

# Persistence of a reef fish metapopulation via network connectivity: theory and data

Allison G. Dedrick<sup>a,\*</sup>

Katrina A. Catalano<sup>a</sup>

Michelle R. Stuart<sup>a</sup>

J. Wilson White<sup>b</sup>

Humberto Montes, Jr.<sup>c</sup>

Malin Pinsky<sup>a</sup>

a. Department of Ecology Evolution and Natural Resources, Rutgers University, 14  
College Farm Road, New Brunswick, NJ 08901 USA;

b. Department of Fisheries and Wildlife, Coastal Oregon Marine Experiment Sta-  
tion, Oregon State University, Newport, OR 97365 USA;

c. Visayas State University

\* Corresponding author; e-mail: agdedrick@gmail.com

*(Author order not yet determined)*

# Introduction

Metapopulation dynamics and persistence depend on connectivity among patches  
3 and the demographic rates at each patch (e.g. Hastings and Botsford, 2006; Hanski,  
1998). Assessing levels of connectivity and demographic parameters has been par-  
ticularly challenging for marine species, where much of the mortality and movement  
6 happens at larval and juvenile stages when individuals are hard to track and have  
the potential to travel long distances with ocean currents (reviewed in White et al.,  
2019). A need to understand metapopulations for conservation and management,  
9 such as siting marine protected areas (e.g. Botsford et al., 2001; White et al., 2010),  
however, has led to a large body of theory describing how marine metapopulations  
might persist.

12 For any population to persist, individuals must on average replace themselves  
during their lifetime. Assessing replacement must take into account the demographic  
processes across the whole life cycle, including how likely individuals are to survive to  
15 the next age or stage, their expected fecundity at each stage, and the survival of any  
offspring produced to recruitment. In a spatially structured population, as many  
marine populations are, in addition to assessing whether the reproductive output  
18 and survival of a population is sufficient, we must also consider how the offspring are  
distributed across space.

Considering both the demographic processes within patches and the connectivity  
21 among them, a metapopulation can persist in two ways: 1) at least one patch can  
achieve replacement in isolation, or 2) patches receive enough recruitment to achieve

replacement through multi-generational loops of connectivity with other patches in  
24 the metapopulation (Hastings and Botsford, 2006; Burgess et al., 2014). In the first  
case (termed self-persistence), enough of the reproductive output produced at one  
patch is retained at the patch for it to persist. In the second (network persistence),  
27 closed loops of connectivity among at least some of the patches - where individuals  
from one patch settle at another and eventually send offspring back to the first in a  
future generation - provide the patch with enough recruitment to persist within the  
30 network. Though it has been challenging to estimate the parameters necessary to  
understand how actual metapopulations persist, a large work of theory developed in  
part to guide marine protected area design helps predict when each type of persistence  
33 is likely to occur (i.e., habitat patches or protected areas that are large relative to  
the mean dispersal distance are likely to be self-persistent, White et al., 2010).

Again, Hameed did not use those new tools (I dont recall off the top of my head  
36 what Carson et al. did in that paper). Perhaps a better way to express this point is to  
say that new tools allow us to measure the dispersal (Almany, D'Aloia) and a better  
appreciation of the relevant population dynamic theory have led to measurement of  
39 the appropriate demographic factors (Carson, Hameed).

New ways of identifying individuals and determining their origins, such as otolith  
and shell microchemistry and genetic parentage analysis (e.g. Wang, 2004, 2014), now  
42 allow us to better measure dispersal (e.g. Almany et al., 2017; D'Aloia et al., 2013)  
and a better appreciation of the relevant population dynamic theory has led to mea-  
surement of the appropriate demographic factors (e.g. Carson et al., 2011; Hameed  
45 et al., 2016) necessary to assess persistence in real metapopulations. We might ex-

pect that populations on isolated islands are the most likely to be self-persistent, as they lack nearby populations with which to exchange larvae and would go locally  
48 extinct if they did not achieve replacement. At isolated Kimbe Island in Papua New Guinea, Salles et al. (2015) find that the population of orange clownfish (*Amphiprion percula*) can likely persist without outside immigration. In contrast, populations of  
51 bicolor damselfish (*Stegastes partitus*) at four study sites nested within a larger reef metapopulation in the Bahamas do not appear able to persist without outside input (Johnson et al., 2018). Persistence has yet to be assessed in the field for an entire  
54 marine metapopulation, such as all of the patches in a coastal metapopulation.

The number of studies estimating demographic rates and connectivity in marine metapopulations is growing (e.g. Carson et al., 2011; Salles et al., 2015; Johnson  
57 et al., 2018; Garavelli et al., 2018), but most use data from one or a few years. Longer data sets enable better estimates of long-term average rates, rather than assuming the demographic and dispersal rates from a particular year or two are  
60 representative. Long data set are also useful for explicitly considering uncertainty, both to assess how well we understand persistence for a population and to assess which parameters contribute most to our uncertainty. Finally, sampling over many  
63 years provides abundance trends to compare with persistence metrics.

Here, we further our understanding of metapopulation dynamics in a network of patches along a coastline through a study of yellowtail clownfish (*Amphiprion clarkii*) in the Philippines. We assess persistence for all patches of habitat within a 30 km stretch of coastline, exceeding estimates of the dispersal spread for this species  
66 (Pinsky et al., 2010), suggesting the network is likely to operate as a contained

69 metapopulation. With seven years of annual sampling data, we are able to estimate  
persistence metrics and replacement over the longer term and investigate abundance  
through time to compare with the replacement-based persistence metrics. We use  
72 our long-term data set from habitat patches on a continuous section of coastline to  
understand persistence within a local network. We find that our sites have stable  
abundances through time but are unlikely to persist as an isolated metapopulation  
75 and require immigration from outside patches to persist.

## Methods

### Persistence theory and metrics

78 For a population to persist, each individual must on average replace itself (e.g. Hastings and Botsford, 2006; Botsford et al., 2019). In non-spatially structured populations, we use criteria such as the average number of recruiting offspring each  
81 individual produces during its life (called  $R_0$  when the population is age-structured and density-independent) or the growth rate of the population (such as the dominant eigenvalue  $\lambda$  of an age-structured Leslie matrix) (Caswell, 2001; Burgess et al., 2014).  
84 For spatially-structured populations, we must also consider the spatial spread of offspring, often represented through a dispersal kernel or connectivity matrix (Burgess et al., 2014).

87 We consider four primary metrics to assess whether and how the population is persistent: 1) lifetime recruit production (LRP), to assess whether the population has enough surviving offspring to achieve replacement, 2) self-persistence (SP), to

90 assess whether any individual patch can persist in isolation without input from other  
patches, 3) network persistence (NP), to assess whether the metapopulation is per-  
91 sistent as a connected unit, and 4) local replacement (LR), as second assessment of  
92 whether individuals replace themselves with recruits retained within population. We  
explain each metric below in detail. To represent the uncertainty in our estimates, we  
calculate each metric 1000 times, sampling each input parameter from a distribution  
96 that represents the uncertainty in each empirical estimate of demographic rates or  
connectivity. In our results, we show our best estimate of each persistence metric  
along with the range of uncertainty values.

99 **Lifetime recruit production**

We find the estimated number of recruits an individual recruit will produce (lifetime  
recruit production: LRP) by multiplying the total number of eggs a recruit-sized  
102 individual will produce in its lifetime (lifetime egg production: LEP) by the fraction  
of those eggs that will survive to become recruits (egg-recruit survival:  $S_e$ ) (Fig. 1):

$$\text{LRP} = \text{LEP} * S_e. \quad (1)$$

If  $\text{LRP} \geq 1$ , the population has the potential for replacement; individuals produce  
105 enough surviving offspring, before considering dispersal. If  $\text{LRP} < 1$ , the individuals  
are not replacing themselves and the population cannot persist without input from  
outside patches and is a sink habitat within a larger metapopulation (Pulliam, 1988).  
108 We use all recruits produced by adults in our population to estimate  $\text{LRP}$ , regardless

of where they settle.

### Self-persistence

- <sup>111</sup> A patch is able to persist in isolation (self-persistent) if individuals produce enough offspring that survive to recruitment (LRP) and settle in the natal patch ( $i$ , with probability of dispersal  $p_{i,i}$ ) to replace themselves:

$$SP_i = \text{LRP}_i \times p_{i,i}. \quad (2)$$

- <sup>114</sup> A patch  $i$  is self-persistent if  $SP_i \geq 1$ . If at least one patch is self-persistent, the metapopulation as a whole is persistent as well (Hastings and Botsford, 2006; Burgess et al., 2014). Our equation for SP is a modification of that used in Burgess et al. (2014), which uses LEP to represent offspring produced and uses local retention (the number of surviving recruits that disperse back to the natal patch divided by the number of eggs produced by the natal patch) to capture egg-recruit survival and dispersal combined:  $\text{LEP} \times \text{local retention} \geq 1$ . We modify this to include egg-recruit survival in the offspring term instead, using LRP in place of LEP.  
<sup>117</sup>  
<sup>120</sup>

### Network persistence

- <sup>123</sup> Network persistence explicitly considers dispersal of individuals among sites, a critical part of the larval life stage, in addition to the reproduction and survival at each site. We represent dispersal with a dispersal kernel, which relates the likelihood of

126 an individual dispersing to the distanced traveled. We find the probabilities of a  
127 recruit dispersing between each set of sites ( $p_{i,j}$ ) by integrating the dispersal kernel  
128 over the distances between sites. We then create a realized connectivity matrix  $C$   
129 by multiplying the dispersal probabilities by the expected number of recruits an  
130 individual produces:  $C_{i,j} = \text{LRP} \times p_{i,j}$  (Burgess et al., 2014, though we include egg-  
131 recruit survival in LRP, rather than in  $p_{i,j}$  as they do). The diagonal entries of  $C$ ,  
132 where the origin and destination are the same site, are the values of self-persistence  
133 for each individual site.

134 Network persistence evaluates the largest real eigenvalue of the realized connec-  
135 tivity matrix  $\lambda_C$ , which must be greater than 1 for the network to persist without  
136 outside input:  $\text{NP} = \lambda_C \geq 1$  (e.g. Hastings and Botsford, 2006; White et al., 2010;  
137 Burgess et al., 2014).

138 **Local replacement**

139 Like network persistence, local replacement (LR) asseses whether the population  
140 is locally self-sustaining. Rather than considering dispersal explicitly as network  
141 persistence does, local replacement modifies LRP to estimate the average number of  
142 recruits produced per individual that return to settle within our sites. We estimate  
143 LR by multiplying LEP by the proportion of eggs produced that survive and return  
144 to recruit at our sites ( $R_e$ ), a modification of egg-recruit survival that implicitly  
145 includes dispersal. If  $LR \geq 1$ , individuals produce enough locally-retained offspring

to replace themselves and the population can persist in isolation.

$$\text{LR} = \text{LEP} \times R_e. \quad (3)$$

## <sup>147</sup> Study system

We focus on a tropical metapopulation of yellowtail clownfish (*Ampiprion clarkii*, Fig. 2c) on the west coast of Leyte island facing the Camotes Sea in the Philippines (Fig. 150 2a). Like many clownfish species, yellowtail clownfish have a mutualistic relationship with anemones that house small colonies of fish (Buston, 2003b; Fautin et al., 1992). Yellowtail clownfish are protandrous hermaphrodites and maintain a size-structured 153 hierarchy; within an anemone, the largest fish is the breeding female, the next largest is the breeding male, and any smaller fish are non-breeding juveniles. The fish move up in rank to become breeders only after the larger fish have died. In the 156 tropical patch reef habitat of the Philippines, yellowtail clownfish primarily spawn from November to May, laying clutches of benthic eggs that the parents protect and tend (Ochi, 1989; Holtswarth et al., 2017). Larvae hatch after about six days and 159 spend 7-10 days in the water column before returning to reef habitat to settle in an anemone (Fautin et al., 1992).

Clownfish are particularly well-suited to metapopulation studies due to their limited movement as adults and well-defined patchy habitat. Once fish have settled, 162 they tend to stay within close proximity of their anemones, which are found on reef patches. This makes fish easier to relocate for mark-recapture studies and simiplifies

the exchange between patches to only dispersal during the larval phase. Patches, whether considered to be the reef patch or the anemone territory of the fish, are discrete and easily delineated (Fig. 2a, b), which makes determining the spatial structure of the metapopulation clear. Additionally, clear patches make it easier to assess how much of the site has been surveyed. These simplifying characteristics in habitat and fish behavior make clownfish and other similarly territorial reef fish useful model systems for studies of metapopulation dynamics and persistence (e.g. Buston and DAloia, 2013; Salles et al., 2015; Johnson et al., 2018). Our focal species of yellowtail clownfish tends to behave more like larger reef fishes, with wider-ranging territories and stronger swimming abilities (Hattori and Yanagisawa, 1991; Ochi, 1989), than the smaller clownfish *A. percula* commonly used in previous metapopulation studies (e.g. Buston et al., 2011; Salles et al., 2015). As we show later, survival in yellowtail clownfish is also lower than *A. percula* and more similar to other damselfishes. MAKE SURE THIS SENTENCE IS TRUE!

