999a, 1999b). Observations by Lafferty et al. (1999b) suggest that

109   colonization may be promoted by localized flooding of small, occupied, streams where

110   the associated increase in longshore current could deliver individuals flushed from one

111   site to another unoccupied site. In this way, flooding may be beneficial to tidewater goby

112   persistence at the rangewide level by allowing colonization of previously extirpated sites.

113   Metapopulation dynamics have been incorporated into management and recovery efforts

114   across their range (USFWS 2005). The most recent recovery plan states that in order to

115   consider tidewater goby for downlisting a metapopulation viability analysis must find a

116   greater than 75% chance of all recovery units surviving out to 100 years. But this

117   requirement does not appear to be the best fit across all tidewater goby recovery units; in

118   northern California, geographically separated populations were found to lack any signal

119   of extinction and colonization dynamics over time scales several decades (Kinziger et al.

120   2015).

121   The objective of this study was to estimate rangewide metapopulation dynamics

122   for endangered tidewater goby by analyzing site occupancy histories generated from

123   eDNA surveys with a *novel* class of *dynamic* multiscale occupancy model that explicitly

124   accounts for non-detection at the three hierarchical levels inherent to eDNA surveys:

125   sites, samples, and qPCR replicates while allowing estimation of extinction and

126   colonization as model parameters with covariates. The use of eDNA survey methods

127   enabled rangewide monitoring of 190 locations in two consecutive years along the entire

128   1350 km of the California coast (Figure 1A). The multiscale dynamic occupancy model

129   was used to examine: (1) rangewide occupancy, extinction, and colonization and how the

130   measured covariates (vegetation, salinity, temperature, dissolved oxygen, estuary size,

131   and distance between neighboring sites) affected these dynamics, (2) the biases resulting

132   from naïve estimates of metapopulation dynamics and their relationship to model

133   estimates, (3) extinction and colonization within recovery units to better understand the

134   spatial variability in tidewater goby metapopulation dynamics, and lastly (4) which

135   covariates (tide, salinity, turbidity, temperature, and dissolved oxygen) affected detection

136   of tidewater goby eDNA in water samples and in qPCR. This work illustrates that eDNA,

137   when used in combination with multiscale occupancy modeling, has the ability to

138   efficiently survey a large number of locations, detect metapopulation dynamics, and

139   evaluate their drivers, at both broad and narrow geographic scales.

140

141 METHODS

142 Data Collection, Field, and Laboratory Methods

143 Site occupancy histories were available for two years, 2016 and 2017. The 2016

144 site occupancy histories are described by Sutter and Kinziger (2019), and the 2017 site

145 occupancy data was generated by the author following Sutter and Kinziger (2019). This

146 created a congruent dataset between the Sutter and Kinziger (2019) survey conducted in

147 2016 and this follow-up 2017 survey thereby allowing analysis of the combined

148 occupancy and environmental data from both surveys. Between May and September of

149 2017, I visited 194 sites, from which I collected between one and six 2.0 L water samples

150 (413 in total). I conducted between three and six qPCR replicates per water sample.

151 These surveys ran along the California Coast from 2 km south of California-Oregon

152 border, to 45 km north of the California-Mexico border (Figure 1A). I used Global

153 Positioning System (GPS) data and photographs from the 2016 survey to collect samples

154 as near as possible to where the previous samples were collected, usually within 100 m. If

155 access or conditions had changed between the two seasons, samples were collected as

156 close as possible to the 2016 locations while duplicating sampling protocol from Sutter

157 and Kinziger (2019); i.e. maintaining distance between multiple samples at a site to avoid

158 the transport of eDNA from disturbed sediment or water from a previous sampling

159 location to another location. Three of the sites sampled in 2016 were not resampled in

160 2017 due to access restrictions. Environmental data collected with each water sample

161 included water temperature (°C), dissolved oxygen (mg/L), salinity (ppt), presence or

162 absence of tidal influence (open or closed to daily tidal flow at time of sampling), and

163   presence or absence of aquatic vegetation. Additional data collected at each sample

164   location included date, time, and latitude and longitude.

165   Water filtering, eDNA extraction, and qPCR protocols were conducted following

166   Sutter and Kinziger (2019). Samples were filtered over a 47 mm diameter polycarbonate

167   track-etched filter membrane with 3.0 μm pore size (Millipore TSTP 04700). Extraction

168   of eDNA from filters was conducted using a DNeasy Blood and Tissue Kit (Qiagen,

169   USA) with modifications for eDNA extraction (Schmelzle and Kinziger 2016, Sutter and

170   Kinziger 2019). Detection of northern and southern tidewater goby was completed using

171   two assays that are designed to target a 119 base pairs region on tidewater goby's

172   mitochondrial cytochrome *b* gene (Schmelzle and Kinziger 2016, Sutter and Kinziger

173   2019). A northern tidewater goby assay was used to test for presence of northern

174   tidewater goby eDNA in samples taken north of Palos Verdes (North Coast, Greater Bay

175   Area, Central Coast, Conception, and Los Angeles/Ventura recovery units) while a

176   southern tidewater goby eDNA assay was used to test for presence of southern tidewater

177   goby eDNA in samples taken from south of Palos Verdes (South Coast recovery unit).

178   Extraction of eDNA from filters and qPCR preparation was performed on separate

179   benchtops in a dedicated laboratory space away from high copy number PCR product.

180   Additionally, qPCR preparations were performed under a hood with high efficiency

181   particulate air (HEPA) filter that was exposed to ultra-violet light for at least 30 minutes

182   prior to use, along with pipettes, centrifuges, and consumables used in reaction

183   preparation.

184   <u>Spatial data for sites</u>

185   Spatial data for sites were obtained from the California Department of Fish and

186   Wildlife (CDFW) and Pacific States Marine Fisheries Council (PSMFC). A California

187   shore type data, which classifies the California coast by habitat using NOAAs

188   Environmental Sensitivity Index (ESI), was obtained from CDFW. This was used to

189   measure the total distance between pairs of sites within sub-units and to measure the

190   distance between pairs of sites within sub-units that are rocky. Distance between sites

191   within sub-units was determined using program *R* and the package RIVERDIST. To

192   generate matrices of shoreline distance the shoreline data was altered to maintain a

193   continuous shoreline; embayments, estuaries, or rivers were removed from the original

194   shoreline data and the data was reconnected across their openings in the shortest path.

195   The distance of rocky coast between sites within subunits was obtained using ArcMap

196   10.5.1 (*Esri Inc.*) and the California shore type dataset. I recorded any length of coast

197   between sites in each sub-unit whose primary ESI category was one of the following: 1A

198   (exposed rocky shore/exposed), 1B (solid man-made structures), 1C (exposed rocky cliffs

199   with boulder talus), 2A (exposed wave-cut platforms in bedrock, mud, or clay), or 2B

200   (exposed scarps and steep slopes in clay). Lastly, estuary size was determined using the

201   West Coast USA Current and Historical Estuary Extent data set from the Pacific Marine

202   and Estuarine Fish Habitat Partnership (PSMFC GIS 2017). Sites were assigned to one of

203   two bins based on estuary size, those sites equal to or less than 1.0 ha were deemed

204   “small” while “large” sites were over 1.0 ha following Lafferty (1999).

205 <u>Occupancy data</u>

206   For the two species assays used in these analyses, Sutter and Kinziger (2019)

207 previously determined their limit of detection and associated critical threshold values

208 (Ct), this value is the cycle number at which a qPCR replicate reaches a threshold of

209 fluorescence for a positive reaction. For the northern tidewater goby assay, a Ct value of

210 40.87 or lower was assumed to be a positive eDNA detection. For the southern tidewater

211 goby assay a Ct of 40.04 or lower was assumed to be a positive detection. This data was

212 passed to the multiscale occupancy model as either a positive (1) or negative (0) detection

213 for each of the qPCR replicates preformed on a sample. I assumed that a single positive

214 detection was indicative of tidewater goby presence at that location, and that there were

215 not false positives.

216   Occupancy Model

217   Data collected by this research was analyzed using a dynamic multiscale

218 occupancy model (see Appendix A for full details). Briefly, the model was used to

219 estimate rangewide occupancy in 2016 (ψ_1) and 2017 (ψ_2), rangewide extinction (ϵ),

220 rangewide colonization (η), and colonization within specified neighborhoods (ω) between

221 2016 and 2017. The model accounted for non-detection errors by estimating the

222 probability of detecting eDNA in a water sample given its presence at a site (θ), and

223 probability of detecting eDNA in a qPCR replicate given its presence in water sample (ρ).

224 Occupancy, detection in a water sample, and detection in a qPCR replicate are estimated

225 using the likelihood function $L(\psi, \theta, p \mid H) \propto \prod_{i=1}^{S} Pr(H_i)$, where H_i represents qPCR

226 detection history at site i, across all possible sites S. Site occupancy state is then predicted

227 for single sites in a single year as either an absence (0) or a presence (1). Occupancy

228 states are represented in the form of a $Z_{i,t}$ value, where $Z_{i,t}$ is the occupancy state at site i

229 at time t. The modeled occupancy state for all sites are compared across time t and $t+1$ to

230 estimate the probability of a site's occupancy state transitioning from state k in time t to

231 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

232 contained in the transition matrix $\Phi_{i,t}$ and represent extinction (ε) and colonization (η)

233 probabilities. If an unoccupied site has occupied neighbors, it is assumed that any

234 colonization at that site stems from its occupied neighbors within its recovery sub-unit,

235 and not from outside the sub-unit. The structure of this model allows separation of

236 metapopulation parameter estimates (occupancy, extinction, and colonization) based on

237 geography. This allows the estimation of occupancy, extinction, and colonization within

238 the tidewater goby management delineated recovery units. Consequently, results are

239 presented at both the rangewide geographic scale as well as the scale of the recovery unit.

240 This model is Bayesian based; the models Markov Chain-Monte Carlo algorithm was run

241 for 110,000 iterations, of which the first 10,000 were discarded.

242 <u>Covariate data</u>

243   Model parameters were modeled as functions of different environmental and

244 spatial covariates. Instead employing a model selection approach where all possible

245 covariate combinations were tested and ranked, each covariate was carefully and

246 individually considered before being included as a possible covariate for a parameter in

247 the analysis. Consideration was based on previously published literature regarding

248   tidewater goby ecology and eDNA methods (Table 1). The environmental covariates

249   examined included salinity (ppt), water temperature (°C), dissolved oxygen (mg/l), depth

250   (cm), turbidity (seconds of filtering time), proportion of sample locations with aquatic

251   vegetation, and presence or absence of tidal influence. The two spatial covariates

252   investigated were the pairwise distance between sites within sub-units and the size of the

253   estuary at each site. The distances between sites within a sub-unit were used as the basis

254   for a neighborhood analysis that examined the probability of colonization from

255   neighboring sites. Further explanation of which covariates were tested for each parameter

256   can be found in Table 1. Model results and the strength of evidence for each parameter-

257   covariate relationship were considered individually via the posterior distribution of each

258   parameter-covariate relationship; evidence of parameter-covariate relationships is

259   presented as the proportion or percent of the posterior distribution above or below zero.

260

261   RESULTS

262   Of the original 194 sites surveyed in both 2016 and 2017, a total of 190 were

263   available for analysis (Figure 1A). One site was removed because samples were collected

264   at a location where tidewater goby occupancy seemed highly improbable, the habitat was

265   riverine in nature and not proximal to the river's estuary. Three other sites were removed

266   due to missing covariate values. The analysis was conducted on a combined data set that

267   included both the northern and southern tidewater goby species; this was done because

268   southern tidewater goby are restricted too few locations to allow analysis using the model

269   employed herein. The parameter-covariate relationships examined by the dynamic

270   multiscale model can be found in Table 1. The posterior distributions for each

271   relationship were considered individually, those distributions that were strongly shifted

272   away from zero were considered to represent a parameter-covariate relationship. Overall,

273   occupancy (ψ) exhibited a relationship to the proportion of sample locations at a site with

274   vegetation, but not with salinity (Figure **2**). Colonization within sub-units (ω) decreased

275   with increasing distance between sites. None of the tested covariates, including estuary

276   size, salinity, temperature, and dissolved oxygen, had a significant effect on probability

277   of extinction (ε, Figure **2**). Detection in a water sample (θ) was found to vary with

278   salinity, dissolved oxygen, and the presence of tide (Figure 3). Turbidity and temperature

279   had no effect on detection in water samples. Detection in a qPCR replicate (ρ) was found

280   to vary with salinity (Figure 3).

281   Rangewide

282   The dynamic multiscale model estimated that the proportion of sites occupied by

283   tidewater goby was the same in 2016 and 2017 (Figure 2). The mean posterior probability

284   of site occupancy in 2016 (ψ_1) was 0.52, with 95% of the posterior observations in the

285   credible interval (95% CRI) 0.51 to 0.56. In 2017, ψ_2 was estimated as 0.51 (95% CRI:

286   0.50 to 0.56). In terms of number of sites, the model estimates that in 2016 there were

287   about 99 (95% CRI: 97 to 107) occupied sites and in 2017 there were about 97 (95%

288   CRI: 94 to 106) occupied sites. Thus, the number of sites occupied by tidewater goby was

289   essentially unchanged between the two years surveyed. The naïve rangewide occupancy

290   rates were lower than the model estimates: 0.44 in 2016 and 0.41 in 2017, or 83 occupied

291   sites in 2016 and 78 occupied sites is 2017 (Figure 2). The posterior distribution for the

292   effect of vegetation shows strong support (99.9% above zero) that probability of

293   occupancy at a site increases as the proportion of sample locations with vegetation

294   increase (Figure 3); whereas there is no evidence (53.2% above zero) that salinity has an

295   effect on site occupancy (Figure 4, Table 2). Probability of site occupancy increased from

296   0.42 (95% CRI: 0.27 to 0.58) at sites with no vegetation to a maximum of 0.68 (95%

297   CRI: 0.55 to 0.79) at sites where vegetation was present at all sample locations (Figure

298   3).

299   The lack of change in proportion of occupied sites between years was not

300   indicative of an absence of metapopulation dynamics. Rather, both extinction and

301   colonization rates were predicted nearly equal rates (Figure 5). The mean of the posterior

302   distribution of the probability of an extinction of an occupied site was 0.11 (95% CRI:

303   0.03 to 0.22). In terms of number of sites, this equates to an extinction at about 10 (95%

304   CRI: 3 to 19) occupied sites. There was no evidence that extinction varied substantially

305   with any of the tested covariates. The extinction-covariate relationships were examined

306   and their posterior distributions were only moderately above or below zero: salinity

307   (75.8% below zero), temperature (53.1% below zero), dissolved oxygen (53.1 below

308   zero), and estuary size (54.8% above zero, Table 2). I examined two additional

309   covariates, the change in salinity at a site from 2016 to 2017 and the absolute value of

310   that change as potential covariates of extinction, but inclusion of these covariates

311   prevented the model from converging, thus they were excluded. Results for the naïve,

312   rangewide extinction rate between the two years was 0.29. This is the equivalent of 24

313   sites occupied in 2016 that were unoccupied 2017. This extinction rate and number of

314   extirpated sites were higher than the model estimates (CRI: 3 to 19, Figure 5).

315   The mean of the posterior distribution of the probability of colonization of an

316   unoccupied site was 0.09 (95% CRI: 0.02 to 0.16), or about 8 (95% CRI: 2 to 13) sites

317   being colonized. There was strong posterior support (88.6% below zero) for the

318   relationship between probability of colonization within a sub-unit (ω) and the distance

319   between neighboring sites; as the distance between sites increases, colonization within

320   sub-units decreases (Figure 6). Colonization rates for unoccupied sites increased with the

321   number of occupied neighbors (Figure 7); the increase in probability per occupied

322   neighbor is a function of the equation $1 - (1 - x)^n$ where x is the colonization rate per

323   neighbor (0.01), and n is the number of occupied neighbors. I explored the possibility of

324   using the distance between sites that was rocky coast as a covariate of colonization as

325   rocky coast is believed to be a hindrance to tidewater goby colonization, but its inclusion

326   prevented the model from converging properly, causing poor model fit. The naïve

327   conditional rangewide colonization rate (0.18) was biased high when compared to the

328   model estimate of colonization; this naïve rate equates to 19 of the unoccupied sites in

329   2016 being occupied by 2017, but this naïve estimate was outside of the model's credible

330   interval of the predicted number of colonizations (Table 3).

<p style="text-align:center">Recovery Units</p>

331

332   The probability of change in occupancy between years was relatively low in most

333   recovery units: four out of the six units had no significant change in occupancy while one

334   recovery unit (Los Angeles/Ventura Units) experienced a significant increase and one

335   (Conception Unit) experienced a significant decrease (Appendix B Figure 12, Table 3).

336   There was no evidence of change in occupancy in the North Coast, Central Coast, and

337   South Coast Units (Appendix B Figure 11, Table 3). The posterior distribution of the

338   probability of change in occupancy in Greater Bay Area was negatively skewed with

339   75.8% of the posterior distribution being less than zero, suggesting a possible decline in

340   the number of occupied sites between the two years of sampling. The Conception unit

341   saw a decrease in occupancy probability (95% CRI: -0.30 to 0.15). Los Angeles/Ventura

342   recovery unit exhibited a significant increase in the probability of occupancy (95% CRI: -

343   0.29to 0.44). Model estimates of change in occupancy within recovery units were not

344   indicative of extinction and colonization dynamics within recovery units (Appendix B

345   Figure 11, Table 3). All recovery units, regardless of the amount of occupancy change

346   within the unit, experienced some level of extinction and colonization. There was no

347 geographic pattern or gradient seen in the model estimates of metapopulation dynamics

348 between the recovery units across the tidewater goby range. Naïve estimates of extinction

349 and colonization within the recovery units were biased high in most recover units

350 (Appendix B Figure 13 and Figure 14).

<div align="center">Environmental DNA</div>

352   Model estimates of detection of eDNA in a water sample (θ) and in qPCR

353 replicates (ρ) were informative of probability of detecting tidewater goby using eDNA

354 methods, as well as which covariates affected those detection probabilities (Figure 8).

355 Specifically, detection in a water sample was affected by salinity, dissolved oxygen, and

356 if a site was open to the tide, whereas turbidity and temperature did not significantly

357 influence eDNA detection in water samples (Figure 8). Given the presence of tidewater

358 goby eDNA at a site, the probability of detecting tidewater goby eDNA in a water sample

359 as 0.76 (95% CRI: 0.66-0.84). The posterior distribution of the covariate dissolved

360 oxygen strongly supports the conclusion that detection in a water sample increases with

361 dissolved oxygen (94.4% above zero; Figures 8 and 9). There was strong evidence, 100%

362 of the posterior distribution below zero, that presence of tide at a site reduced the

363 probability of detection in a water sample (Figures 8 and 9). Likewise, there was strong

364 evidence (98.9% below zero) that detection in a water sample decreased significantly

365 with increasing salinity; an increase in salinity of 20 ppt results in a decrease in the

366 probability of detection in a water sample of between 0.11 and 0.12 depending on

367 whether a site is open to the tide or not (Figure 10). There was no evidence of effect of

368 turbidity (60.6% above zero) or temperature (67.8% below zero) on detection in water

369   samples. Detection in a qPCR replicate given presence in a water sample (ρ) was strongly

370   impacted by salinity. Given the presence of tidewater goby eDNA in a water sample, the

371   probability of qPCR detection was 0.59 (95% CRI: 0.56-0.63); there was strong evidence,

372   that this probability decreased with increasing salinity (100% of the posterior distribution

373   being less than zero; Figure 11).

374

375   DISCUSSION

376   Occupancy and Dynamics

377   Model results indicated that rangewide occupancy was stable across the two

378   survey years, and that the balanced rates of extinction and colonization led to no

379   significant loss or gain in the number of occupied sites. This suggests that for the single

380   transition analyzed here, tidewater goby were at a dynamic equilibrium. For a species of

381   conservation concern, like tidewater goby, these results may be considered promising as

382   they show no sign of decline between these two years, but these results must be

383   considered in light of the fact that they are limited in temporal scope, representing only

384   two years of data and capturing only a single transition; I do not expect that the

385   occupancy, extinction, and colonization rates found here would continue beyond the time

386   period in which they were found.

387   The two years in which field work for this project occurred coincided with a

388   substantial climatic transition in California; 2016 and several years before it are

389   considered dry or drought years, with 2011 to 2016 being the driest years on record for

390   California. In contrast, the winter of 2016/2017 was exceptionally wet, breaking rainfall

391   records across the state (Wang et al. 2017). It has been predicted that tidewater goby

392   occupancy would be negatively impacted with drought conditions causing site to dry

393   whereas increased rainfall and river flow might promote increased colonization (Lafferty

394   1999a, 1999b). Based on these observations, it might be hypothesized that over my study

395   period I would find high colonization rates and low extinction rates rather than the more

396   comparable rates found. Transient time, or the time it takes for population dynamics to

397   return to equilibrium following a perturbation, is typically longer when perturbations are

398   strong, species are close to their persistence threshold, species have slow turnover, or

399   when a site network is composed of only a few important sites (Ovaskainen and Hanski

400   2002). If the prolonged drought conditions represent a significant perturbation to

401   tidewater goby metapopulation equilibrium, it is unlikely that the relative stability found

402   in my results represents rangewide stability that could be expected to continue into the

403   future. It is more likely that rangewide, tidewater goby dynamics are recovering from a

404   perturbation and not yet fully stabilized, and so their extinction and colonization rates are

405   likely to vary in future years.

406   The rangewide spatial coverage of this research has allowed us to examine

407   metapopulation dynamics within the recovery units delineated by the tidewater goby

408   recovery plan (USFWS 2005). Occupancy, extinction, and colonization for the recovery

409   units were variable and tended not to reflect the rangewide model estimates precisely,

410   suggesting minor variation in dynamics by region or recovery unit. While occupancy,

411   extinction, and colonization varied across recovery units, they did not appear to follow

412   any geographic cline or environmental gradient along the California coast. However, my

413   results emphasize the importance of the spatial structure of localized patches; the

414   probability of colonization was higher within sub-units when occupied and unoccupied

415   sites are closely spaced. Lafferty (1999b) suggested that tidewater goby might exist in a

416   core-satellite structure and my results appear to support this hypothesis; colonizing

417   individuals are more likely to originate from occupied neighbors, as opposed to distantly

418   located occupied sites.

419   <u>Covariates</u>

420   The results from this analysis support the assertion of Armstrong (2005) that

421   incorporation of both the metapopulation paradigm and the habitat paradigm can provide

422   robust and detailed results that could be of more use to managers than studies conducted

423   under only one of these paradigms. Within the metapopulation paradigm, the results of

424   this study outline the importance of occupied site density; I show that probability of

425   colonization for an unoccupied location decreases rapidly as the distance from an

426   occupied neighbor increases (Figure 6), and that colonization probability can be higher

427   within a recovery sub-unit where sites are closely spaced, as opposed to a rangewide

428   colonization process. These results are indicative of the well-supported premise in

429   metapopulation ecology that colonization is higher between more closely located habitat

430   sites as opposed to isolated sites (Hanski 1999) and suggest metapopulation resiliency in

431   areas where occupied sites are proximally located. The results showing a positive

432   correlation between tidewater goby occupancy and increased vegetation at a site fall

433   under Armstrong's (2005) habitat paradigm. Past research on tidewater goby has

434   repeatedly found a positive association between tidewater goby occupancy and the

435   presence of aquatic vegetation (McGourty 2006, Swenson 1999, Worcester 1992).

436   The failure to find a significant relationship between site level environmental

437   covariates (salinity, dissolved oxygen, and temperature) and extinction might be

438   attributed to the limited scope of the environmental data used in this analysis. The

439   habitats surveyed in this study are dynamic systems where environmental conditions can

440   be subject to hourly, daily, and seasonal fluctuations. The two surveys used here

441   measured each of these covariates once per water sample collected at a site, and most

442   sites had all water samples from a single year collected in a single day. Thus, the

443   recorded environmental conditions did not capture data that might be informative of a

444   metapopulation dynamic process that occurs at an annual scale. Use of covariates that

445   summarized the annual trends in these conditions at each site might have yielded

446   significant results. Unfortunately, most of the sites surveyed here are small, poorly

447   studied systems where longer term or continuous environmental monitoring data is not

448   available.

449   Imperfect detection

450   The results of the dynamic multiscale occupancy model reinforce the findings of

451   previous research emphasizing that failure to account for non-detection results in biased

452   estimates of metapopulation dynamics (MacKenzie et al. 2002, MacKenzie 2003,

453   Moilanen 2002). Rangewide, naïve occupancy estimates were consistently low compared

454   to model estimates. At both the rangewide and recovery unit scale, naïve estimates of

455   extinction and colonization were consistently biased high compared to model estimates,

456   indicating extinction and colonization would be overestimated unless non-detection

457   errors are unaccounted for, consistent with the findings of Moilanen (2002).

458   In instances where the informed and naïve estimates of extinction or colonization

459   were similar, naïve estimates were zero and model estimates were low. This occurred in

460   two recovery units: Los Angeles/Ventura and Conception. The data from Los

461   Angeles/Ventura show no extinctions between the two years surveyed, and there was no

462   evidence of colonizations in the Conception recovery unit. Because of this, the naïve

463   estimates for these recovery units cannot be higher than zero. But, by accounting for non-

464   detections in the analysis, the model is able to detect a probability of extinction and

465   colonization in both of these recovery units. Once accounted for, detection probability

466   highlights the possibility of occupancy at sites where target species were not detected, as

467   well as the possibility of extinctions or colonizations in sites.

468   Three of the six recovery units (North Coast, Central Coast, and South Coast) had

469   model estimates of occupancy change (from 2016 to 2017) similar to their naïve

470   estimates. In these three recovery units, the number of occupied sites in the survey data

471   was the same for both 2016 and 2017. While the change in proportion of occupied sites

472   from one year to the next can be informative in an analysis of metapopulation dynamics,

473   this particular metric may not convey the details in site occupancy changes from one

474   season to the next. For instance, both the naïve estimate and the model results may

475   similarly indicate relatively low changes in the proportion of occupied sites between the

476   two seasons, but this may not be reflective of the extinction and colonization dynamics.

477   This phenomenon is evident in these three units where both model and naïve estimates

478   indicate that site occupancy was the same or similar between 2016 and 2017.

479   Examination of the extinction and colonization estimates within these recovery units

480 demonstrates that each of these recovery units experienced at least one extinction,

481 colonization, or multiple instances of both. The loss of resolution when examining

482 interannual occupancy on its own could lead to extinctions and colonization dynamics

483 going unnoticed. This may be of particular concern for an endangered species, or any

484 species managed under a metapopulation paradigm, where the density of occupied sites is

485 of particular importance.