## Field data collection

We focus on a set of nineteen patch reef sites spanning 30 km along the western coast of Leyte island (Fig. 2a). The sites consist of rocky patches of coral reef separated by sand flats, with reef patches covering approximately 20% of the sampling region (Fig. 2b). On the north edge, the sites are isolated from nearby habitat with no substantial reef habitat for at least 20 km.

Since 2012, we have sampled fish and habitat at most of the sites annually (Table A2). During sampling, divers using SCUBA and tethered to GPS readers swam the

extent of each site. Divers visited each anemone inhabited by yellowtail clownfish and tagged anemones. At each anemone, the divers attempted to catch all of the 189 yellowtail clownfish 3.5 cm and larger, took a small tail fin-clip from each for use in genetic analysis, measured the fork length, and noted the tail color (as an indicator of life stage). Starting in the 2015 field season, fish 6.0 cm and larger were also tagged 192 with a passive integrated transponder (PIT) tag, unless already tagged. Divers also looked for eggs around each anemone and measured and photographed any clutches found. In total, we took fin clips from and genotyped 2772 fish and PIT-tagged 1929 195 fish across all years and sites combined, marking 3413 individual fish.

## Estimating demographic and dispersal parameters from empirical data

### 198 Parentage analysis and dispersal kernel

Over seven years of sampling, we genotyped 1719 potential parents and 785 juveniles and found 62 parent-offspring matches (details in Catalano et al., in prep). We use a 201 distance-based dispersal kernel fit from the parent-offspring matches (Catalano et al., in prep), where the relative dispersal  $p(d)$  is a function of distance  $d$  in kilometers and parameters  $\theta = 1.19$  and  $z = e^{k_d=-2.33}$  that control the shape and scale of the kernel: 204  $p(d) = ze^{-(zd)^\theta}$ . The dispersal kernel estimates the relative dispersal by distance for fish that have survived and recruited so does not estimate pre-settlement mortality. To find the probability of fish dispersing among our sites, we numerically integrate 207 the dispersal kernel using the distance from the middle of the origin site ( $i$ ) to the

closest ( $d_1$ ) and farthest ( $d_2$ ) bounds of the destination site ( $j$ ):

$$p_{i,j}(d) = \int_{d_1}^{d_2} ze^{-(zd)^\theta} dd. \quad (4)$$

To account for uncertainty in the dispersal kernel, we use sets of the shape parameter  $\theta$  and the scale parameter  $k_d$  that represent the span of the 95% confidence interval when  $k_d$  and  $\theta$  are estimated jointly (Catalano et al., in prep).

### Growth and survival: mark-recapture analyses

We marked fish with both genetic samples and PIT tags, allowing us to estimate growth and survival through mark-recapture. In total, we have 3413 marked fish with size and stage data for each capture time.

For growth, we estimated the parameters of a von Bertalanffy growth curve (Fabens, 1965) in the growth increment form relating the length at first capture  $L_t$  to the length at a later capture  $L_{t+1}$  (Hart and Chute, 2009), where  $L_\infty$  is the average asymptotic size across the population and  $K$  controls the rate of growth:

$$\begin{aligned} L_{t+1} &= L_t + (L_\infty - L_t)[1 - e^{(-K)}] \\ &= e^{(-K)}L_t + L_\infty[1 - e^{(-K)}]. \end{aligned} \quad (5)$$

We see from eqn. 5 that we would expect the first length  $L_t$  and the second length  $L_{t+1}$  to be related linearly (Hart and Chute, 2009). From the slope  $m = e^{(-K)}$  and y-intercept  $b = L_\infty[1 - e^{(-K)}]$ , we calculated the von Bertalanffy parameters, such

that  $K = -\ln m$  and  $L_\infty = \frac{b}{(1-m)}$ . We used the first and second capture lengths for fish that were recaptured after a year (within 345 to 385 days) to estimate  $L_\infty$  and  $K$ . We have some fish that were recaptured more than two times so we randomly selected only one pair of recaptures from each to use to estimate the parameters 225 1000 times. We found the mean estimates and mean standard error of those fits, 228 then sampled from within than range to generate a set of von Bertalanffy growth curves to use in our metric uncertainty calculations.

We used the full set of marked fish to estimate annual survival  $\phi$  and probability of 231 recapture  $p_r$  using the mark-recapture program MARK implemented in R through the package **RMark** (Laake, 2013). We fit several models with year, size, and site effects on the probability of survival on a log-odds scale (see full list in Table A4). 234 For fish that were not recaptured in a particular year, we estimated their size using our growth model (eqn. 5) and the size recorded or estimated in the previous year. Fish are not well-mixed at our sites and divers needed to swim near an anemone to 237 have a reasonable chance of capturing the fish on it so we also considered a distance effect on recapture probability. Using diver GPS tracks, we estimated the minimum distance between a diver and the anemone for each tagged fish in each sample year 240 to include as a factor. We compared the fit of the models using a modified version of the Akaike information criterion that reduces the potential for overfitting with small sample sizes (AICc) and selected the model with the lowest AICc value. (Table A4).

243 **Fecundity**

We used a size-dependent fecundity relationship determined using photos of egg clutches and females (Yawdoszyn, in prep), where the number of eggs per clutch 246 ( $E_c$ ) is exponentially related to the length in cm of the female ( $L$ ) with size effect  $\beta_l = 2.388$ , intercept  $b = 1.174$ , and egg age effect  $\beta_e = -0.6083$  dependent on if the eggs are old enough to have visible eyes. We multiplied the number of eyed eggs 249 per clutch by the number of clutches per year  $c_e = 11.9$  (estimate from Holtswarth et al., 2017) to get total annual fecundity  $f$  of a female of length  $L$ :

$$f = c_e * e^{\beta_l \ln(L) + \beta_e [\text{eyed}] + b}. \quad (6)$$

### Lifetime egg production

252 We used an integral projection model (IPM) (Ellner et al., 2016) with size as the continuous structuring trait  $z$  to estimate lifetime egg production (LEP), the expected total number of eggs produced by one recruit. We initialized the IPM with 255 one recruit-sized individual at the initial annual time step ( $t = 0$ ), then projected forward for 100 time steps. We used the size-dependent survival (eqn. B.3) and growth (eqn. 5) functions as the probability density functions in the kernel to describe the survival and growth of the individual into the next time step. We get the size distribution ( $v_z$ ) at each time step, which represents the probability that the individual has survived and grown into each of the possible size categories, ranging 258 from a minimum of  $L_s = 0$  cm to a maximum of  $U_s = 15$  cm divided into 100 equal 261 from a minimum of  $L_s = 0$  cm to a maximum of  $U_s = 15$  cm divided into 100 equal

size bins. The probability that the individual is still alive and of any size decreases as the time steps progress; by using a large number of steps, we are able to avoid arbitrarily setting a maximum age and instead let the probabilities become essentially zero.

We then multiplied the size-distribution  $v_z$  at each time by the size-dependent fecundity  $f_z$  described above (eqn. 6) to get the total number of eggs produced at each time step. Integrating across time and size gives the total number of eggs one individual is likely to produce in its lifetime:

$$\text{LEP} = \int_{t=0}^{100} \int_{z=L_s}^{z=U_s} v_{z,t} f_z dz dt. \quad (7)$$

When entering the starting individual into the matrix, we use 0.1 as the standard deviation of size to spread the starting individual across size bins. When projecting the distribution of the size of the fish in the next year, we used the size determined by the growth curve (eqn. 5) as the mean along with an estimate of spread to account for differences in fish growth rates. We used our recapture data to estimate the standard deviation ( $\text{size}_{sd}$ ) of the distribution of sizes in the next year of fish starting from one size (A1).

We only considered reproductive effort once the fish has reached the female stage and use the average size of first observation as female for recaptured fish as the transition size  $L_f = 9.32\text{cm}$ . To incorporate uncertainty, we sampled directly from the sizes at which recaptured fish were first captured as female (5.2 cm - 12.7 cm) (Fig. 3d).

282 **Survival from egg to recruit**

We estimate survival from egg to recruit ( $S_e$ ) using parentage matches to find the number of surviving recruits produced by genotyped parents (similar to the method in Johnson et al., 2018). We scale the number offspring we assign back to parents 285 ( $R_m = 62$ ) by various ways we could have missed offspring in our sampling ( $P_h$ ,  $P_c$ ,  $P_d$ , and  $P_s$ , described below and in Fig. B.1), then divide by the estimated number of eggs produced by genotyped parents (the number of genotyped parents  $N_g = 1719$  288 multiplied by the expected LEP for a fish of parent size  $\text{LEP}_p$ :

$$S_e = \frac{\frac{R_m}{P_h P_c P_d P_s}}{N_g \text{LEP}_p}. \quad (8)$$

We scale the number of matched recruits we find by the cumulative proportion of 291 habitat in our sites we sampled over time ( $P_h = 0.41$ , details in B.1), the probability of capturing a fish if we sampled its anemone ( $P_c = 0.56$ , see B.2 for details), and the proportion of the total dispersal kernel area from each of our sites covered within 294 our sampling region ( $P_d = 0.28$ , calculation in B.3). Finally, because our dispersal kernel gives the probability of dispersal given that a recruit settled somewhere but our sampling region is not all habitat, we scale by the proportion habitat in our sampling 297 region ( $P_s = 0.20$ , details in B.3.0.1) to avoid counting this mortality twice.

To estimate local replacement, we scale only by the proportion of habitat we cumulatively sample in our sites and the probability of capturing a fish to estimate 300 the survival and retention of recruits back to our sites:  $R_e = \frac{\frac{R_m}{P_h P_c}}{N_g \text{LEP}_p}$ .

To incorporate uncertainty in our estimate of egg-recruit survival, we consider

uncertainty in the number of offspring assigned to parents during the parentage  
303 analysis ( $R_m$ ) and in the probability of capturing a fish ( $P_c$ ). We generate a set  
of values for the number of assigned offspring using a random binomial, where the  
number of trials is the number of genotyped offspring (745) and the probability of  
306 success on each trial is the assignment rate of offspring from the parentage analysis  
(0.079) (Catalano et al., in prep). For the probability of capturing a fish, we sample  
values from a beta distribution that captures the mean and variance of capture  
309 probabilities across recapture dives (details in B.2).

### Defining recruit and census stage

When assessing persistence, it is important to consider mortality and reproduction  
312 that occurs across the entire life cycle to determine whether an individual is replacing  
itself with an individual that reaches the same life stage (Burgess et al., 2014). We  
define a recruit to be a juvenile individual that has settled on the reef within the  
315 previous year, which also encompasses the size we are first able to sample (3.5-6.0 cm  
for parentage studies) (Fig. 1). In theory, it does not matter how we define a recruit  
as long as we use that definition in our calculations of both egg-recruit survival and  
318 LEP. In our system, however, while it is straightforward to calculate LEP from any  
size, we did not have enough tagged recruits to reliably estimate survival to different  
recruit sizes. Instead, we choose the mean size of offspring matched in the parentage  
321 study as our best estimate of the size of a recruit ( $\text{size}_{\text{recruit}}$ ) and test sensitivity  
to different recruit sizes by sampling from a uniform distribution over the sizes the  
recruit stage covers (3.5-6 cm, Table A1).

<sup>324</sup> **Accounting for density-dependence**

Ideally we would assess persistence metrics when the population is at low abundance and not limited by density-dependence. Clownfish have strong social hierarchies and juveniles on an anemone will prevent others from settling there as well (seen in <sup>327</sup> *A. percula*, Buston, 2003a). Each anenome, therefore, can house only one settling clownfish, with anemones already occupied by *A. clarkii* settlers essentially unavailable as habitat. This density-dependent mortality will artificially reduce the apparent survival of new recruits, biasing persistence metrics. We account for this density-dependent mortality by multiplying our estimate of settling recruits (the numerator <sup>330</sup> of eqn. 8) by the proportional increase (DD) in unoccupied anemones at our sites if all of the *A. clarkii* anemones were unoccupied, where  $p_A$  is the proportion of anemones occupied by *A. clarkii* and  $p_U$  is the proportion of unoccupied anemones: <sup>333</sup>  $DD = \frac{(p_U + p_A)}{p_U}$ . We present results with this density-dependence modification (with subscript DD: LRP<sub>DD</sub> in the main text and without in the appendix (Figs. C.2, C.3)).