486 <center>Implications for eDNA surveys</center>

487 Model results of detection probabilities in water samples and qPCR replicates are

488 useful in informing future surveys of how environmental conditions might affect their

489 ability to detect eDNA. The results from this analysis highlight the need to incorporate

490 some previous knowledge about a study system into sampling design; I show that

491 detection probability in water samples decreases at sites that are open to the tide and at

492 sites with higher salinity. Future surveys may need to account for tidal conditions and site

493 level salinity when determining how many water samples need to be collected at a site in

494 order to consider it thoroughly surveyed. Using the detection probability from these

495 results ($\hat{\theta} = 0.76$), before accounting for salinity, based on the equation $1 - (1 - \hat{\theta})^n =$

496 0.95, three water samples (n) need to be collected to have a detection probability greater

497 than 0.95. Factoring in salinity and presence of tide would increase the number of water

498 samples needed to reach the same level of detection. Detection in qPCR replicates yields

499 similar results. Using the same formula, $1 - (1 - \hat{\rho})^n = 0.95$, before accounting for

500 salinity, given eDNA presence in a water sample and a $\hat{\rho}$ value of 0.593, there needs to be

501   at least four qPCR replicates to have greater than a 95% chance of detection. Again, this

502   number could increase as salinity increases due to increase in $\hat{\rho}$

503   Management Implications

504   The results from the research suggest that there is no latitudinal gradient in

505   tidewater goby metapopulation dynamics. The lack of extinction and colonization signal

506   in some northern population segments noted by Kinziger et al. (2015) may be a product

507   of the core-satellite population structure of tidewater goby; genetic samples used

508   Kinziger et al. (2015) may have originated from stable core populations, as opposed to

509   less stable satellite populations, and so would offer no evidence of recent founder events

510   or bottlenecks. Future genetic work may offer some insight into tidewater goby

511   metapopulation dynamics by targeting recently re-colonized area for which historic

512   samples exist.

513   This research demonstrates several tools that would be useful in future tidewater

514   goby monitoring and conservation. Environmental DNA survey techniques and dynamic

515   multiscale occupancy modeling could serve as a foundation for conduction a

516   metapopulation viability analysis as required by the species recovery plan. The

517   colonization and extinction rates found here represent metapopulation dynamics for two

518   years of data and one transition and are unlikely to characterize future transitions;

519   however, if the rates found here remained consistent these estimates could be

520   extrapolated across longer time scales using the formula $1 - (1 - x)^n$ where x is the rate

521   in question, extinction or colonization, and n is the number of years. Examining the

522 dynamic rates found here across a decadal scale shows that individual occupied sites may

523 have a high probability of extinction, 0.67, and that unoccupied sites have a high

524 probability of colonization, 0.59. At the 100-year mark, there is 99.9% probability of both

525 colonization of unoccupied sites and extinction of occupied sites. These simplistic

526 estimates do not take into account the presumed core-satellite structure, assume that all

527 sites have similar extinction or colonization probabilities, and come from data that is

528 limited in scope; with further monitoring and the incorporation of additional years of

529 data, these probabilities could be tailored to and predicted for sites within recovery units

530 and offer a higher resolution view of the metapopulation viability within recovery units.

531 Additionally, inclusion of data from more time periods with varying climate and

532 environmental conditions might offer insight into the environmental drivers of extinction

533 and colonization not found here, and could act as a basis for modeling tidewater goby

534 metapopulation dynamics into the future as functions of climate change and larger scale

535 weather patterns such as the El Ninõ/Southern Oscillation or the Pacific Decadal

536 Oscillation.

537   The two relationships found here, between (1) occupancy and vegetation and (2)

538 colonization and distance between sites, highlight the underpinnings of Armstrong's

539 (2005) call for incorporation of metapopulation paradigms and habitat paradigms into

540 management and conservation. My results suggest that in order to best conserve tidewater

541 goby, management may need to incorporate themes of both the metapopulation paradigm;

542 (e.g., distance between sites) as well as the habitat paradigm (e.g., quality habitat with

543 presence of aquatic vegetation). For example, restoration efforts that include

544 consideration of aquatic vegetation would have little impact on tidewater goby

545 populations if colonization rates for the restored habitat are low due to large gaps of

546 unoccupied locations along the coast. Likewise, increasing the number of occupied sites

547 in those gaps of unoccupied locations by planting tidewater goby may not lead to

548 permanent occupancy if the habitat is not suitable.

549   In this study, I modeled two years of eDNA occupancy data using a dynamic

550 multiscale occupancy model that explicitly accounts for non-detection at three levels:

551 site, sample, and in qPCR detection. The model developed for this analysis was used to

552 examine the probabilities of occupancy, extinction, and colonization of tidewater goby at

553 two scales: rangewide and in distinct management units. Additionally, the model allowed

554 testing of spatial and environmental covariate effects on occupancy, extinction, and

555 colonization, as well as covariate effects on eDNA detection at both the water sample and

556 qPCR replicate level. This study design and associated dynamic multiscale occupancy

557 model are widely applicable and likely useful in situations where eDNA survey

558 techniques or a hierarchical sampling design are an option. The increasing use of eDNA

559 will likely provide opportunities to apply this model over increasing temporal scales in an

560 increasing number of ecological settings to aid in monitoring and conservation of

561 endangered species like the tidewater goby.

562

564
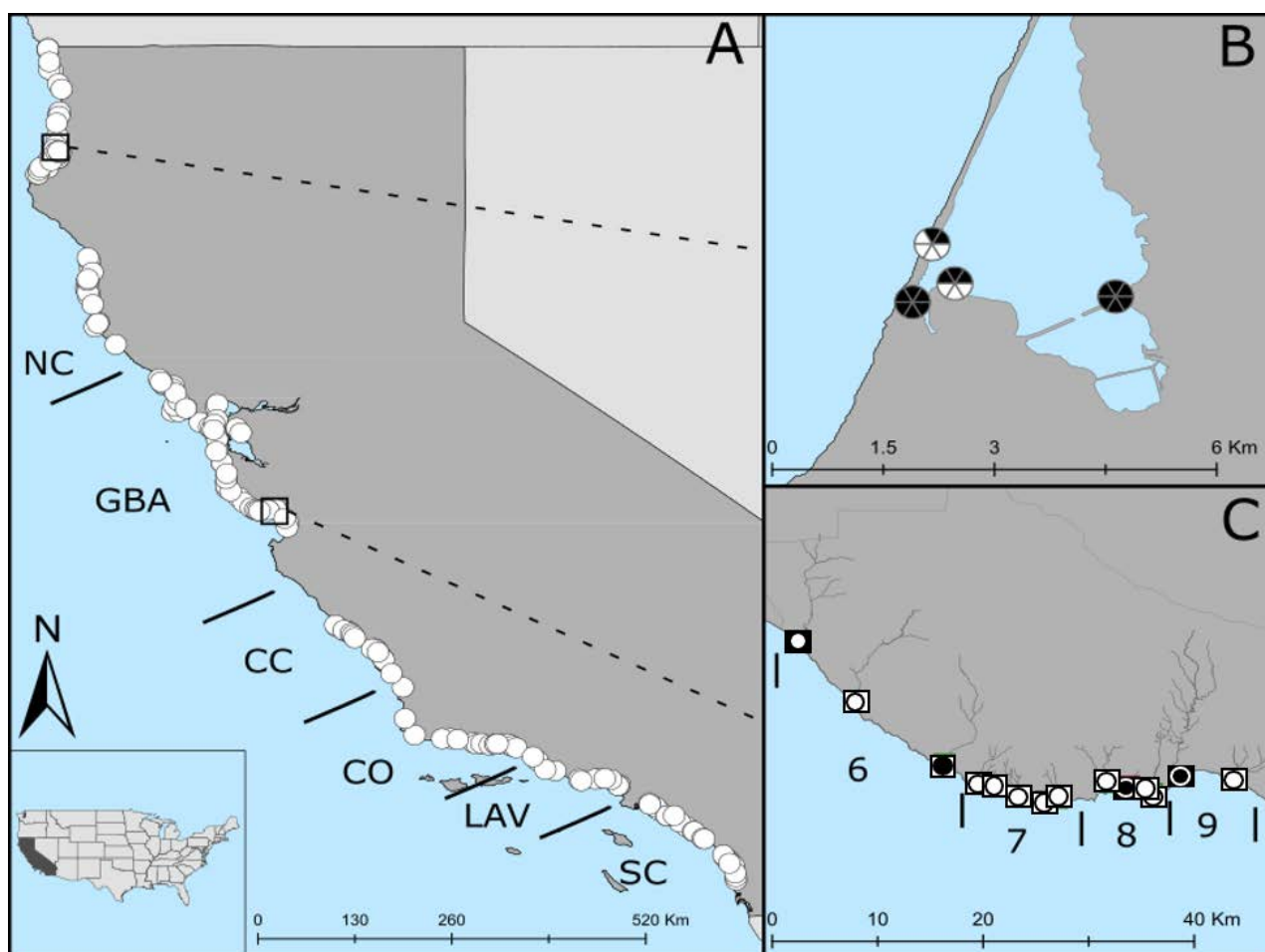
Figure 1 Tidewater goby sampling locations. The hierarchical nature of eDNA sampling is demonstrated by the transition from panel A, showing the 190 sites surveyed for tidewater goby along the 1350 km California coast, to panel B where each pie represents the location of individual water samples collected 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

579 used in the dynamic multiscale occupancy model. Square symbols represent sites
580 surveyed in 2016 and the circles represent sites surveyed in 2017. White indicates a
581 detection in at least one qPCR reaction at a site whereas indicates non-detection at
582 a site at all water samples and qPCR replicates.
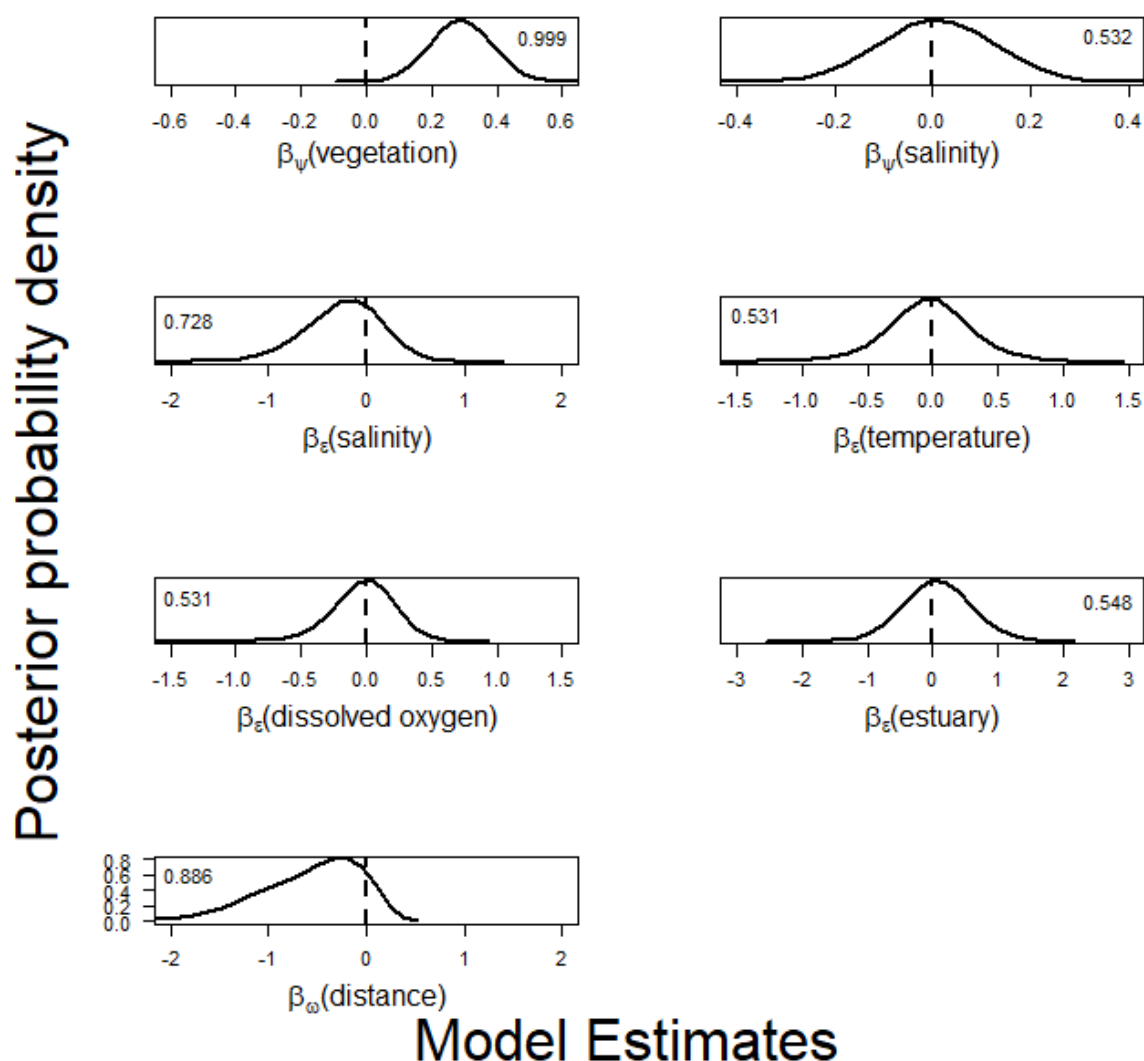
583

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

589
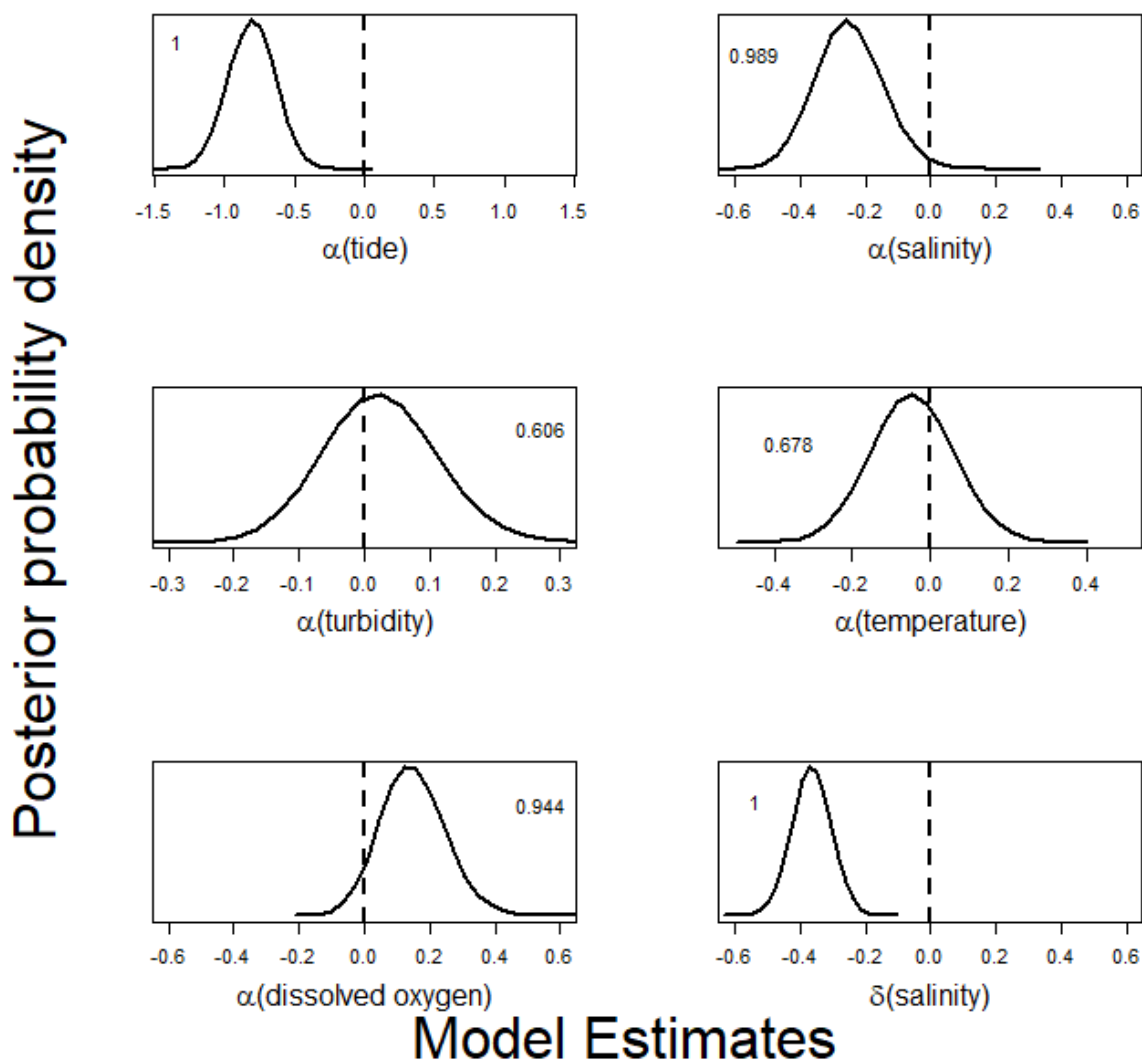
Figure 3 Posterior probability densities for covariates of detection in a water sample (α) and detection in qPCR replicates (δ) that were examined as part of an analysis of two years (2016 and 2017) of rangewide eDNA occupancy data of tidewater goby *Eucyclogobius spp*. Data was analyzed using a dynamic 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).

597
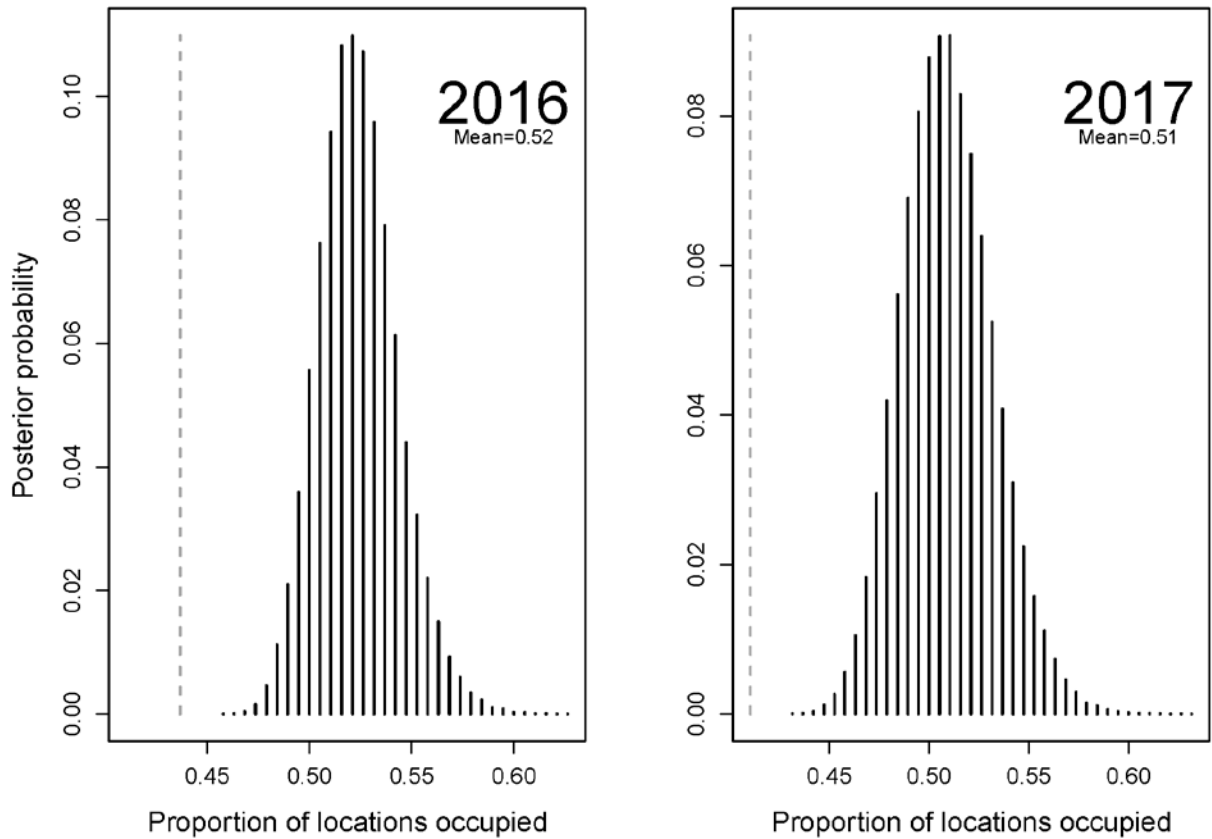
Figure 4 Posterior distribution of tidewater goby *Eucyclogobius spp.* rangewide occupancy. The proportion of occupied sites were estimated using dynamic multiscale occupancy model to analyze eDNA survey data for 2016 (left) and 2017 (right). The dashed line represents the naïve occupancy rate.

602
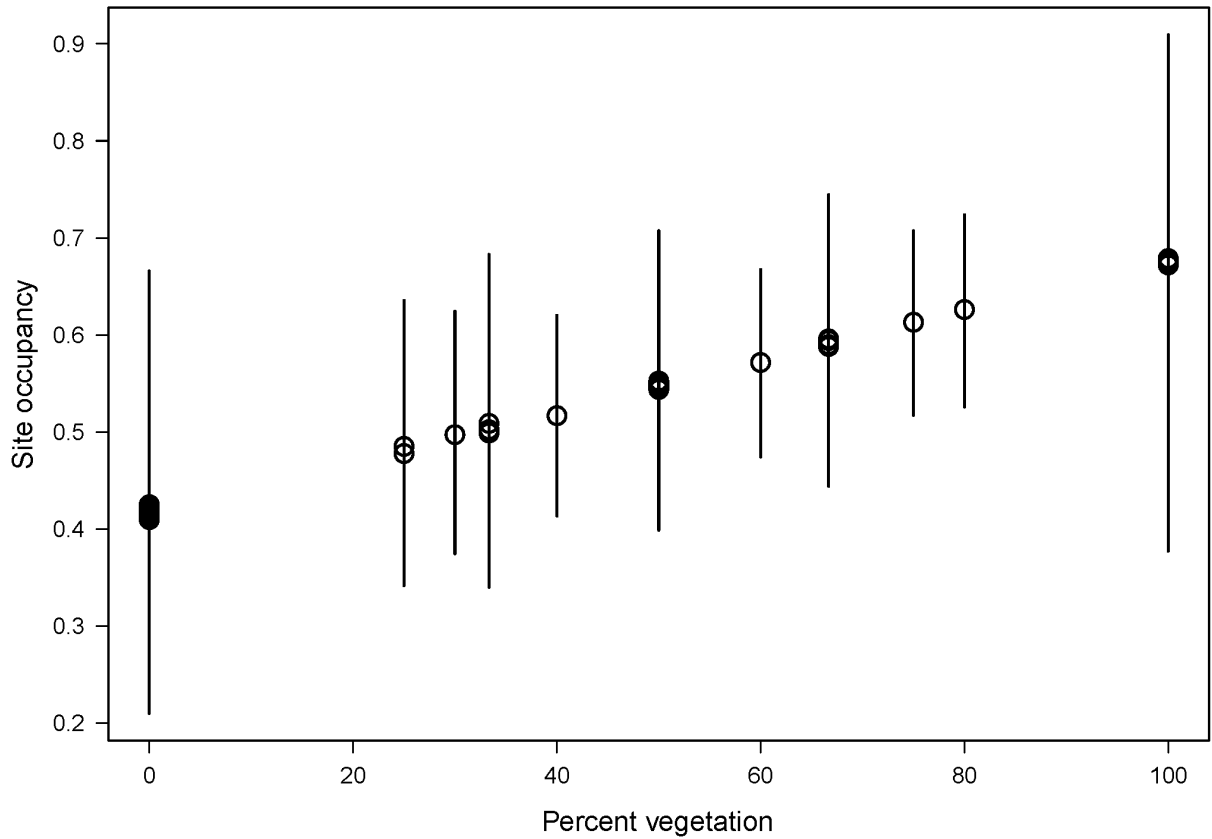
Figure 5 Effects of vegetation levels on site occupancy. The presence of aquatic vegetation was tested as a covariate of tidewater goby *Eucyclogobius spp.* occupancy at a site (ψ) in an analysis of two years, 2016 and 2017, of eDNA survey data using a multiscale occupancy model. The *x*-axis is the 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.