**Estimated abundance over time**

<sup>339</sup> We also consider trends in abundance of breeding females at each site over time ( $F_{i,t}$ ) to compare to our replacement-based estimates of persistence. Similarly to as we do for offspring, we scale up the number of females caught ( $F_c$ ) at each site  $i$  in each <sup>342</sup> sampling year  $t$  by the proportion of habitat sampled in that site and year  $P_{h_{i,t}}$  and

by the probability of capturing a fish  $P_c$ :

$$F_{i,t} = \frac{F_{c_{i,t}}}{P_{h_{i,t}} P_c} \quad (9)$$

We fit a mixed effects model to estimate the number of fish in each year as a  
 345 Poisson variable  $\lambda_a$  with site as a random effect  $m_i$  using the package `lme4` in R  
 (Bates et al., 2015):

$$\begin{aligned} F_{i,0} &\sim Poisson(\lambda_a) \\ F_{i,t} &= (\lambda_a + m_i)^t. \end{aligned} \quad (10)$$

We estimate  $\lambda_a$  for an average site as well as the individual sites. The population is  
 348 increasing over time if  $\lambda_a > 1$  and decreasing if  $\lambda_a < 1$ .

(11)

## Results

From the mark-recapture analysis of tagged and genotyped fish, we estimated mean  
 351 values of  $L_\infty = 10.71\text{cm}$  (range of estimates 10.50 - 10.90 cm) and  $K = 0.864$  (range  
 of estimates 0.785 - 0.944) for the von Bertalanffy growth curve parameters (eqn. 5,  
 Fig. 3b, Table A1). For juvenile and adult (post-recruitment) survival on a log-odds  
 354 scale, the best-fit model has an effect of site and a positive effect of size (eqn. B.3,  
 Table A1, Figs. 3c, B.3). The accompanying best-fit model for recapture probability

has negative effects of size and diver distance from anemone (eqn. B.4, Table A1,  
357 Fig. B.4).

For our persistence metrics, we present the estimate using the input values without  
357 considering uncertainty (the mean of the distributions, with the exception of the  
size range for recruit), as the best estimate and the range with uncertainty shown in  
360 brackets. Using our best estimates for growth, survival, and fecundity, we calculated  
an average value of LEP across sites of XX eggs [XX, XX], Fig. 4b). Our best esti-  
363 mates of LEP at individual sites range from XX eggs to XX eggs (Fig. C.4). Adult  
survival has the most effect on the value of LEP (Fig. C.7), with higher values of  
LEP with higher annual survival of adults because adults survive longer have more  
366 opportunities to reproduce and produce more eggs per clutch when they are larger.

We estimated egg-recruit survival  $S_{e_{DD}}$  to be 0.0013 [2.1e-04, 0.057] when we  
correct for density-dependence in our data. Uncertainty in the size of transition to  
369 breeding female  $L_f$  has the largest effect on egg-recruit survival (Fig. C.9); the larger  
the transition size to female, the fewer tagged eggs we estimate were produced by our  
genotyped parents and the higher our estimate of egg-recruit survival. This differs  
372 from our finding above that adult survival has the largest effect on LEP because the  
starting size of the individual considered is lower when we estimate LEP for a recruit  
than LEP for a parent (6.0cm). Fish considered as parents in our parentage analysis  
375 have already survived one or more years since recruiting so the transition to breeding  
female plays a larger role in the number of eggs they are likely to produce than for  
fish who have just recruited.

378 We estimated average lifetime recruit production ( $LRP_{DD}$ ) across sites, the prod-

uct of  $LEP_i$  and  $S_{e_{DD}}$ , to be 1.42 [0.48, 6.66] (Fig. 4c). Our best estimates of  $LRP_{DD}$  at individual sites range from XX to XX (Fig. C.5). Both best estimates, XX% of the  
381 estimates of LRP with density-dependence compensation are one or larger, above the threshold necessary for replacement before considering dispersal. This means that individuals at our sites produce enough surviving offspring before considering dispersal to be able to replace themselves, but LRP does not tell us whether those offspring  
384 will settle within our sample sites and drive persistence.

Our estimate of local replacement ( $LR_{DD}$ ), which estimates replacement for recruits from our sites returning to our sites, implicitly including dispersal, is 0.16 [0.05, 0.76] when we correct for density-dependence in our data. With a value well below one, this suggests individuals at our sites do not replace themselves with recruits  
390 that settle in our sites, and that our sites do not persist as an independent network. When we calculated LR using all arriving recruits to our sites, however, rather than just those originating there, our best estimate was  $> 1$  (2.06 with XX%  $\geq 1$ ), suggesting  
393 that there is recruit-recruit replacement at our sites when we include immigrant recruits.

We did not find any sites with a best estimate of  $SP_{DD} \geq 1$  (Fig. 5a), though  
396 at the sites Haina ( $SP_{DD} = 0.13$  [XX, XX]) and Wangag ( $SP_{DD} = XX$  [XX, XX]), XX% and XX%, respectively, of our estimates are  $\geq 1$ . We also saw estimates  $> 1$  for Caridad Cemetery but those were due to the lack of recaptures at that site and the  
399 resulting high uncertainty in adult survival and are unlikely to indicate persistence. This suggests that Haina and Wangag could persist on their own at the upper limit of uncertainty in adult survival at the sites, but in general our results do not indicate

402 that any site could persist in isolation.

Most of the connectivity occurs among the sites in the northern part of our sample area, from Palanas to Caridad Cemetery, and at the southern part of our sample area from Tomakin Dako to Sitio Baybayon (Fig. 5b), where the sites tend to be the largest, have higher abundances, and have higher survivals (though not entirely). Sites at the edges of our sampling region seem to be the strongest sub-populations, which means many of the recruits they produce could be exported away from our sites rather than into our sampling region.  
405  
408

For network persistence, our best estimate of the dominant eigenvalue of the realized connectivity matrix  $\lambda_{c_{DD}}$  is 0.36 [0.12, 1.58],  $p(\lambda_{c_{DD}} \geq 1) = XX$  when we correct for density-dependence(Fig. 5c). Network persistence for our sites is unlikely but not impossible.  
411

414 Based on our estimates of  $LRP_{DD}$ ,  $LR_{DD}$ ,  $SP_{DD}$ , and  $NP_{DD}$ , it is possible but unlikely that our set of sets is able to persist in isolation as a closed system. With our existing site configuration and dispersal kernel estimate, we would need a value of LRP of 3.99 (an egg-recruit survival of 0.0038 with our estimated value of LEP or a value of LEP of 5095 with our estimated value of  $S_{e_{DD}}$ ) to have a best estimate of  $\lambda_{c_{DD}} = 1$  and network persistence. Alternately, increasing the amount of habitat in our region (currently about 20%) to 55%, with equally sized and spaced sites with adult survival of an average site, gives a best estimate of  $\lambda_{c_{DD}} = 1$  and suggests network persistence (Fig. 6a). Considering uncertainty, 95% of the estimates of  $\lambda_{c_{DD}}$  are  $> 1$  at 85% habitat (Fig. 6b).  
417  
420  
423

Our estimated abundance of females over time has a slight positive trend for the

average site ( $\lambda_a = XX$ , Fig. 4a), suggesting a slight increase in population size for  
426 the population overall through time. Most individual sites also show a slight positive trend in female abundance through time, though one large site shows declines (Fig. 4a, Fig. C.1s).

## 429 Discussion

In this first assessment of demographic persistence of a coastal marine metapopulation, we did not find strong evidence for either self-persistence of an individual patch  
432 or network persistence of the entire system as an isolated region. Self-persistence of one of the larger sites or network persistence is possible at the upper end of our estimates with uncertainty, but neither is suggested by our best estimates or XX%  
435 of the estimate range with uncertainty. This inability to persist as an isolated region does not necessarily mean that the populations at our sites are declining, however. Our assessments of population trends - both abundance over time and replacement  
438 of recruits when we include immigrants - find that the population levels at our sites are stable or increasing slightly. Taken together, our metrics suggest that the sites in our region have stable populations on average but require input of immigrants  
441 to persist. The portion of coastline we sampled is likely a sink portion of a larger metapopulation, given that there does not seem to be a long-term deline in the population.

444 For our sites to be able to persist as a network on their own, either the number of surviving recruits produced by an average recruit (LRP) would need to be higher or more recruits would need to be retained within the region. With our existing site

447 configuration and estimated connectivity, LRP would need to be at least 3.99 to see  
a best estimate of network persistence among our sites, which is within our range of  
uncertainty (top XX %) but about 3-5 times higher than our best estimate. RELATE  
450 TO LITERATURE! HOW REASONABLE ARE THESE VALUES FOR A REEF  
FISH? Reproduction and survival at our sites might be lower than they could be  
due to generally low reef health and habitat quality in our sites from anthropogenic  
453 effects such as pollution and silt from a nearby gravel mine and habitat disturbance  
from storms. Our sites are in an area that was hit in 2013 by Typhoon Haiyan, one  
of the strongest typhoons ever to make landfall, which destroyed much of the reef  
456 habitat in one of our northern sampling sites. This recent disturbance and generally  
low habitat quality could contribute to low production of surviving recruits in our  
sites (seen in other populations with low habitat quality, e.g. Hayashi et al., 2019)  
459 necessitating subsidization by outside populations. Alternately, higher connectivity  
and retention of offspring among our sites could lead to network persistence. Lack  
of sufficient connectivity and retention was thought to inhibit network persistence  
462 for a collection of reserves for eastern oysters (*Crassostrea virginica*) in the Pamlico  
Sound in North Carolina, despite demographic rates high for the species within the  
reserves (Puckett and Eggleston, 2016).

465 We do not find clear evidence for network persistence for our sites despite esti-  
mates of the mean dispersal distance of *A. clarkii* from previous genetic work (11  
km, Pinsky et al., 2010) and from our samples (Catalano et al., in prep) that are well  
468 within the 30 km span of our sites. Though the length of our sampling region is more  
than twice the mean dispersal distance, usually sufficient for persistence of a popula-

tion in an isolated reserve (e.g. Lockwood et al., 2002), our sampling region contains  
471 only about 20% habitat, rather than a continuous stretch. Our largest site, Haina, is  
only about 0.8km wide, about 10 (CHECK!) times smaller than the mean dispersal  
distance, so our sites are likely too small relative to dispersal distances to see self-  
474 persistence in this system. In our site with the highest LEP, a width of XX would be  
required for self-persistence. It is also possible that the overall habitat coverage along  
the coastline in our sampling region is too low to support network persistence. Our  
477 sensitivity test to proportion habitat suggests that about 2.75 times more habitat in  
our sampling region would give a best estimate with network persistence, with 50%  
of our estimates showing persistence ( $\lambda_{C_{DD}} \geq 1$ ) at XX proportion habitat. The  
480 suggestion of a habitat shortage in our sampling region is partially dependent on  
our assumption that larvae land in non-habitat between patches and die, however.  
Larvae have some navigational and habitat-finding capabilities (CITATIONS), so  
483 we could be underestimating their ability to find habitat in our calculations, which  
would decrease the amount of habitat required for network persistence.

We suggest that our region is a sink area of a larger metapopulation but the area  
486 required for the larger metapopulation depends on the production and connectivity  
of outside patches. If surrounding patch populations have a similar LRP and level of  
connectivity as our sites, increasing the area of the network to include them would  
489 also not achieve network persistence. If nearby sites have higher egg production or  
egg-recruit survival, however, it might not take much of an increase in area considered  
to create a persistent network. Nearby reef sites such as Cuatro Islas, with higher  
492 coral cover and less silt, could have higher survival of fish and be contributing recruits

to our sites.

An alternative to our suggestion that our sampled sites are a sink portion of a  
495 larger metapopulation is that variability in demography on a longer scale than our  
sampling time could lead to persistence. For example, in rockfish on the west coast  
of the United States, XXX Our study, though relatively long term, might not have  
498 captured this kind of variability. TALK ABOUT SYSTEMS WHERE WE SEE  
THIS (ROCKFISH!) AND WHY WE WOULD/WOULDN'T EXPECT TO SEE  
THIS IN THIS SYSTEM.

501 We see considerable uncertainty in our estimate of persistence metrics depending  
on the particular input values we use (Figs. 4, C.10). Our highest estimate for LRP is  
about 24 times more than our lowest estimate and our highest NP estimate is about  
504 22 larger than our lowest, spanning the range between network persistence for our set  
of sites to far from it. Measuring demographic and dispersal parameters in the field  
is challenging; in the face of limited and imperfect data, characterizing uncertainty  
507 and propagating it from our estimates of demographic and dispersal inputs through  
to our estimates of persistence metrics is important to contextualize our results. In  
our study, uncertainty in egg-recruit survival (a commonly challenging parameter  
510 to estimate, e.g. Johnson et al., 2018; Hameed et al., 2016), partially driven by  
uncertainty in how likely we are to capture recruits during sampling (Fig. C.9), has  
a large effect on whether or not we think our populations are persistent. For a marine  
513 metapopulation, our system is relatively uncomplicated and yet still hard in which  
to concretely ascertain persistence. As we accumulate more empirical assessments  
of metapopulations to compare to our expectations from theory and models, we will

516 have to think carefully about how to handle uncertainty as we move to tackling larger  
and more complicated systems.

Persistence criteria, such as those detailed in Hastings and Botsford (2006) and  
519 Burgess et al. (2014), ask whether a population at low abundance can grow and  
recover rather than going extinct. Density-dependence is often ignored at low abun-  
dances (Botsford et al., 2019) so is not explicitly considered in persistence metrics.  
522 In real populations, however, it can be challenging to estimate density-independent  
demographic rates, as density-dependence is occurring in the population as it is  
sampled. In *A. clarkii*, density-dependence is likely most important immediately  
525 post-settlement, as for many fish species, but is also relatively easy to measure at  
that point and accounted for in our analyses. Density-dependence could continue to  
be important throughout the life history, however, due to the social hierarchies in  
528 colonies of clownfish (e.g. Buston and Elith, 2011). In other species of clownfish, indi-  
viduals on the same anemone maintain strict size spacing, restricting their food intake  
and growth to avoid encroaching on the position of another fish and being attacked  
531 or evicted (seen in *A. percula*, Buston, 2003a,b). This suggests that while fish are in  
the pre-reproductive queue, density-dependence may lower growth rates compared to  
the growth of fish alone on an anemone, as would be the case in a population at low  
534 abundance. We attempt to account for the primary effect of density-dependence on  
our estimate of egg-recruit survival but other estimates, particularly growth and sur-  
vival, would also likely be higher in the absence of density-dependence and increase  
537 LRP.

Our estimates of persistence metrics suggest that it is possible but not likely that

the region of sites we sample persist as a network without outside input, despite  
540 covering an area more than twice the estimated mean dispersal distance for our focal  
species. Our estimate of LRP near the threshold of one required for replacement  
(slightly  $< 1$  when we do not correct for density-dependence, slightly  $> 1$  when we  
543 do), suggests that dispersal is not likely the primary reason our sites do not persist  
as a network. If density-dependence is strongly present in our data such that our  
corrected estimate is the best, then our sites could persist if there were no losses  
546 to dispersal. Otherwise, our sites do not produce enough offspring for replacement  
regardless of dispersal patterns, possibly due to worsening habitat quality. This is  
a reminder that dispersal is only part of the persistence story for metapopulations;  
549 even areas that seem large enough to contain a persistent network based on dispersal  
distance will not be able to persist in isolation if they have low production and  
survival of offspring. We do find recruits coming back to our region, and even to  
552 their natal site, but broader connectivity to more productive sites likely enables our  
sites to persist.

These are somewhat high values of egg-recruit survival compared to what we see  
555 elsewhere in the literature (e.g. Rumrill, 1990; Metaxas and Saunders, 2009) (though  
not unreasonable, e.g. White et al., 2014; Johnson et al., 2018) because we scale up  
by the amount of habitat in our sampling area and count mortality due to dispersal  
558 to non-habitat in the dispersal probability, rather than in  $S_e$ .

## Figures

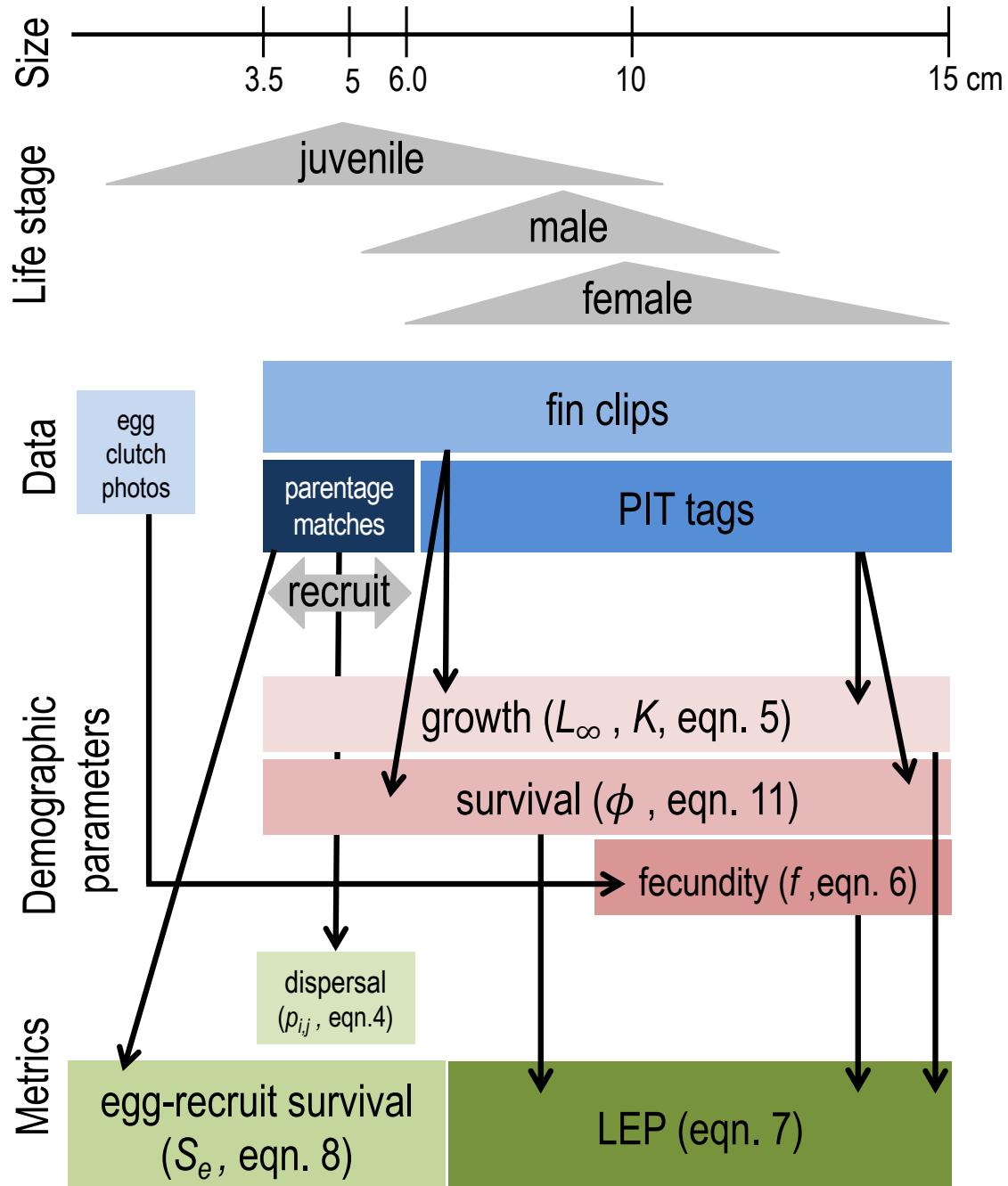


Figure 1: The data collected for fish at each life stage and how they match to the equations and metrics estimated. We consider recruits to be offspring in their first year of settlement, represented by the 3.5–6.0 cm range.

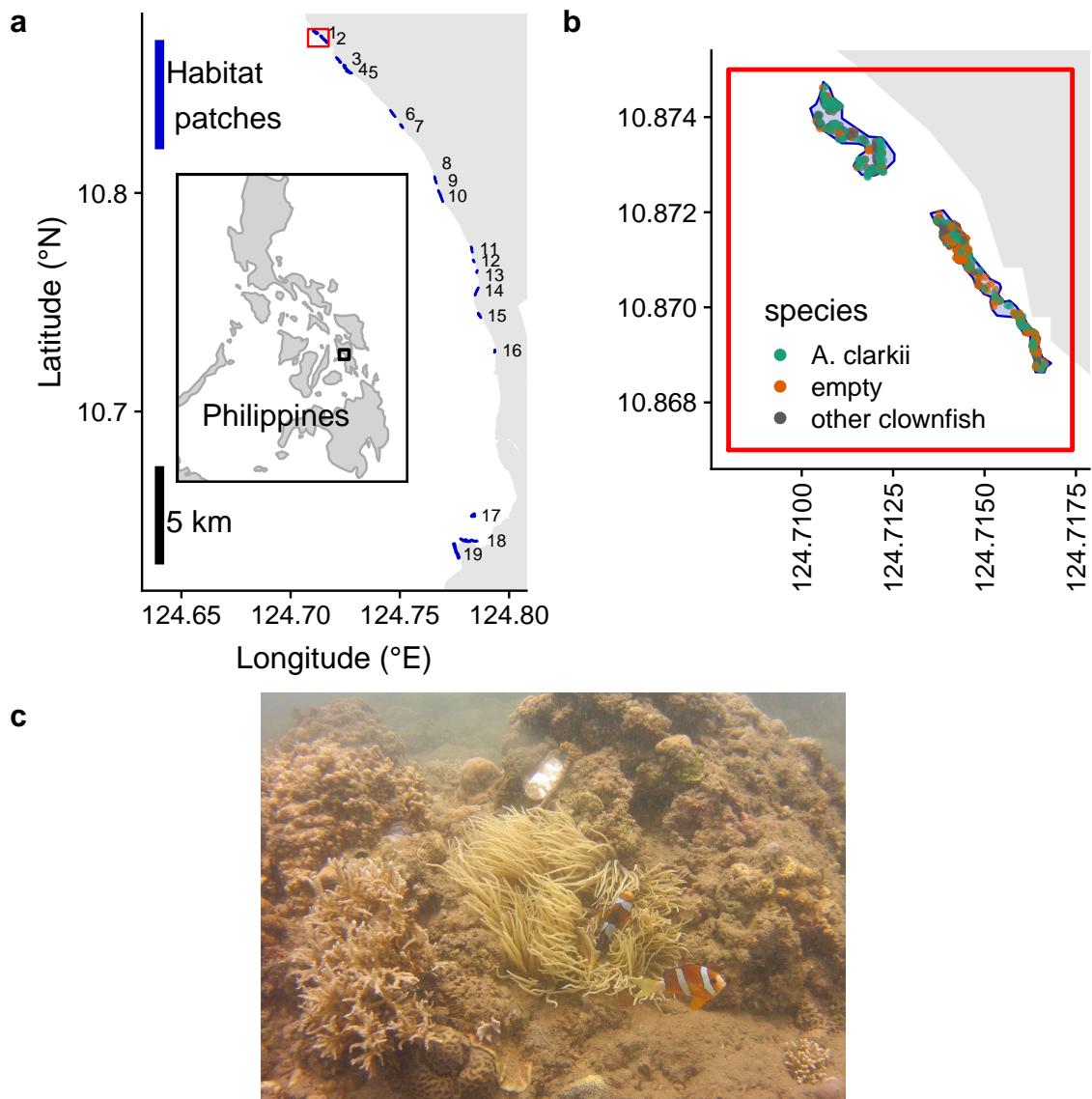


Figure 2: a) Map of the sites along the coast of Leyte in the Philippines. From north to south, the patches are: 1) Palanas, 2) Wangag, 3), North Magbangon, 4) South Magbangon, 5) Cabatoan, 6) Caridad Cemetery, 7) Caridad Proper, 8) Hicop South, 9) Sitio Tugas, 10) Elementary School, 11) Sitio Lonas, 12) San Agustin, 13) Poroc San Flower, 14) Poroc Rose, 15) Visca, 16) Gabas, 17) Tomakin Dako, 18) Haina, and 19) Sitio Baybayon. b) Zoomed-in map of the two northern-most sites, Palanas and Wangag, to show anemone arrangement with anemones colored as occupied by *A. clarkii* (green), occupied by other clownfish species (orange), or unoccupied by clownfish (grey). c) An example anemone occupied by *A. clarkii* in a typical habitat at the sites. The metal anemone tag is visible just above the anemone on the rock.