609
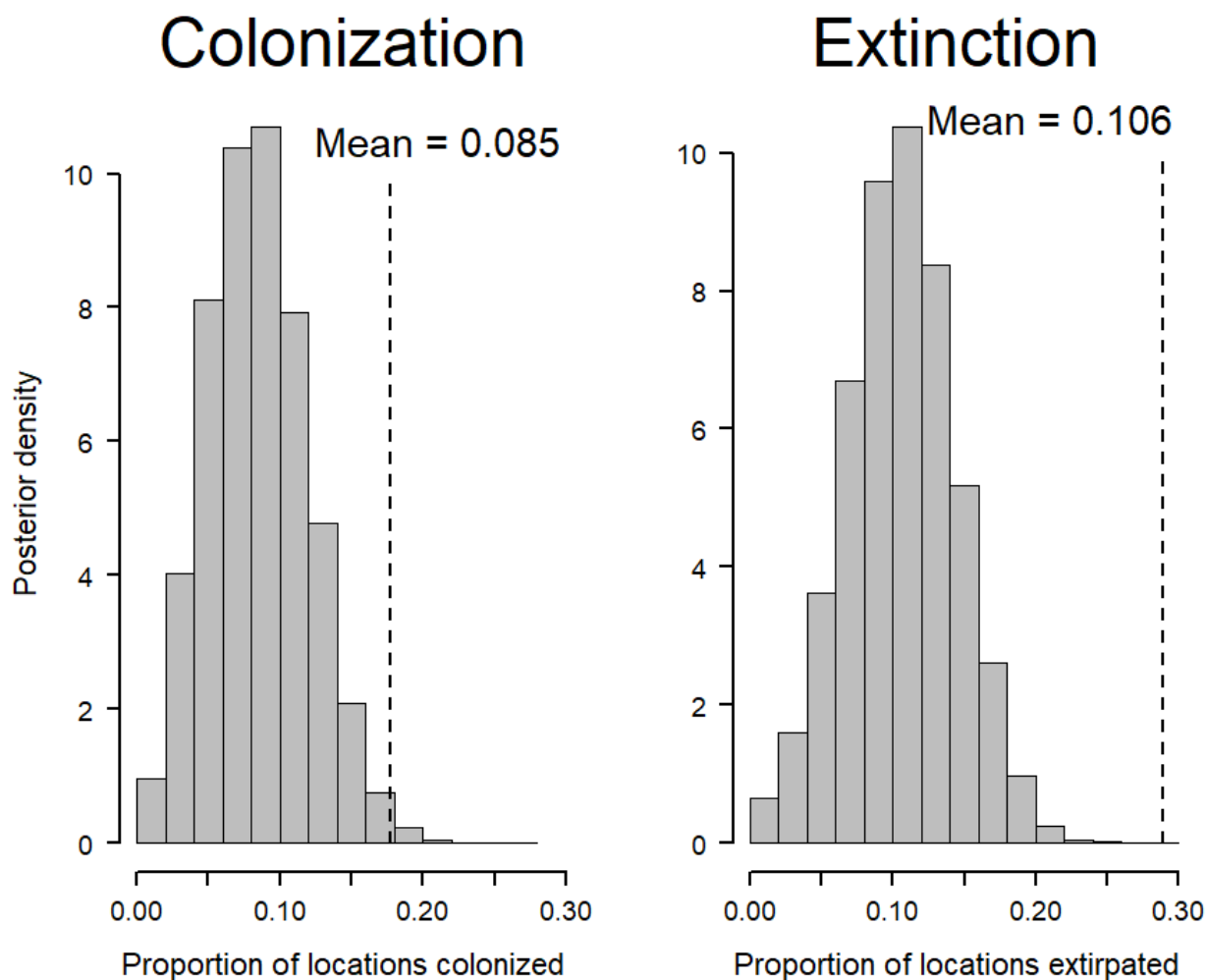
610

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

617
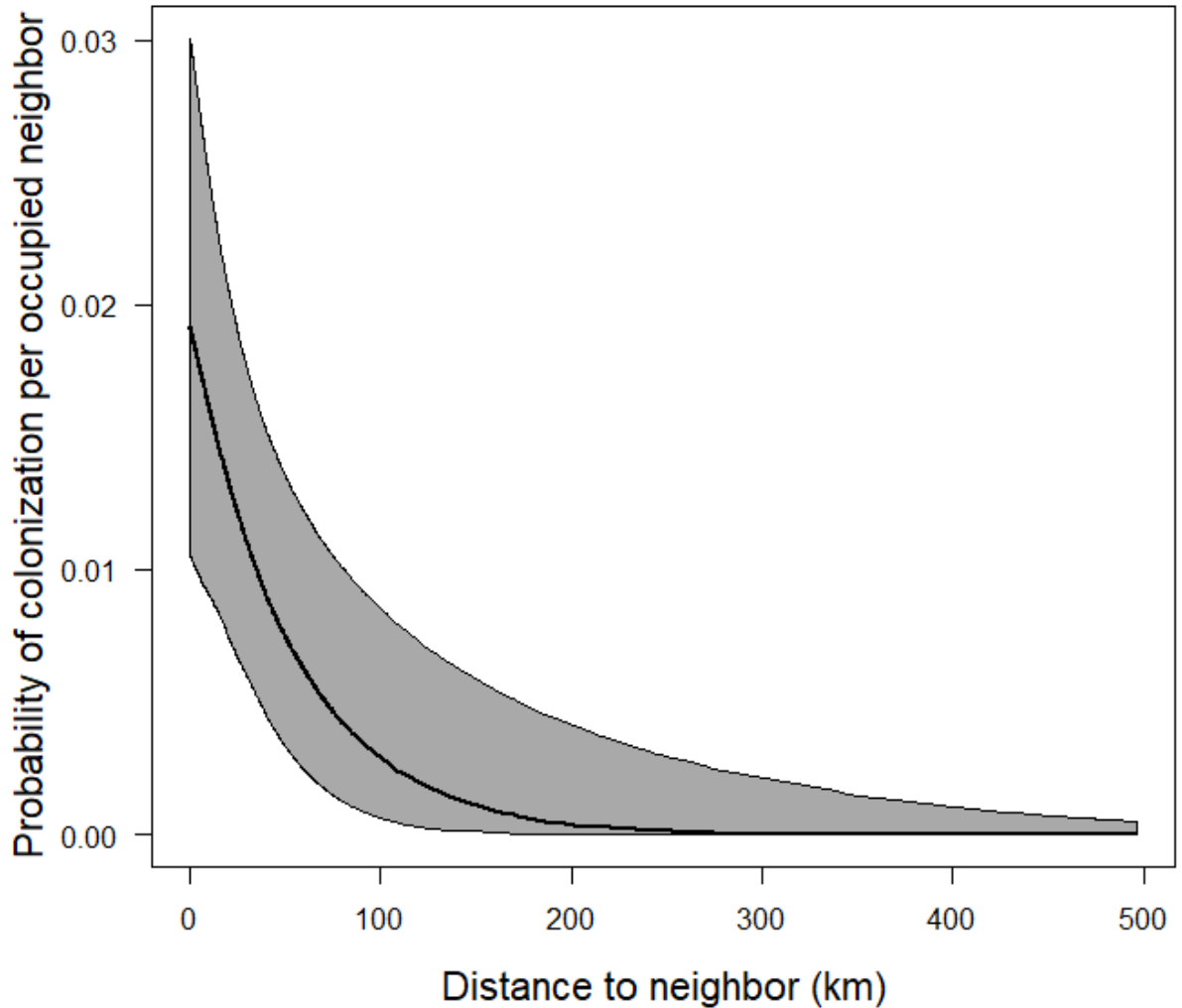
Figure 7 **Plot of the shoreline distance between neighbors versus the probability of**
colonization of an unoccupied site from an occupied neighbor (ω). Analysis was
done using a dynamic multiscale occupancy model to analyze two years, 2016 and
2017, of tidewater goby *Eucyclogobius spp.* eDNA occupancy data. The shaded area
represents the 95% credible interval.

623
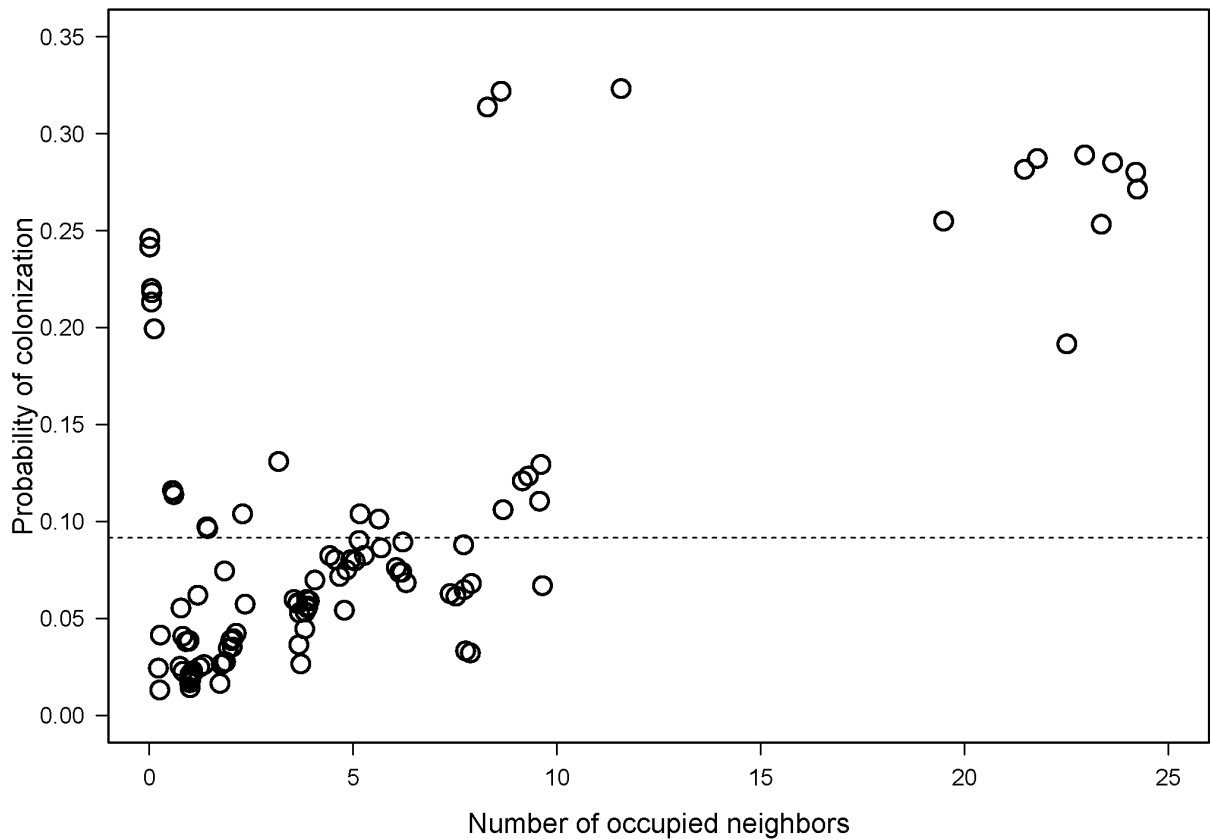
Figure 8 Probability of colonization of unoccupied sites versus the number of occupied neighbors.   Colonization at sites with occupied neighbors is only a process of localized, neighborhood colonization (ω), while sites with no occupied neighbors are subject to the rangewide colonization rate (η). Probability 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.

632

633
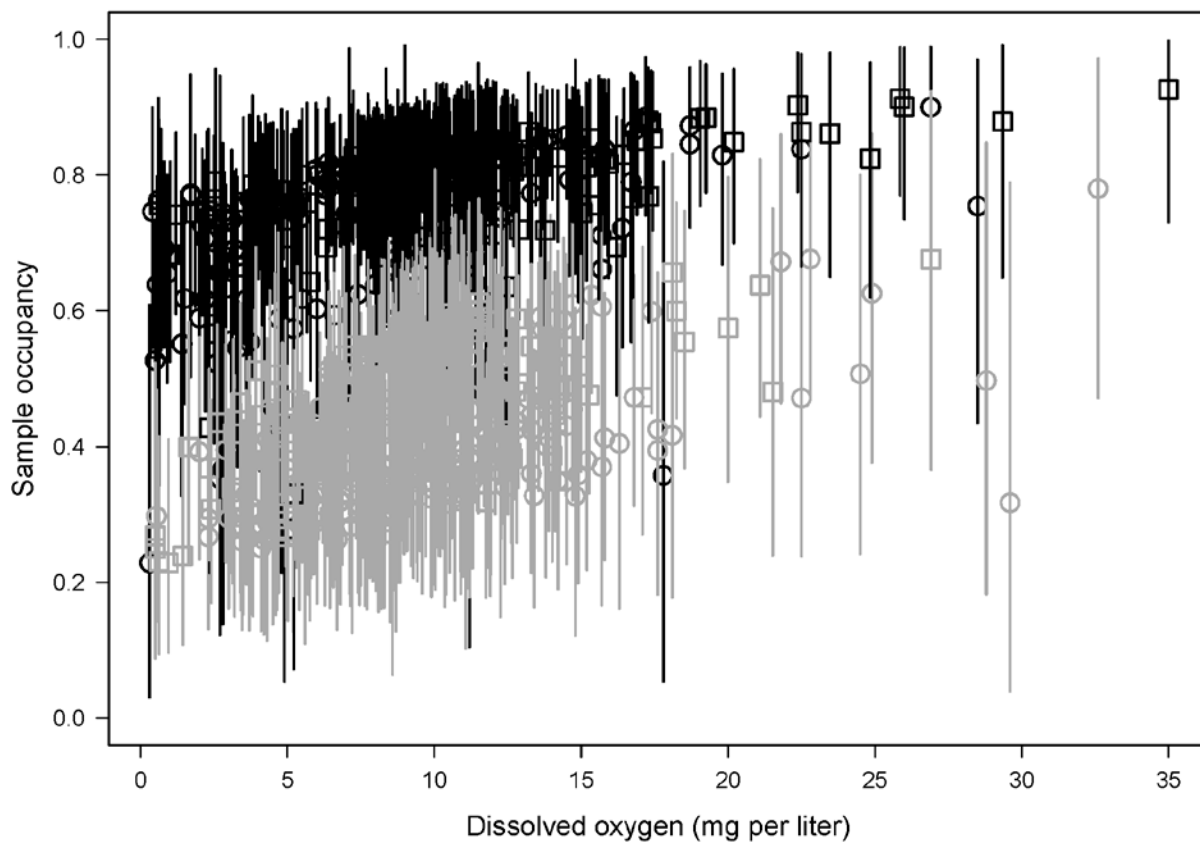
Figure 9 Effect of dissolved oxygen and tidal presence on eDNA presence (with 95% credible intervals) in a water sample. Water samples with tidal influence are in grey, water samples without tidal influence are in black. Water samples were collected in 2016 (squares) and 2017 (circles) as part of a rangewide eDNA survey of tidewater goby *Eucyclogobius spp.* and analyzed using a dynamic multiscale occupancy model.