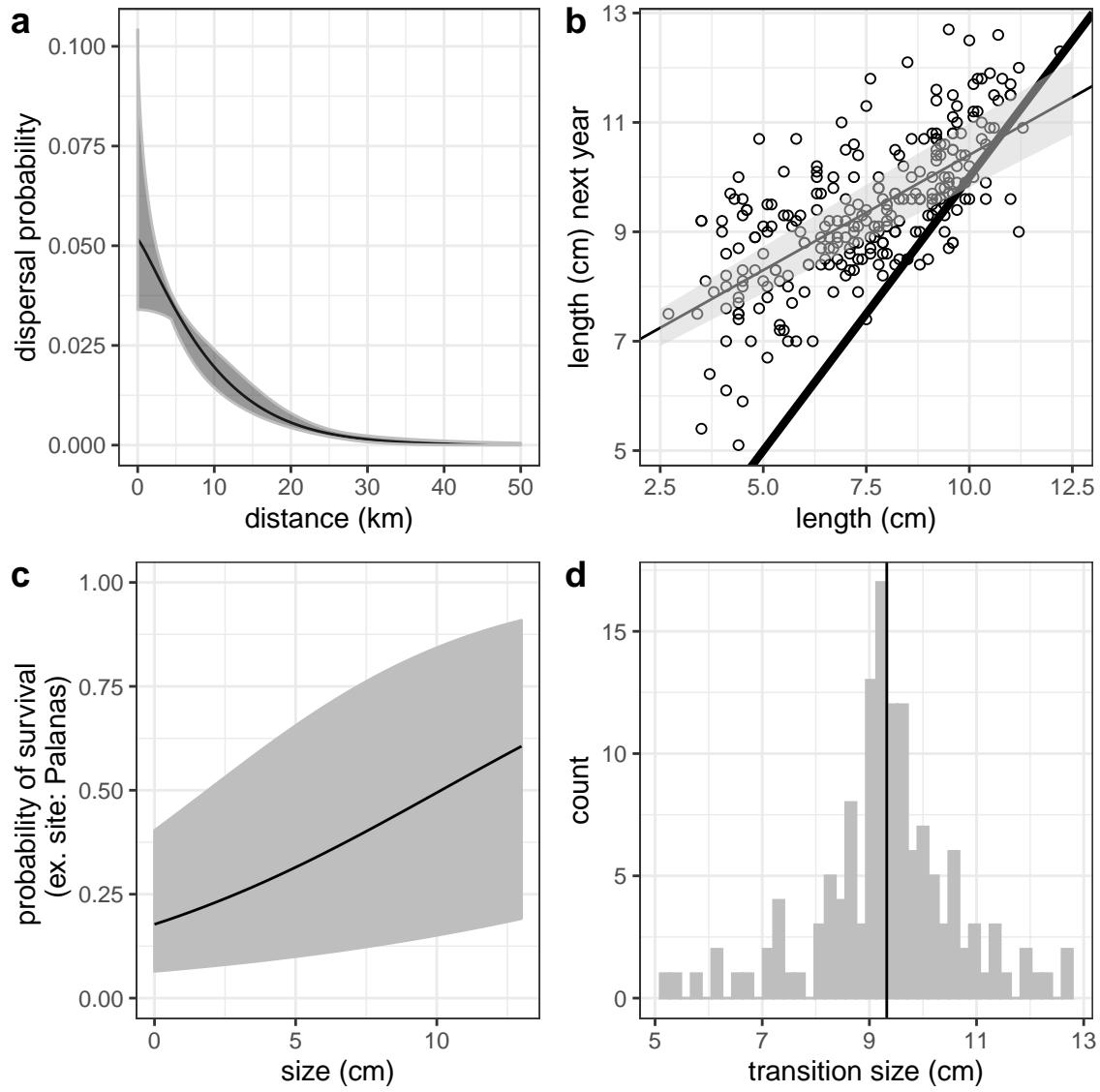


Figure 3: Best estimates (solid black line) and uncertainty (grey) for dispersal (a), growth, including the 1:1 line in thick black (b), post-recruit annual survival at Palanas as an example site (c), and size at female transition (d) parameters. Best est

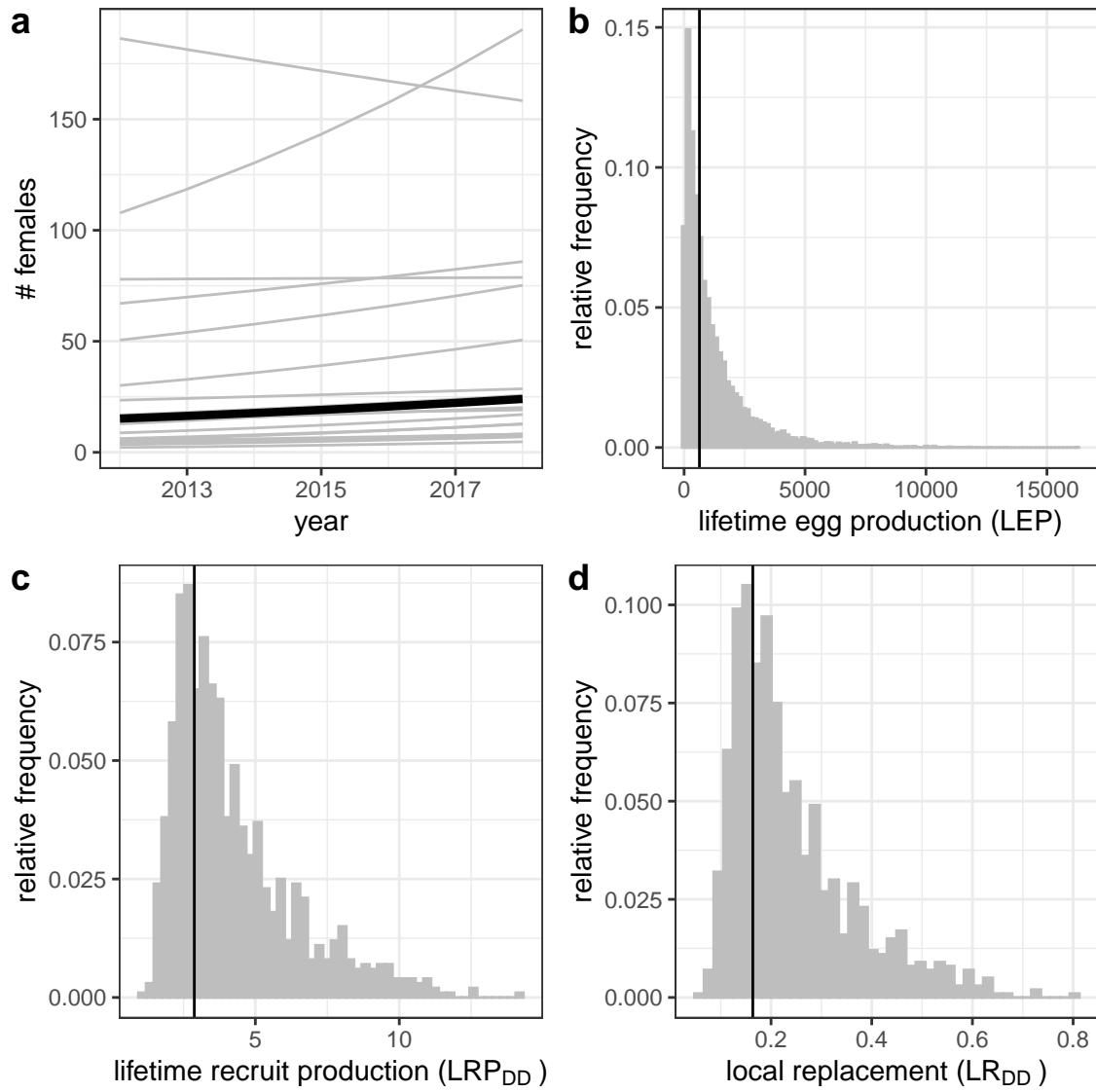


Figure 4: Estimates of a) estimated abundance of females over time at each individual site (grey lines) and for an average site (black line), b) individual-site LEP for all sites with the best estimate averaged across sites (black line), c) average  $LRP_{DD}$  across sites, and d) local replacement, showing the best estimate (black solid line) and range of estimates considering uncertainty in the inputs (grey). Estimates of LRP and local replacement include compensation for density-dependent mortality in early life stages.

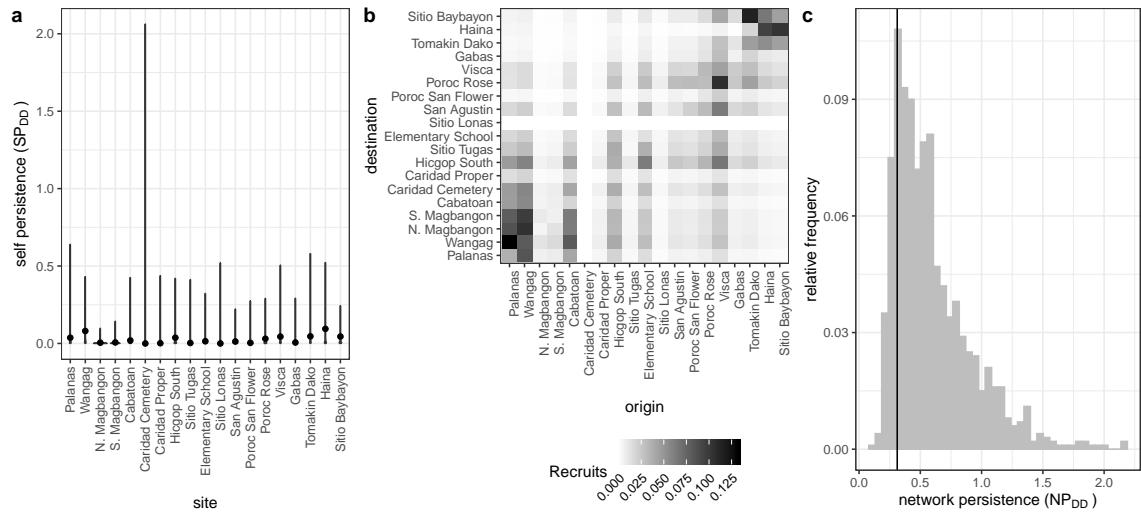


Figure 5: Values of self-persistence (a), realized connectivity among sites (b), and network persistence (c). All estimates include compensation for density-dependence in early life stages. For self-persistence (a) and network persistence (c), the best estimate is shown with in black (point for (a), line for (c)) and the range of estimates considering uncertainty is shown in grey.

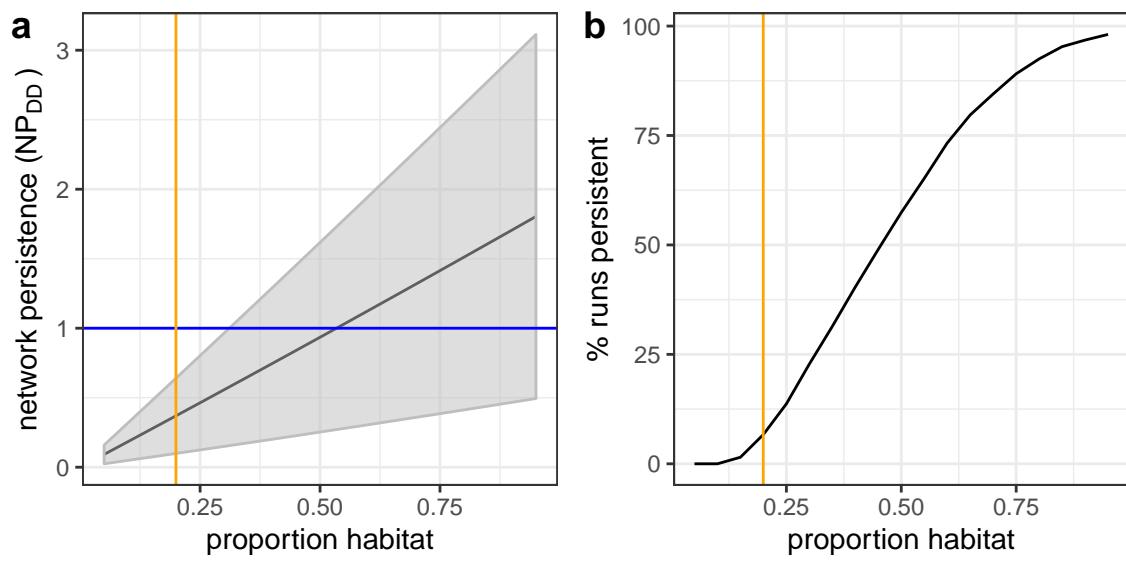


Figure 6: Sensitivity of network persistence to the proportion of the sampling region that is habitat. (a) The best estimate of  $\lambda_{cDD}$  with the standard deviation of the estimates with uncertainty for 19 patches of equal size and spacing with adult survivals for an average patch. (b) The percentage of estimates from the runs in (a) with  $\lambda_{cDD} \geq 1$  with increasing proportion habitat.

# Appendix

## <sup>561</sup> A Summary of parameters

Parameter	Description	Best estimate	Range in uncertainty runs	Notes
$k_d$	scale parameter in dispersal kernel	-2.11	-2.36 to -1.96	estimated using methods in Bode et al. (2018) in Catalano et al. (in prep)
$\theta$	shape parameter in dispersal kernel	1	NA	estimated using methods in Bode et al. (2018) in Catalano et al. (in prep)
$L_\infty$	average asymptotic size in von Bertalanffy growth curve	10.70 cm	10.50 to 10.90 cm	

$K$	growth coefficient in von Bertalanffy growth curve	0.864	0.785 to 0.944	
$b_\phi$	intercept for adult survival	-1.82	$\pm 0.231$ standard error	on a log-odds scale
$b_a$	size effect for adult survival	0.169	$\pm 0.028$ standard error	on a log-odds scale
$b_{pr}$	intercept for recapture probability from mark-recapture analysis	2.10	$\pm 0.849$ standard error	on a log-odds scale, not used in persistence estimates
$b_1$	size effect for recapture	-0.161	$\pm 0.088$ standard error	on a log-odds scale, not used in persistence estimates
$b_2$	distance effect for recapture	-0.196	$\pm 0.023$ standard error	on a log-odds scale, not used in persistence estimates

size <sub>recruit</sub>	size (cm) of recruited offspring	mean of size of offspring in parentage analysis = 4.37 cm	3.5 - 6.0 cm	drawn from uniform distribution across range
size <sub>recruit, sd</sub>	standard deviation of size of a recruit	0.1		used in discretization of IPM for LEP
size <sub>sd</sub>	standard deviation distribution of sizes of a fish in the next year	1.45		used in discretization of IPM for LEP, estimated from range of sizes for fish starting at 7.4-7.6 cm and recaptured a year later
$b_e$	coefficient for eyed eggs	-0.608		Yawdoszyn et al. (in prep)
$b_l$	size effect in eggs-per-clutch relationship	2.39		Yawdoszyn et al. (in prep)
$b$	intercept in eggs-per-clutch relationship	1.17		Yawdoszyn et al. (in prep)

$c_e$	egg clutches per year	11.9		Holtswarth et al. (2017)
$L_f$	size at transition to female	9.32cm	5.2 - 12.7cm	drawn from distribution in data
$P_h$	proportion of sites sampled cumulatively across time	0.41		details in B.1
$P_d$	proportion of dispersal kernel area from each site covered by our sampling region	0.57		details in B.3
$P_c$	probability of capturing a fish	0.56	drawn from beta distribution with parameters $\alpha_{P_c} = 1.44$ and $\beta_{P_c} = 1.13$	details in B.2

$P_s$	proportion of our sampling region that is habitat	0.20		details in B.3.0.1
DD	proportion of habitat that would be available without density-dependence at settlement	1.71		
$p_U$	proportion of anemones unoccupied by clownfish	0.53		used to estimate DD
$p_A$	proportion of anemones occupied by <i>A. clarkii</i>	0.38		used to estimate DD
$L_s$	minimum size in LEP IPM	0		eqn. 7

$U_s$	maximum size in LEP IPM	15		eqn. 7
-------	----------------------------	----	--	--------

Table A1

## B Method details

### B.1 Proportion of habitat sampled

We used tagged anemones to estimate the proportion of habitat sampled at each site in each year ( $P_{h_{i,t}}$ ). We tagged each anemone that is home to *A. clarkii*, with a metal tag, which is relatively permanent and easy to re-sight (the anemone tag is visible above the anemone in Fig. 2c), so we consider the total number of metal tags at each site to be the total number of anemones that are habitat. We divide the number of tagged anemones visited each sampling year by the total number of metal tags at that site to get the proportion of habitat sampled. We use proportion of anemones rather than proportion of total site area because anemones, and therefore habitat quality, are unevenly distributed across the site; areas we did not visit are likely to have a lower density of anemones than the areas we did sample.

For scaling the number of tagged recruited offspring to account for areas of our sites we did not sample, we use the overall proportion habitat sampled across all sites and sampling years ( $P_h$ ). We sum the metal-tagged anemones we visited across all sites and years to get the total number of metal-tagged anemones we visited while sampling. We then divide that by the number of anemones we could have sampled,

Site	# Total anems	% Habitat surveyed						
		2012	2013	2014	2015	2016	2017	2018
Cabatoan	26	42	58	58	65	73	0	62
Caridad Cemetery	4	0	75	50	0	50	50	50
Elementary School	8	0	100	38	88	88	88	100
Gabas	9	0	0	0	44	44	67	0
Haina	104	0	6	13	13	10	56	80
Hicgop South	18	0	67	22	28	56	83	78
N. Magbangon	105	5	12	40	63	63	0	5
S. Magbangon	34	41	56	32	0	65	0	71
Palanas	137	29	58	47	63	85	86	86
Poroc Rose	13	100	100	69	31	23	69	69
Poroc San Flower	11	100	82	73	73	55	82	64
San Agustin	17	94	65	71	65	100	82	76
Sitio Baybaon	260	0	14	30	33	30	41	80
Tomakin Dako	50	0	24	22	36	34	60	68
Visca	13	100	100	23	38	62	85	62
Wangag	296	18	32	42	34	26	49	68

Table A2: Table showing the percent of metal-tagged anemones surveyed at each site in each sampling year.

579 the sum of total metal-tagged anemones across all sites multiplied by the number of sampling years, to get the overall proportion habitat sampled across our sites and sampling years.

582 **B.2 Probability of capturing a fish, from recapture dives**

We use mark-recapture data from recapture dives done within a sampling season to estimate the probability of capturing a fish. During some of the sampling years, 585 portions of the sites were sampled again within a few weeks of the original sampling dives. We assume there is no mortality of tagged fish between the original sampling dives and the recapture dives because they are so close in time and that fish do not 588 change their behavior or response to divers, so therefore assume that the probability of recapturing a fish is the same as the probability of capturing a fish on a sampling dive. For each recapture dive, we use GPS tracks of the divers to identify the anemones 591 covered in the recapture dive and the set of PIT-tagged fish encountered on those anemones during the original sampling dives. We estimate the probability of capture  $P_c$  as the number of tagged fish caught during the capture dive  $m_2$  divided by the 594 total number of fish caught on the recapture dive  $n_2$ :  $P_c = \frac{m_2}{n_2}$ . The value of  $P_c$  from each recapture dive is reported in Table A3.

We use the mean  $P_c$  across all 14 recapture dives, covering XX sites in 3 sampling 597 seasons (2016, 2017, 2018), as our best estimate. Because there are so few recapture dives compared to the number of times we calculate the metrics to show the range of uncertainty, we represent the probability of capture as a distribution, rather than 600 sampling directly from the values calculated for each recapture dive. The distribution of capture probabilities across the 14 dives is quite skewed so we represent it as a beta distribution, using the mean  $\mu_{P_c}$  and variance  $V_{P_c}$  of the set of 14 values to find 603 the appropriate  $\alpha_{P_c}$  and  $\beta_{P_c}$  parameters, where

$$\alpha_{P_c} = \left( \frac{1 - \mu_{P_c}}{V_{P_c}} - \frac{1}{\mu_{P_c}} \right) \mu_{P_c}^2 \quad (\text{B.1})$$

$$\beta_{P_c} = \alpha_{\mu_{P_c}} \times \frac{1}{\mu_{P_c} - 1}. \quad (\text{B.2})$$

The mean of the individual capture probability values is  $\mu_{P_c} = 0.56$ , with variance  $V_{P_c} = 0.069$ , which gives beta distribution parameters  $\alpha_{P_c} = 1.44$  and  $\beta_{P_c} = 1.13$ . We  
 606 sample 1000 values from the beta distribution, then truncate the sample to only values larger than the lowest value of  $P_c$  estimated in an individual dive (0.20), to avoid extremely low values that are sometimes randomly sampled from the distribution  
 but are unrealistically low. We then sample with replacement from the truncated set to get a vector of values the length of the number of runs.

*Talk to Katrina to fill out table with data that went into recapture calcs*

Site	Year	$m_2$	$n_2$	$P_c$
				0.56
				0.26
				0.89
				0.67
				0.20
				0.83
				0.47
				0.20
				0.83
				1.0
				0.33
				0.58
				0.63
				0.41

Table A3: Table showing the site, year, number of fish caught on sample dives on the anemones resampled during the recapture dive ( $m_2$ ), number of fish caught on the recapture dive ( $n_2$ ) for each recapture dive.

### 612 **B.3 Scaling up recruits**

To estimate the total number of offspring produced by our genotyped parents that survive to recruitment, we scale up the number of matched offspring caught during 615 sampling ( $R_m$ ) to account for recruits we could have missed (Fig. B.1). We scale up by 1) the cumulative proportion of habitat we sampled at our sites over time ( $P_h$ ) to account for recruits at anemones we did not sample (details in B.1), 2) the 618 probability of capturing a fish if we sampled its anemone ( $P_c$ ) to account for fish that escaped during sampling (details in B.2), 3) the proportion of the dispersal kernel from our sites within of our sampling region ( $P_d$ ) to account for fish that dispersed 621 outside of our sampling area (details in B.3), and 4) the proportion habitat in our

sampling region ( $P_s$ ) to avoid counting mortality of fish dispersing to non-habitat within our region twice (in both the estimate of total recruits and in the integrated dispersal kernel) (details in B.3.0.1), and 5) the proportion of anemones occupied by *A. clarkii* (DD) to account for density-dependent mortality of settling recruits.

### How could we have missed potential recruits originating from our sites?

- 1) Failed to catch recruit when sampling ( $P_c$ )
- 2) Missed sampling some habitat areas within our sites ( $P_h$ )
- 3) Recruit dispersed outside our study region ( $P_d$ )
- 4) Recruit dispersed to non-habitat within our region ( $P_s$ )
- 5) Recruit died due to density-dependent competition with other settlers (DD)

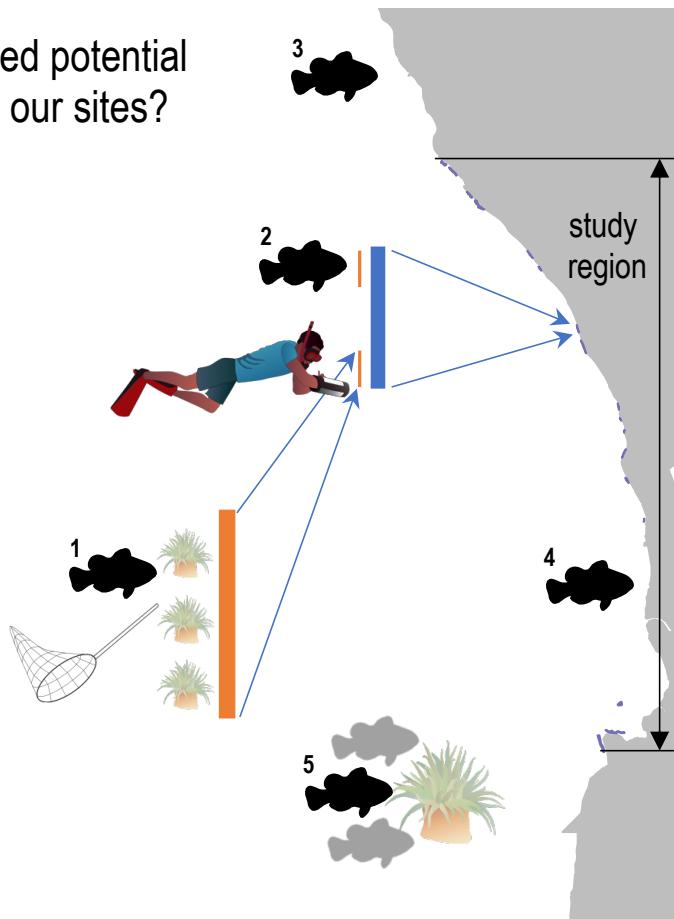


Figure B.1: Schematic of five ways we could have missed recruits while sampling and used to scale up our raw estimate of recruits from matched offspring.

## **Proportion of dispersal kernel area sampled**

- 627 To estimate the proportion of the dispersal kernel from each site covered by our sampling area, we divide the sum of the total area covered by the dispersal kernels centered
- 630 [Add in description of calculation and equation]

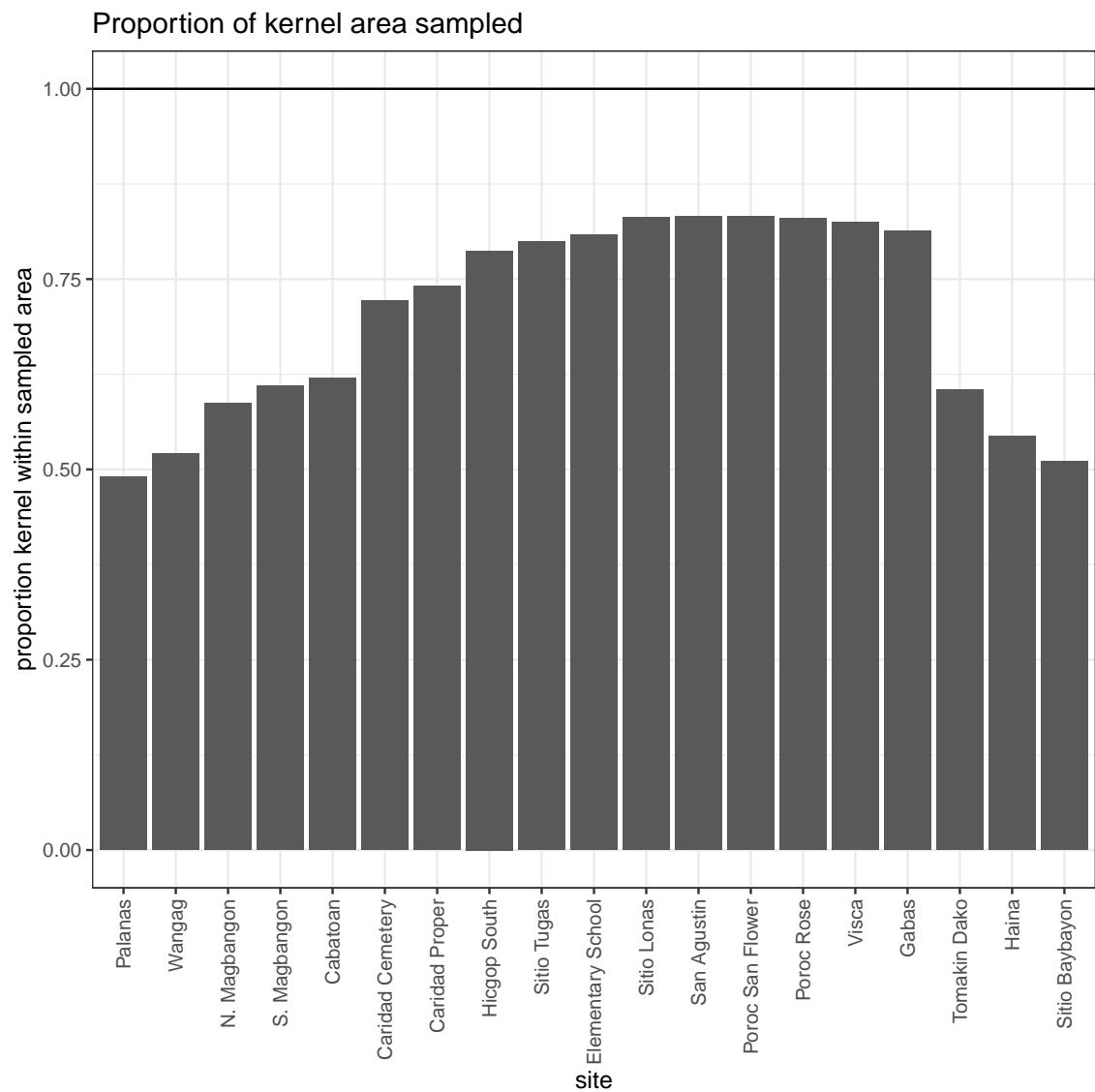


Figure B.2: Proportion of the dispersal kernel area from the center of each site covered by our sampling.