639
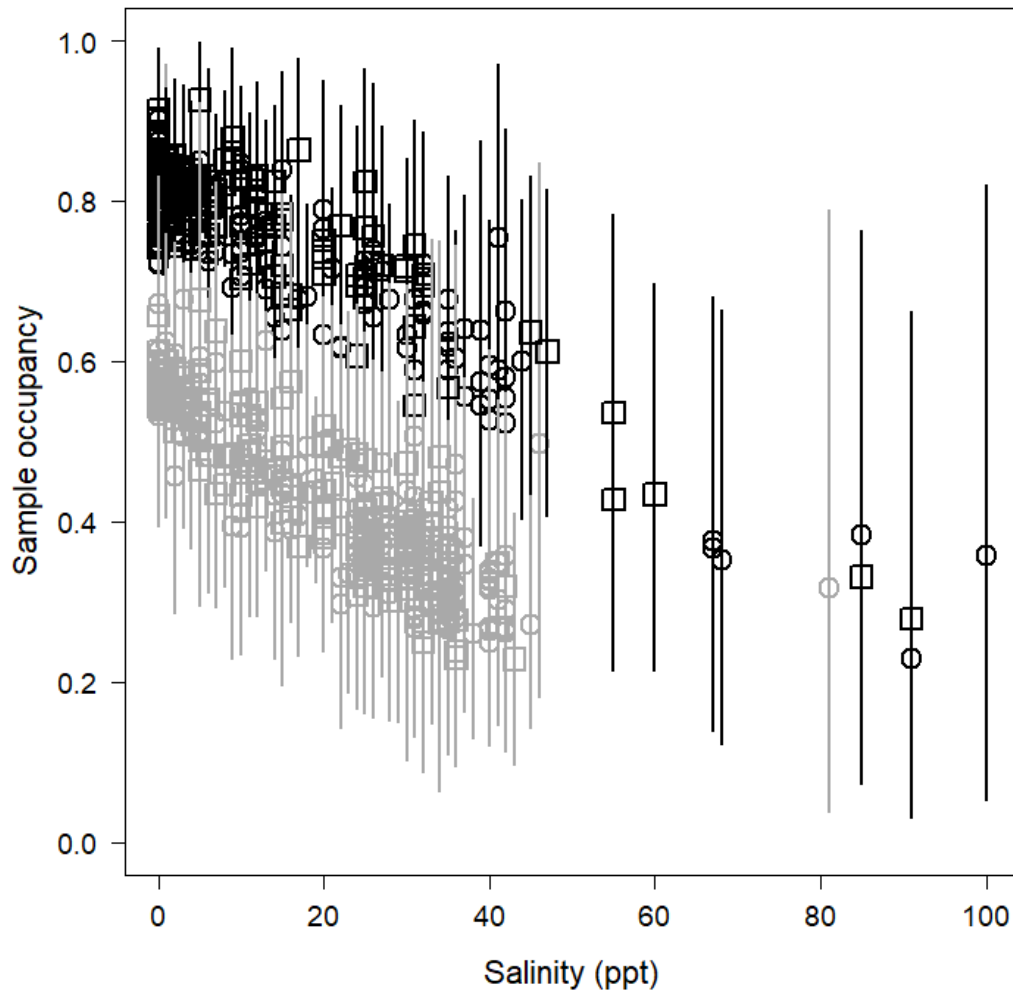
Figure 10 Effect of salinity and tidal presence on eDNA presence (with 95% credible intervals) in a water sample. Water samples with tidal presence are in grey, water samples without tidal presence are in black. Water samples were collected in 2016 (squares) and 2017 (circles) as part of a rangewide eDNA survey of tidewater goby Eucyclogobius spp. and analyzed using a dynamic multiscale occupancy model.
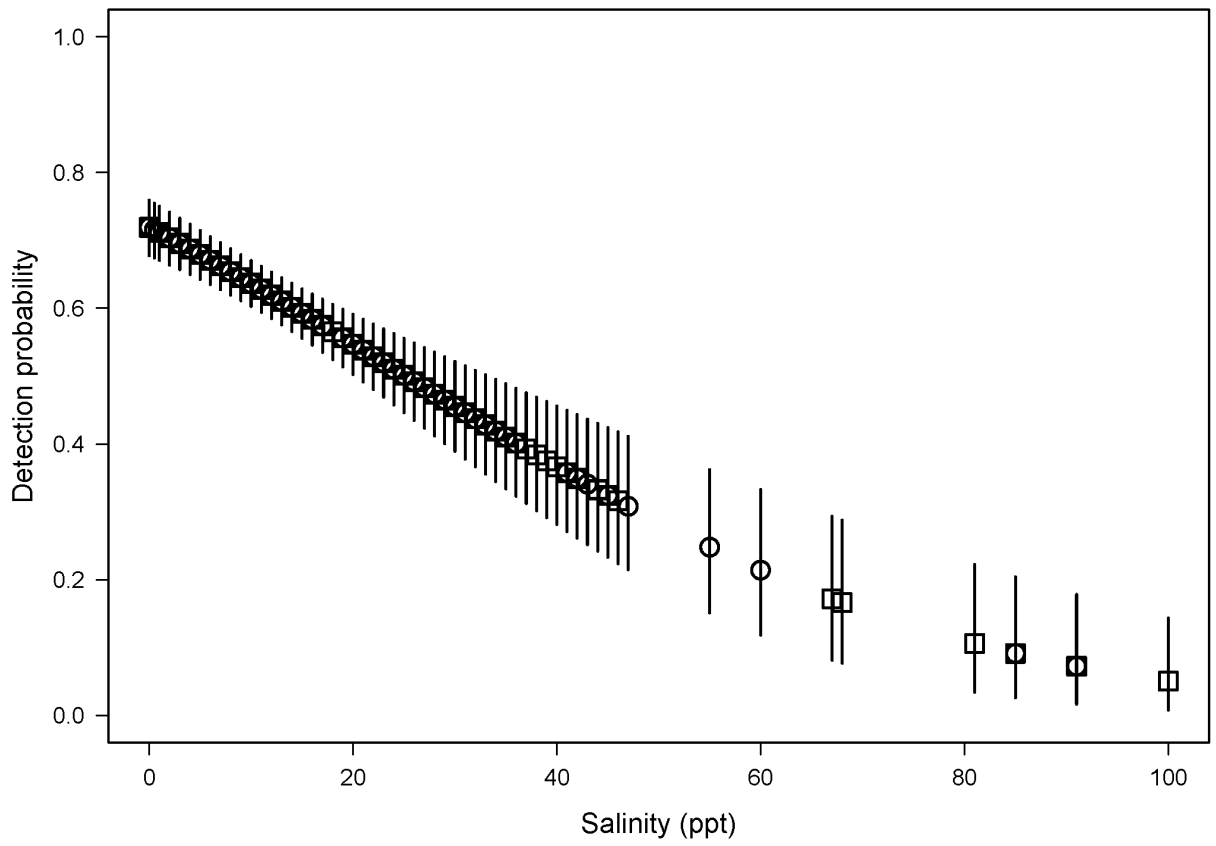
646 **Figure 11 Estimates of probability of detection of eDNA (with 95% credible**
647 **intervals) in qPCR replicates from water samples of with varying salinities. Water**
648 **samples were collected in 2016 (squares) and 2017 (circles) as part of a rangewide**
649 **eDNA survey of tidewater goby *Eucyclogobius spp*. Data were analyzed using a**
650 **dynamic multiscale occupancy model that accounts for non-detection at site, water**
651 **sample, and qPCR replicate levels.**

652 **Table 1 The dynamic multiscale model allows for the incorporation of covariates that may drive patterns in tested parameters.**
653 **Here I present the covariates tested while analyzing tidewater goby *Eucyclogobius spp.* occupancy data from two consecutive**
654 **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).

655

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

668 **Table 3 Comparison of naïve and model based estimates of colonization and extinction**
669 **rates for tidewater goby *Eucyclogobius spp.* across their range, separated by Recovery**
670 **Units (NC = north coast, GBA = greater bay area, CC = central coast, CO = Conception,**
671 **LAV = Los Angeles/Ventura, SC = south coast) as outline in the Tidewater goby Recovery**
672 **Plan (2005 USFWS). Occupancy change represents the change in poropotion of occupied**
673 **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

674

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.

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

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

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

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

726 Foote, A. D., P. F. Thomsen, S. Sveegaard, M. Wahlberg, J. Kielgast, L. A. Kyhn, A. B. Salling,
727   A. Galatius, L. Orlando, and M. T. P. Gilbert. 2012. Investigating the potential use of
728   environmental DNA (eDNA) for genetic monitoring of marine mammals. PLOS ONE
729   7(8):e41781.

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

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

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

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

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

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

745 Hanski, I. 1994. A practical model of metapopulation dynamics. The Journal of Animal Ecology
746   63(1):151.

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

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

750 Hanski, I., and M. Gilpin. 1991. Metapopulation dynamics: brief history and conceptual domain.
751   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.

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

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

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

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

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

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

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

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

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

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

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

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

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

825 Strickler, K. M., A. K. Fremier, and C. S. Goldberg. 2015. Quantifying effects of UV-B,
826   temperature, and pH on eDNA degradation in aquatic microcosms. Biological Conservation
827   183:85–92.

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

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

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

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

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

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

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

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

847   Vörös, J., O. Márton, B. R. Schmidt, J. T. Gál, and D. Jelić. 2017. Surveying Europe's Oonly
848   cave-dwelling chordate species (*Proteus anguinus*) Using Environmental DNA. PLOS ONE
849   12(1):e0170945.

850   Wang, S., A. Anichowski, M. K. Tippett, and A. H. Sobel. 2017. Seasonal noise versus
851   subseasonal signal: forecasts of California precipitation during the unusual winters of
852   2015-2016 and 2016-2017: S2S forecast of CA precipitation. Geophysical Research Letters
853   44(18):9513–9520.

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

857   Ying, Y., Y. Chen, L. Lin, and T. Gao. 2011. Risks of ignoring fish population spatial structure
858   in fisheries management. Canadian Journal of Fisheries and Aquatic Sciences 68(12):2101–
859   2120.

860   APPENDICES

861   APPENDIX A

862   Occupancy Analysis – Authored by Dr. Robert Dorazio

863   Multiscale occupancy models (Nichols et al., 2008; Mordecai et al., 2011) provide a

864   useful framework for the analysis of data collected in eDNA surveys. In fact, the hierarchical

865   relationships assumed between parameters of these models are exactly those induced by the

866   three-level, nested sampling design of eDNA surveys (Dorazio and Erickson, 2018). These

867   models have been used to analyze eDNA surveys of several species, including a fungal pathogen

868   (Schmidt et al., 2013), the Burmese python (Hunter et al., 2015), the tidewater goby (Schmelzle

869   and Kinziger, 2016, Sutter and Kinziger 2019), and a cave-dwelling aquatic salamander (Vörös

870   et al., 2017).

871   Multiscale occupancy models can be used to estimate the spatial distribution of a species

872   during a relatively short period of sampling. If surveys are repeated at sample locations (say,

873   annually or seasonally), changes in occupancy state, such as colonization of a previously

874   unoccupied location or extinction of an occupied location, can be inferred by analyzing each

875   sampling period's data separately. However, this approach does not allow potential drivers of the

876   processes responsible for changes in occupancy state to be identified. To solve this problem, we

877   propose a class of *dynamic* multiscale occupancy models in which temporal changes in

878   occupancy states are specified as a function of explicit colonization and extinction processes. In

879   the following section we first describe a versatile class of multistate occupancy models. We

880   follow this section by describing models of sampling and observation processes induced by the

881   sampling designs used in eDNA surveys.

882 <u>Multistate models of occupancy dynamics</u>

883   Suppose each of *M* survey locations is sampled during *T* disjoint (non-overlapping)

884   periods and that we wish to model changes in occupancy state at these locations between the

885   periods of sampling. Let $Z_{i,t}$ denote a random variable for the occupancy state of the *i*th survey

886   location ($i = 1, \ldots, M$) during the *t*th sampling period ($t = 1, \ldots T$). We assume that $Z_{i,t}$ can

887   have one of three values:

888   1 if the survey location is occupied,

889   2 if the survey location is unoccupied but has been occupied during the previous

890   sampling period, or

891   3 if the survey location is unoccupied and has not been occupied previously.

892   State 3 distinguishes locations that have never been occupied from those that have previously

893   been colonized but are temporarily unoccupied.

894   Our models of occupancy state dynamics include two distributional assumptions: one for

895   the initial occupancy state during sampling period 1

$$Z_{i,1} \sim \text{Cat}(\psi_i) \tag{1}$$

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

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

897   where Cat denotes the categorical distribution and where $t = 1, \ldots, (T - 1)$. In Eq. 1

898   $\psi_i = (\psi_i, 0, 1-\psi_i)'$ denotes a vector containing the probabilities of each occupancy state during

899   period 1. (We use the prime superscript to denote the transpose of a vector.) The parameter ψ_i is

900   the probability that the *i*th location is occupied during period 1 ($Pr(Z_{i,1} = 1)$). If the *i*th location is

901   not occupied during this period, $Z_{i,1}$ must equal 3 with probability $1-\psi_i$ because no previous

902   sampling has occurred.

903   We assume that the initial occupancy probability may be formulated as a function of

904   covariates whose values can vary spatially as follows:

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

905   where $x_{\psi,i}$ is a vector of regressors that codify the covariate measurements taken at the ith survey

906   location during sampling period 1, and where $\boldsymbol{\beta}_\psi$ is a vector of parameters that specify the effects

907   of the covariates on occupancy during sampling period 1. The function F provides a one-to-one

908   mapping of a real-valued argument to a real number that lies on the closed interval $(0, 1)$. We use

909   the cumulative distribution function (cdf) of a Gaussian distribution for F but others, such as the

910   cdf of a logistic distribution, also could be used without loss of generality.