### B.3.0.1 Proportion habitat in sampling area

We assume that larvae are unable to navigate to habitat if they attempt to settle  
633 on an unsuitable patch, though clownfish larvae do likely have some ability both to  
sense good settlement areas, either by detecting host anemones (Elliott et al., 1995;  
Arvedlund et al., 1999) or conspecifics (e.g. Lecchini et al., 2005, for coral reef fish  
636 more broadly), and swim in particular direction (e.g. Bellwood and Fisher, 2001;  
Fisher, 2005). To avoid counting mortality due to settling on non-habitat twice -  
once in scaling up our matched recruits, which only includes those who settled on  
639 habitat, and once in integrating the dispersal kernel, we scale our estimate of total  
surviving recruits from our patches by the proportion of our sampling region that  
is habitat ( $P_s$ ). We find  $P_s$  by summing the lengths of all of our sites, which run  
approximately north-south, and dividing that by the total distance north-south of  
642 our sampling region, giving  $P_s = 0.20$ .

Model	Model description	AICc	dAICc
$\phi \sim S, p \sim S + D$	survival size, recapture size+distance	3348.861	0
$\phi \sim S, p \sim D$	survival size, recapture distance	3359.998	-11.1371
$\phi, p \sim D$	survival constant, recapture distance	3383.175	34.3141
$\phi, p \sim S + D$	survival constant, recapture size+distance	3384.959	36.0981
$\phi \sim t, p$	survival time, recapture constant	3408.342	59.4816
$\phi \sim i, p$	survival site, recapture constant	3440.842	91.98112
$\phi \sim i, p \sim S + D$	survival site, recapture size+distance	3440.842	91.98112
$\phi, p \sim t$	survival constant, recapture time	3453.609	104.74839
$\phi \sim S, p \sim S$	survival size, recapture size	3527.710	178.84940
$\phi, p$	survival constant, recapture constant	3570.908	222.04690

Table A4

## B.4 Full set of MARK models

645 We consider the following set of models in MARK for survival ( $\phi$ ) and recapture ( $p$ ) probability, including effects of size ( $S$ ), minimum distance from diver to anemone during surveys ( $D$ ), time ( $t$ ), and site ( $i$ ) (Table A4):

648 The best model for post-recruitment annual survival  $\phi$  has a positive size effect ( $b_a = 0.169 \pm 0.028$  SE UPDATE THESE NUMBERS!) with intercepts varying by site (eqn. B.3, Fig. B.3). The best model for recapture probability  $p_r$  has a negative effect of size ( $b_1 = -1.816 \pm 0.080$  SE UPDATE THESE NUMBERS!) and a negative effect of diver distance from anemone ( $b_2 = -0.171 \pm 0.021$  SE UPDATE THESE NUMBERS!), with intercept  $b_{p_r} = 17.93 \pm 0.858$  SE UPDATE THESE NUMBERS! (eqn. B.4, Fig. B.4), suggesting divers are less likely to recapture larger fish and those

at anemones far from areas sampled.

$$\log\left(\frac{\phi}{1-\phi}\right) = b_{\phi_i} + b_a \text{size}. \quad (\text{B.3})$$

$$\log\left(\frac{p_r}{1-p_r}\right) = b_{p_r} + b_1 \text{size} + b_2 d. \quad (\text{B.4})$$

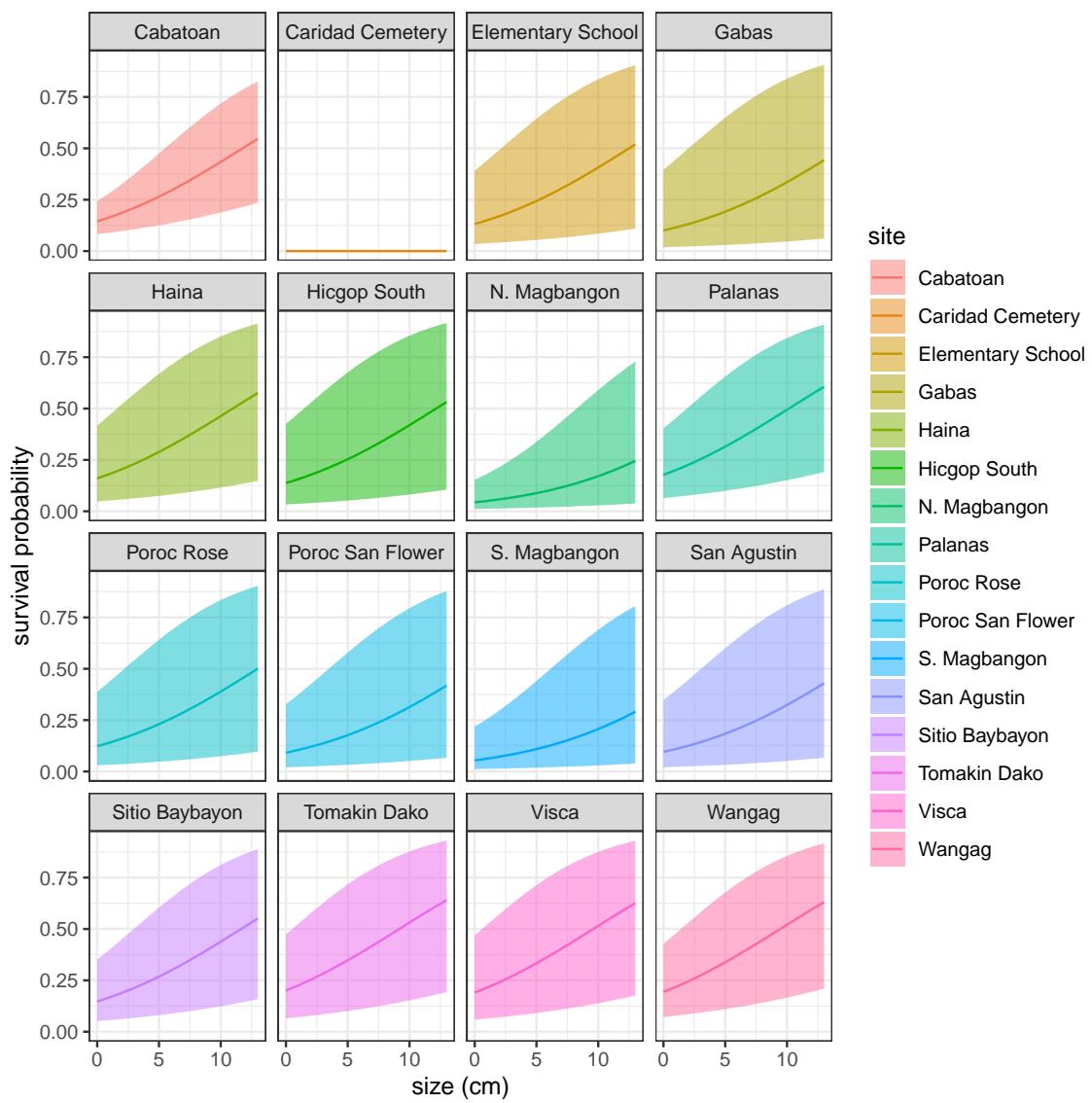


Figure B.3: Annual survival by size at each site.

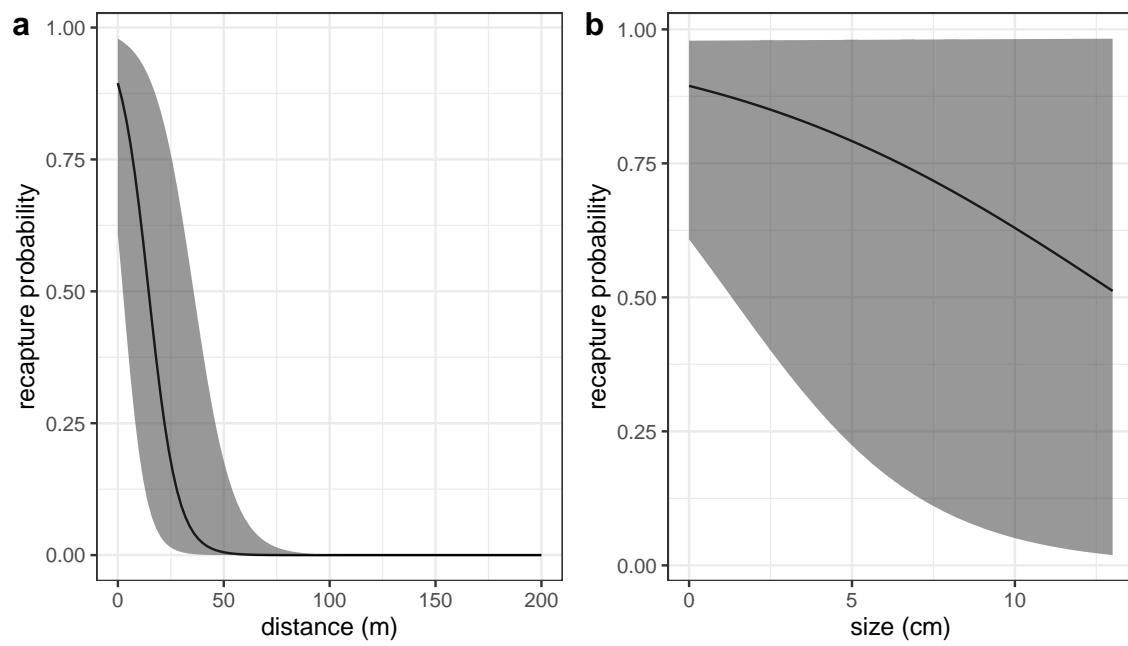


Figure B.4: Effects of a) minimum distance between divers and the anemone where the fish was first caught and b) fish size on the probability the fish will be recaptured, estimated along with survival in a mark-recapture analysis. Size has a slightly negative effect on the probability of recapture - larger fish are stronger swimmers and more likely to flee the anemone when divers approach, but with high uncertainty at larger sizes. Distance also has a negative effect on recapture, as fish tend to stay close to their anemones and are not likely to be recaptured if divers do not get close to their anemone.

## C Result details and sensitivity

### <sup>657</sup> C.1 Abundance trends by site

We use the number of females captured at each site in each sampling year, scaled by the proportion of habitat sampled at that site in that year and by the probability of <sup>660</sup> capturing a fish, to estimate abundance trends for each site (Fig. C.1).