911   In Eq. 2 the vector of probabilities $\boldsymbol{\phi}_{k,i,t}$ corresponds to the kth row of a matrix of

912   transition probabilities $\boldsymbol{\phi}_{i,t}$ that we now define. Transitions among occupancy states correspond

913   to *colonizations* of unoccupied locations and to *extinctions* of occupied locations. These

914   transitions are assumed to be time- and state-dependent, as indicated by the following matrix of

915   transition probabilities:

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

916

917   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

918   the interval between sampling periods t and $t + 1$; therefore, each row of $\boldsymbol{\phi}_{i,t}$ sums to one by

919   construction. For example, the first row includes probabilities of extinction $\varepsilon_{i,t}$ (a change from

920   state 1 to state 2) and persistence $1 - \varepsilon_{i,t}$ (remaining in state 1) when the ith location is occupied

921   during period t. Once a survey location is occupied it cannot change to state 3 (by definition), so

922   the third element of row 1 is zero. Similarly, the second and third rows of $\Phi_{i,t}$ include

923 probabilities of colonization when the ith survey location is unoccupied during sampling period t.

924 The third row contains the probability that this location is *first* colonized $\eta_{i,t}$ (a change from state

925 3 to state 1), whereas the second row contains the probability of recolonization $\gamma_{i,t}$ (a change

926 from state 2 to state 1) of the ith survey location.

927 This model of occupancy-state dynamics is extremely versatile. Each of the probabilities

928 of colonization and extinction may be formulated as functions of covariates whose values can

929 vary spatially and/or temporally. In addition, the effects of dispersal of individuals from

930 neighboring locations of each surveyed site can be used to specify colonization and extinction

931 probabilities.

932 In the model that we developed for tidewater gobies, the probability of extinction at a

933 survey location is specified as a function of location- and time-specific covariate values as

934 follows:

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

935 where $x_{\varepsilon,i,t}$ is a vector of regressors that codify the covariate measurements taken at the ith survey

936 location during sampling period t, and where $\boldsymbol{\beta}_{\varepsilon}$ is a vector of parameters that specify the effects

937 of the covariates on the probability of extinction. The probabilities of colonization are formulated

938 similar to that of extinction except that we also specify the effects of dispersal of individuals

939 from neighboring locations. To be specific, we assume that colonization of an unoccupied survey

940 location may occur from one of two processes depending on the occupancy states of the

941 location's neighbors. If none of the neighboring locations are occupied, we assume that

942 colonization rates are functions of location- and time-specific covariate values (that is, we adopt

943 the approach used to specify extinction probability). However, if one or more neighboring

944 locations are occupied, we assume that colonization occurs by the movements of individuals

945   from occupied neighbors, that is, colonization is a localized process . Let $\omega_{i,n}$ denote the

946   probability that the ith survey location is colonized by movements of individuals from a

947   neighboring location (indexed by n) during the period between sampling intervals t and $t + 1$.

948   The probability that the ith survey location is colonized by individuals from at least one of its

949   occupied neighbors during this time interval is

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

950   where $\mathcal{N}_i$ denotes the set of location indices corresponding to the neighbors of the ith survey

951   location and where I is an indicator function whose value equals one for a true argument and

952   zero otherwise. Note that the expression in Eq. 5 equals zero if none of the ith survey

953   location's neighbors are occupied (i.e., if $Z_{n,t} \neq 1$ for all $n \in N_i$). In other words, if a survey

954   location's neighbors are all unoccupied, we assume that local colonization's cannot occur and

955   that any colonization must stem from the other (non-local) colonization process. Figure 16

956   contains a diagrammatic illustration of our model's specification of the effects of dispersal from

957   occupied neighboring locations to an unoccupied survey location. In this figure only two of four

958   neighboring locations are occupied, so the probability of being colonized by at least one of the

959   four neighbors depends only on the colonization probabilities of the two occupied neighbors, as

960   specified in Eq. 5.

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

962   colonization processes as follows:

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

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

963

964

965   where $q_{i,t} = \sum_{n \in N_i} I(Zn, t = 1)$ denotes the number of neighbors of the ith survey location

966   that are occupied during period t. As noted earlier, the second parenthesized, $I(q_{i,t} = 0)$, term in

967   these equations equals zero whenever $q_{i,t}$ equals zero.

968   Our formulation of the local colonization process parameterized by $\omega_{i,n}$ is similar to the

969   approach described by Broms et al. (2016). This formulation provides considerable flexibility by

970   allowing different models of $\omega_{i,n}$ to be constructed. For example, in the simplest model we

971   assume $\omega_{i,n}$ is a constant (say, ω) that does not differ among survey locations. In this case the

972   expression in Eq. 5 simplifies to

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

973   where $q_{i,t}$ is the number of occupied neighbors of the ith survey location during sampling period

974   t. In this case the probability of colonization of the ith survey location is a mono- tone increasing

975   function of the number of occupied neighbors of that location. In a more complicated model, we

976   might assume that colonization between locations depends on physical features of the path

977   between locations (e.g., measures of connectedness) or on habitat gradients. In this case we could

978   specify $\omega_{i,n}$ as a function of these covariates as follows:

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

979   where $\boldsymbol{x}_{\omega,i,n}$ is a vector of regressors that codify the covariate measurements assumed to influence

980   colonization between the ith and nth survey locations and where $\boldsymbol{\beta}_\omega$ is a vector of parameters that

981   specify the effects of these covariates on $\omega_{i,n}$.

982   Models of eDNA occurrence and detection

983   Thus far, we have only modeled occupancy states in the first level of sampling in eDNA surveys

984   (i.e., at the location level). In most eDNA surveys multiple samples are collected at each

985   location, and any eDNA that may be present in each of these samples is detected using multiple

986   PCR replicates. Therefore, it necessary to model the presence of eDNA in samples and the

987   detection of eDNA in PCR replicates. Following Dorazio and Erickson (2018), we model the

988   occurrence and detection of eDNA in samples using nested conditional distributions.

989   Suppose $J_{i,t}$ samples are collected independently from the ith survey location during

990   period t. Let $A_{i,j,t}$ denote a random variable whose values indicate the presence ($A_{i,j,t} = 1$) or

991   absence ($A_{i,j,t} = 0$) of eDNA in the jth sample ($j = 1, \ldots, J_{i,t}$). We assume

$$A_{i,j,t}|Z_{i,t} \sim \text{Bernoulli}(\theta_{i,j,t}I(Z_{i,t} = 1)) \hspace{3cm} (8)$$

992   where the parameter $\theta_{i,j,t}$ is the conditional probability that eDNA is present in the jth sample of

993   location i during period t given that this location is occupied (more correctly, given that eDNA of

994   the target species is present) during period t. Note that $A_{i,j,t}$ equals zero with probability one if the

995   ith survey location is unoccupied during period t.

996   Suppose $K_{i,j,t}$ independent replicates are extracted from the jth sample of survey location i

997   during period t and are amplified using PCR chemistry. Let $Y_{i,j,t}$ denote a random variable for the

998   number of these replicates in which eDNA is detected. We assume

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

999   where the parameter $p_{i,j,t}$ is the conditional probability that eDNA is detected in a single replicate

1000   of the jth sample given that eDNA is present in this sample. Eq. 9 implies that $Y_{i,j,t}$ equals zero

1001   with probability one if eDNA is absent from the sample.

1002   As we have done previously, the parameters in Eqs. 8 and 9 are specified as functions of

1003   covariates that are thought to be informative of the occurrence or detection of eDNA in samples.

1004   For example, adopting the notation used by Dorazio and Erickson (2018), we assume

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

1005   for sample occurrence probability and

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

1006   for the probability of detecting eDNA in a PCR replicate. Covariates of $\theta_{i,j,t}$ and $p_{i,j,t}$ are codified

1007   in the vectors of regressors ($w_{i,j,t}$ and $v_{i,j,t}$, respectively) and offer ample opportunity to identify

1008   factors that are thought to influence eDNA occurrence and detection at the sample level.

1009   **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 $\rho_{i,j,t}$

1010
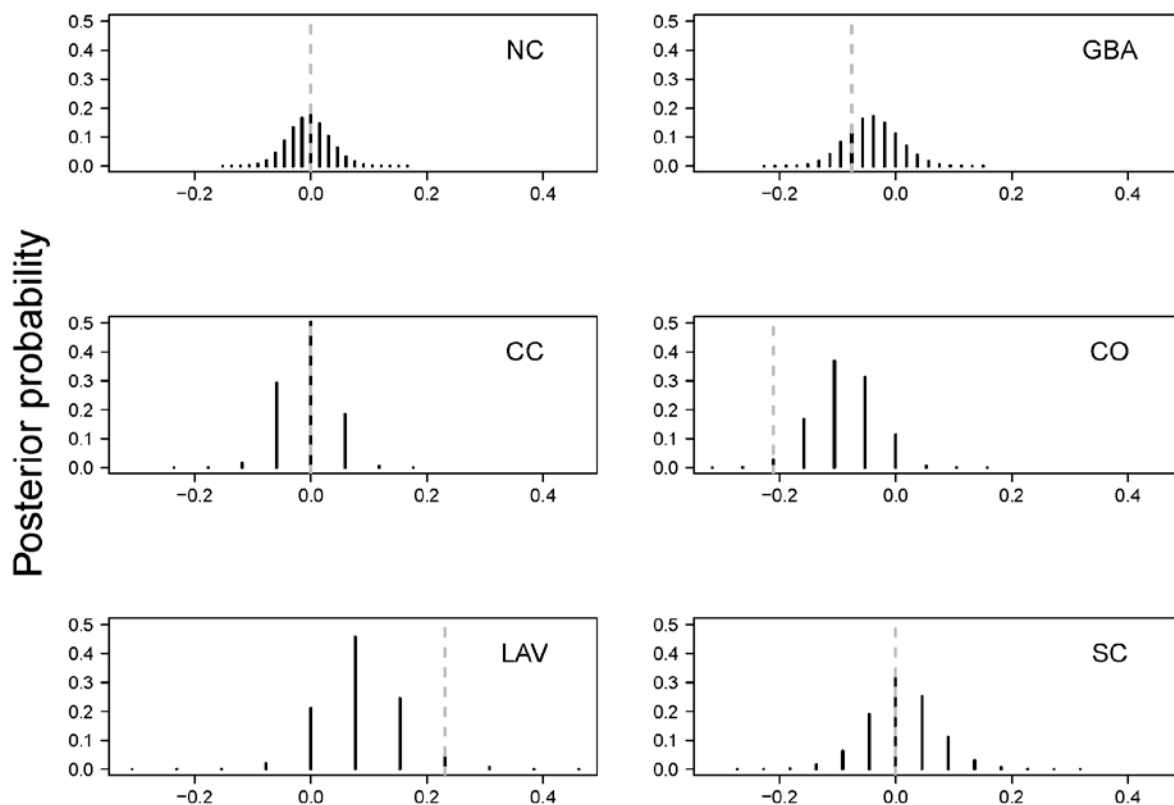
Change in proportion of survey locations containing eDNA

1013   **Figure 12 Posterior probability of change in proportion of sites containing eDNA from**
1014   **2016 to 2017 within the Recovery Units (NC = North Coast, GBA = Greater Bay Area, CC**
1015   **= Central Coast, CO = Conception, LAV = Los Angeles/Ventura, SC = South Coast)**
1016   **delineated by the Tidewater goby Recovery Plan (USFWS 2005) along the California coast.**
1017   **A shift in the positive direction on the x-axis represents in increase in the number of**
1018   **occupied sites. Bar density is representative of the site density in the recovery units. Naïve**
1019   **rates of change in proportion of sites containing eDNA are represented by the dashed grey**
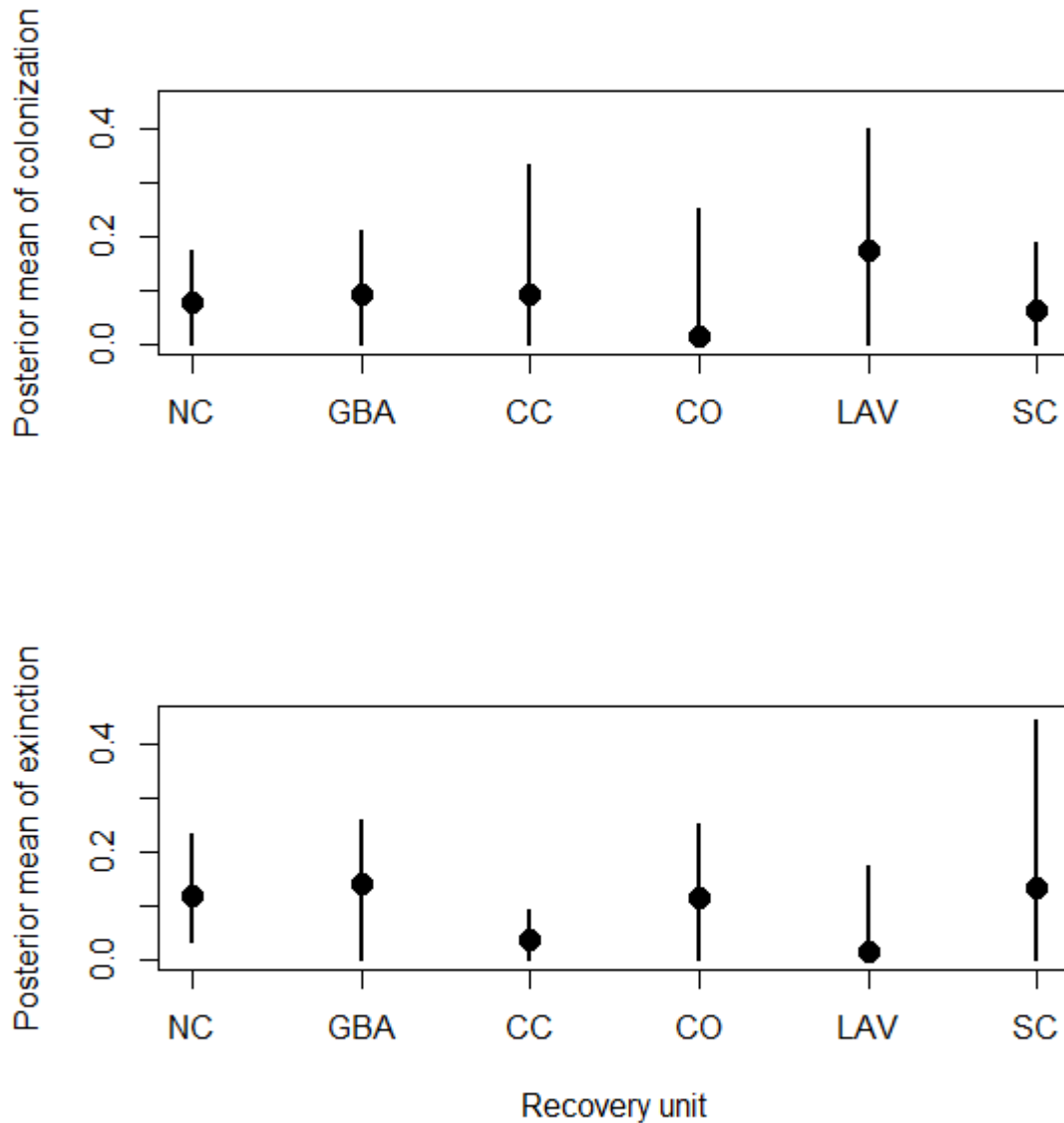1020   **line. Data were analyzed using a dynamic multiscale occupancy model.**
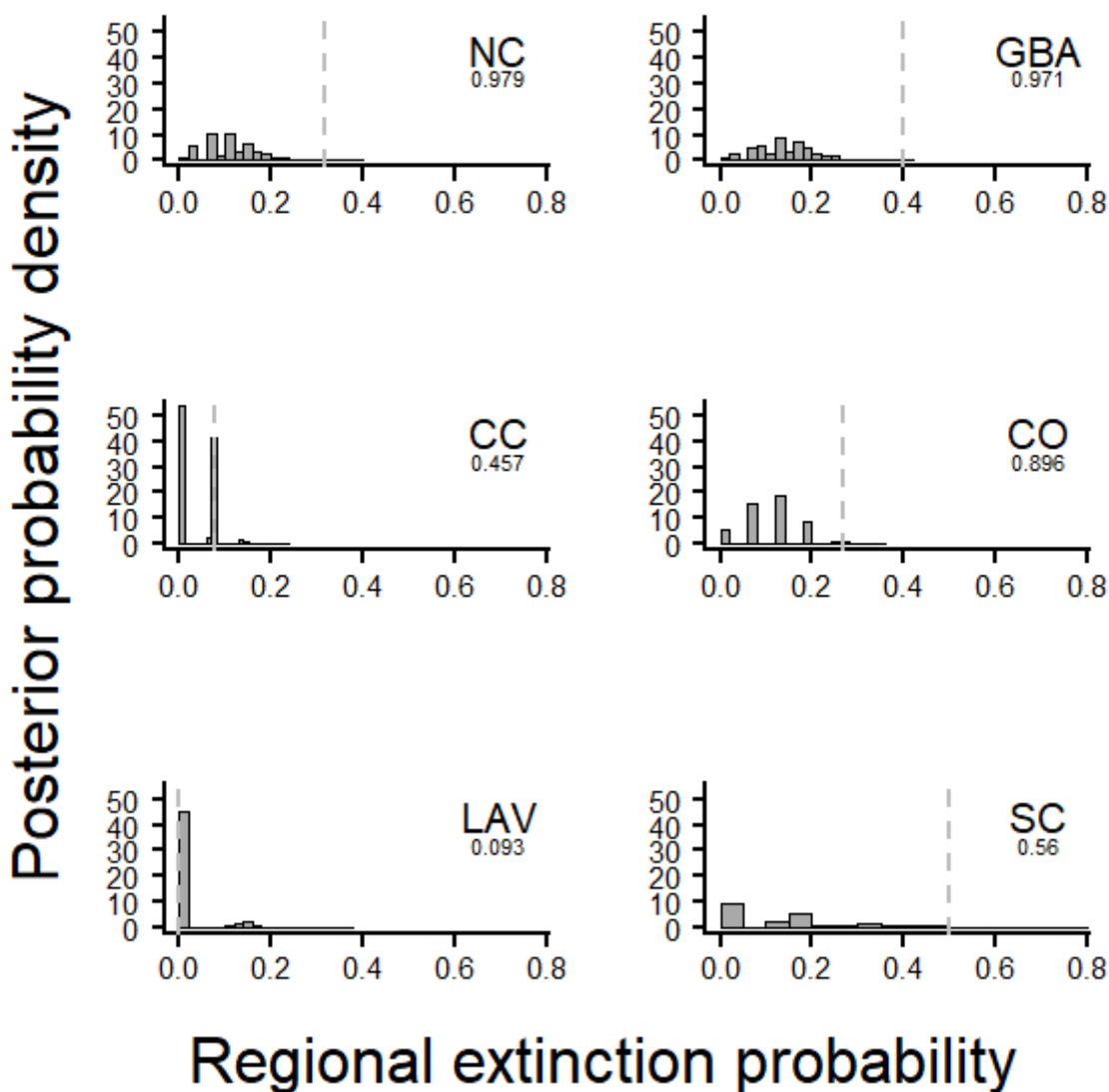
Figure 13 Estimates of probability of extinction within Recovery Units (NC = North Coast, GBA = Greater Bay Area, CC = Central Coast, CO = Conception, LAV = Los Angeles/Ventura, SC = South Coast) delineated by the Tidewater Goby Recovery Plan from a rangewide eDNA survey conducted in 2016 and 2017. Conditional 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.
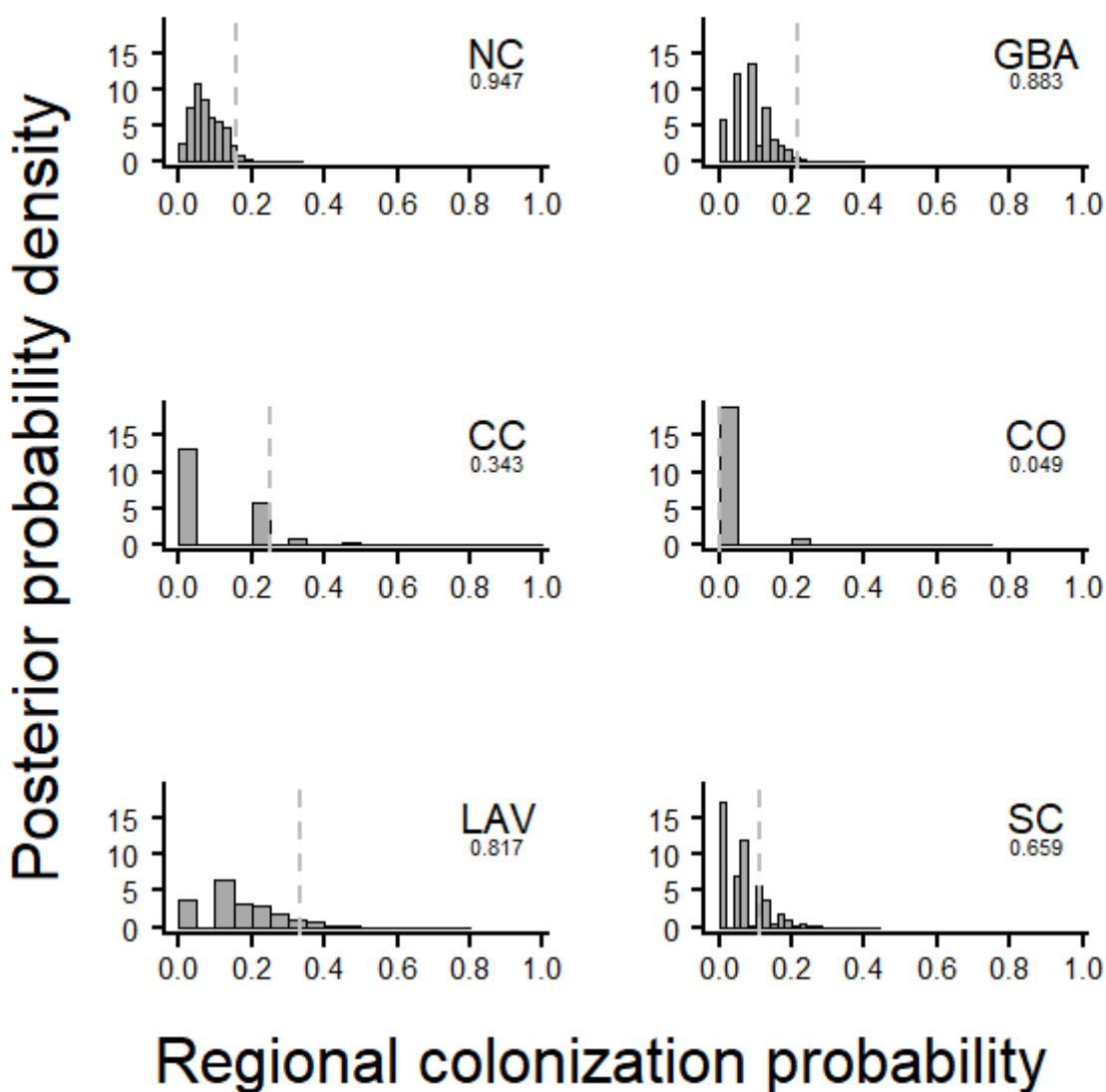
**Figure 14 Estimates of probability of colonization within Recovery Units (NC = north coast, GBA = greater bay area, CC = central coast, CO = Conception, LAV = Los Angeles/Ventura, SC = south coast) delineated by the Tidewater goby Recovery Plan from a rangewide eDNA survey conducted in 2016 and 2017. 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.**

1039

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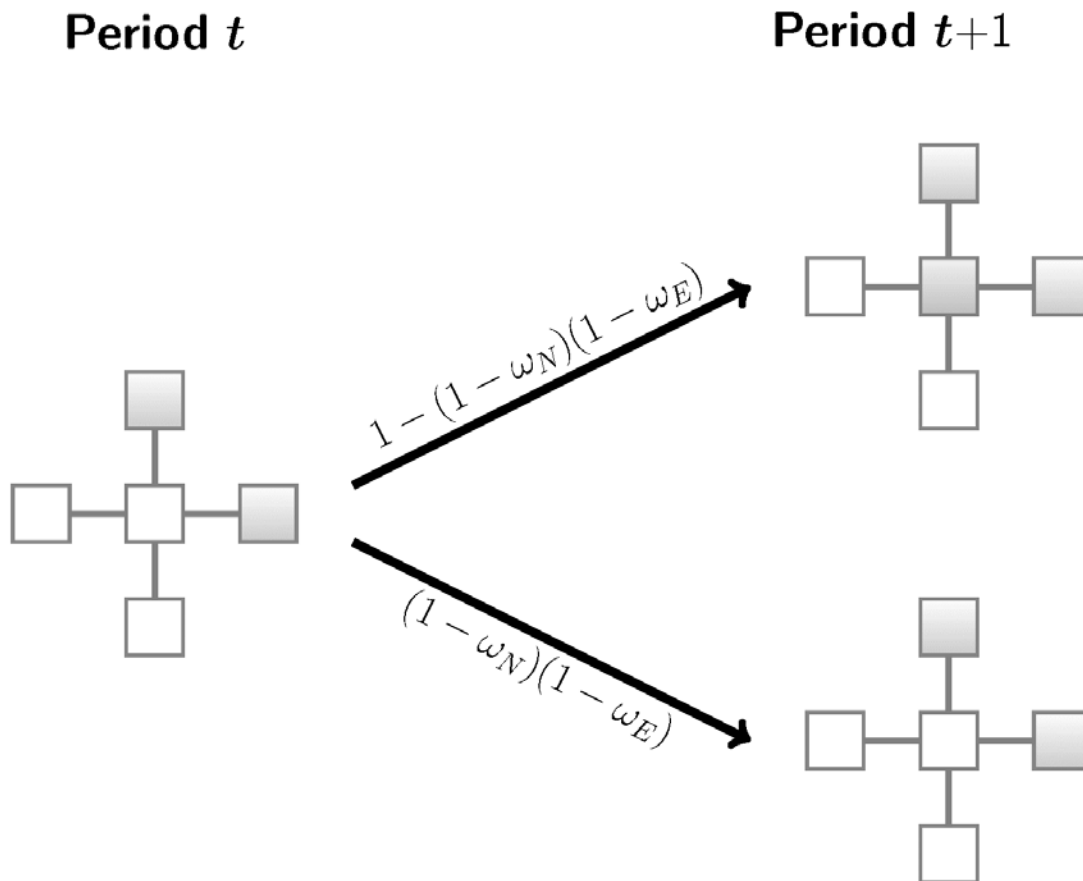
Figure 16 Diagrammatic illustration of our model's specification of the effects of dispersal from occupied neighboring locations to an unoccupied survey location. In this figure only two of four neighboring locations are occupied, so the probability of being colonized by at least one of the four neighbors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