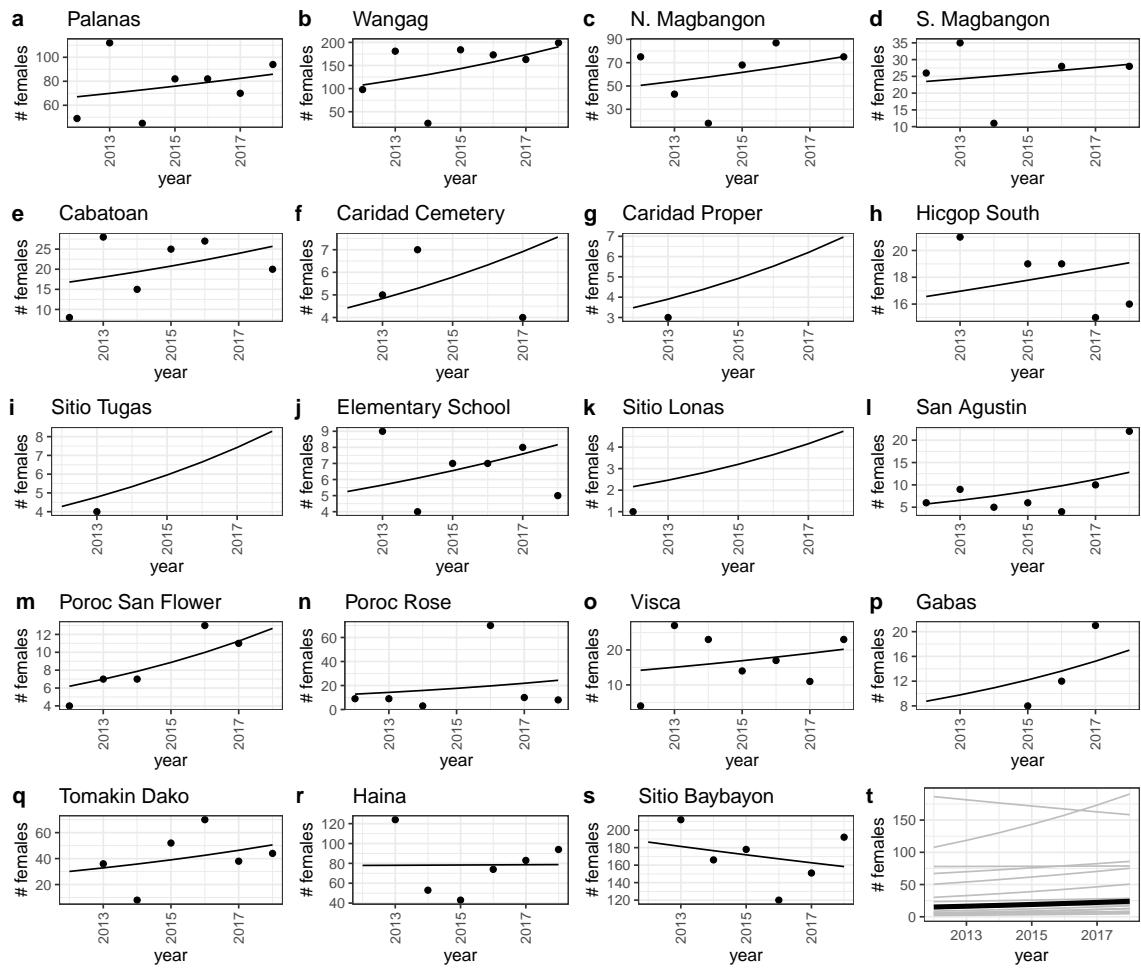


Figure C.1: Scaled number of females captured (black dots) and abundance trends (black lines) by site from a mixed effects model with site as a random effect.

## C.2 Compensating for density dependence

Estimating persistence metrics without compensating for density-dependence in our  
663 data gives us an understanding of whether individuals at our sites are able to replace  
themselves and whether our sites persist as in isolation at the current abundance  
levels, rather than at low abundance. Without compensation for early life density-  
666 dependence, all of our metrics show that the set of sites we sample is less likely to  
persist as an isolated network. We estimate egg-recruit survival ( $S_e$ ) to be 7.8e-04  
[1.2e-04, 0.033] and average lifetime recruit production (LRP) across sites to be 0.83  
669 [0.28, 3.89], with XX% of LRP estimates  $\geq 1$ . (Fig. C.2c). Our estimate of local  
replacement (LR), which estimates replacement for recruits from our sites returning  
to our sites implicitly including dispersal, is 0.09 [0.03, 0.44].

672 When we calculate LR using all arriving recruits to our sites, however, rather  
than just those originating there, the best estimate is  $> 1$  (1.22, with XX% of values  
with uncertainty  $\pm 1$ ), suggesting that there is recruit-recruit replacement at our sites  
675 when we include immigrant recruits.

We do not find any sites with a best estimate or uncertainty range of  $SP > 1$   
(Figs. C.3a), with the exception of the wide uncertainty bounds on SP for Caridad  
678 Cemetery. Our best estimate of the dominant eigenvalue of the realized connectivity  
matrix  $\lambda_c$  is 0.21 [0.07, 0.92] with  $p(\lambda \geq 1 = XX)$  (Fig. C.3c).

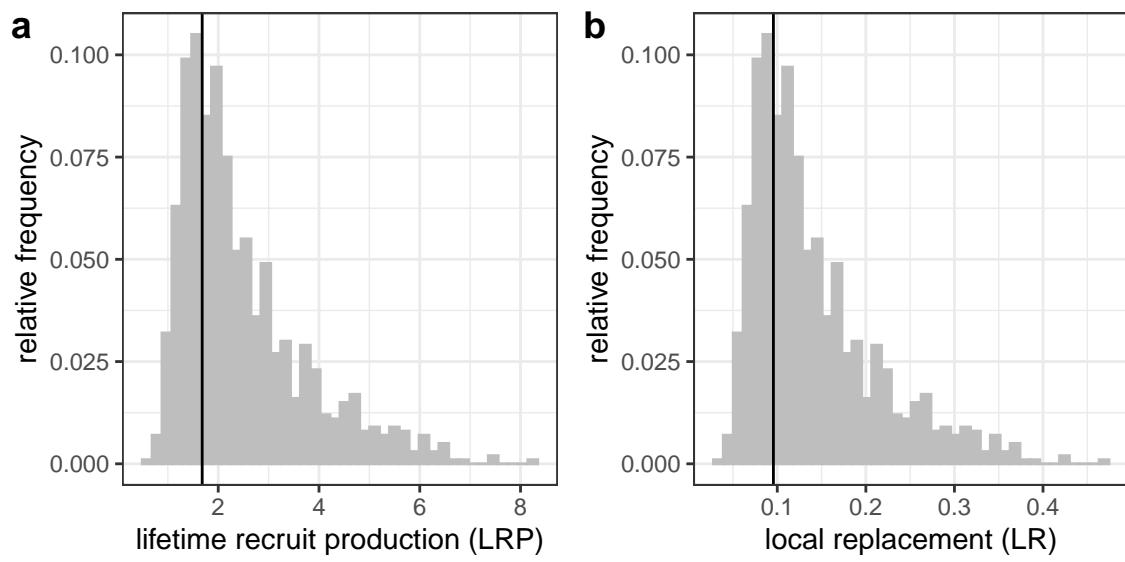


Figure C.2: Estimates of a) LRP, and b) local replacement without compensating for density dependence in early life stages in our data, showing the best estimate (black solid line) and range of estimates considering uncertainty in the inputs. These estimates compare to those in 4c,d, where we correct for additional mortality in early life due to density dependence.

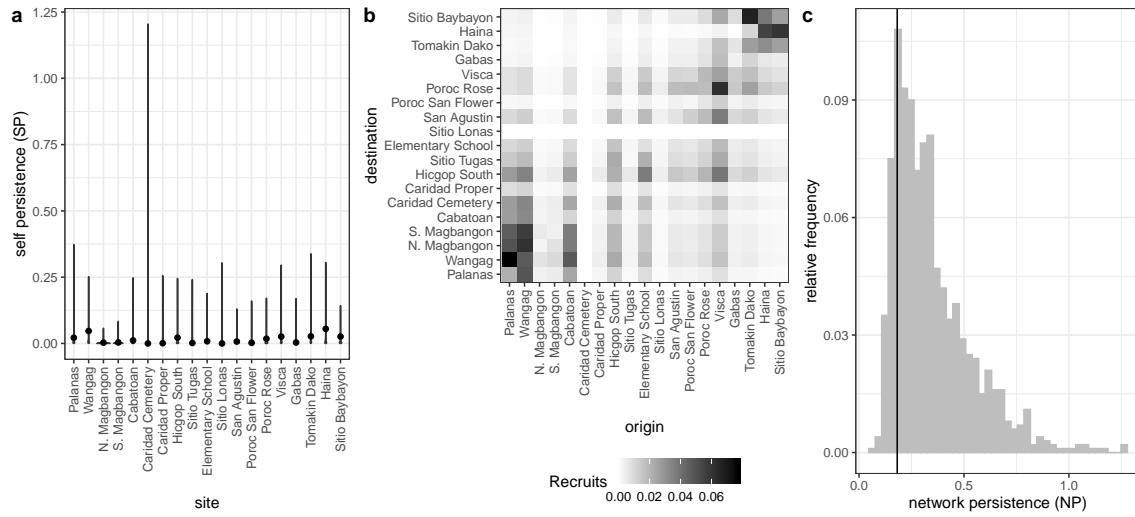


Figure C.3: Values of self-persistence (a), realized connectivity among sites (b), and network persistence (c) without compensation for density-dependence in early life stages in our data. For self-persistence (a) and network persistence (c), the best estimate is shown with in black (point for (a), line for (c)) and the range of estimates considering uncertainty is shown in grey. These estimates compare to those in 5 where we attempt to compensate for density dependence in early life stages.

### C.3 LEP and LRP by site

681 WRITE SOME TEXT!

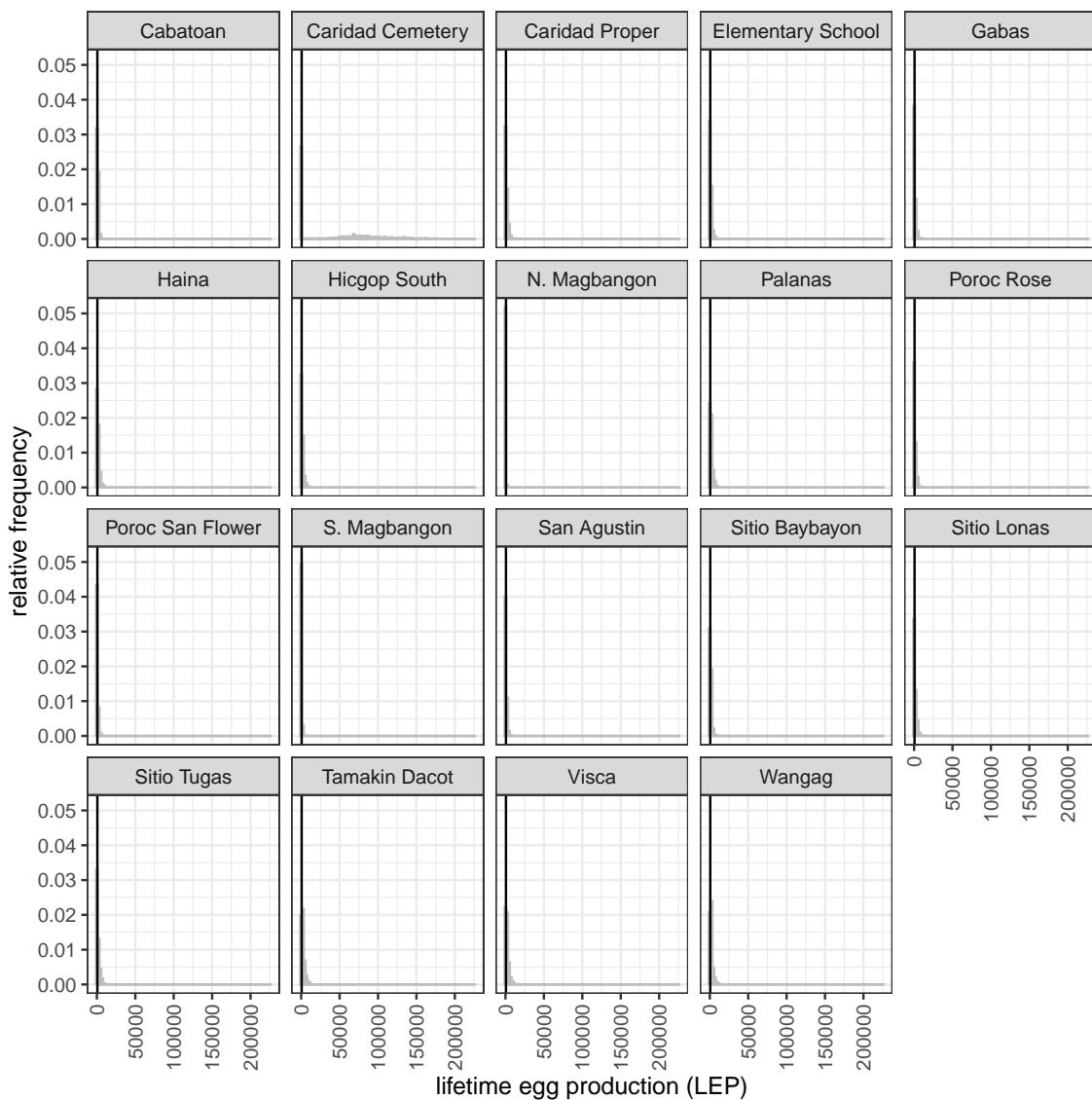


Figure C.4: Write a caption.

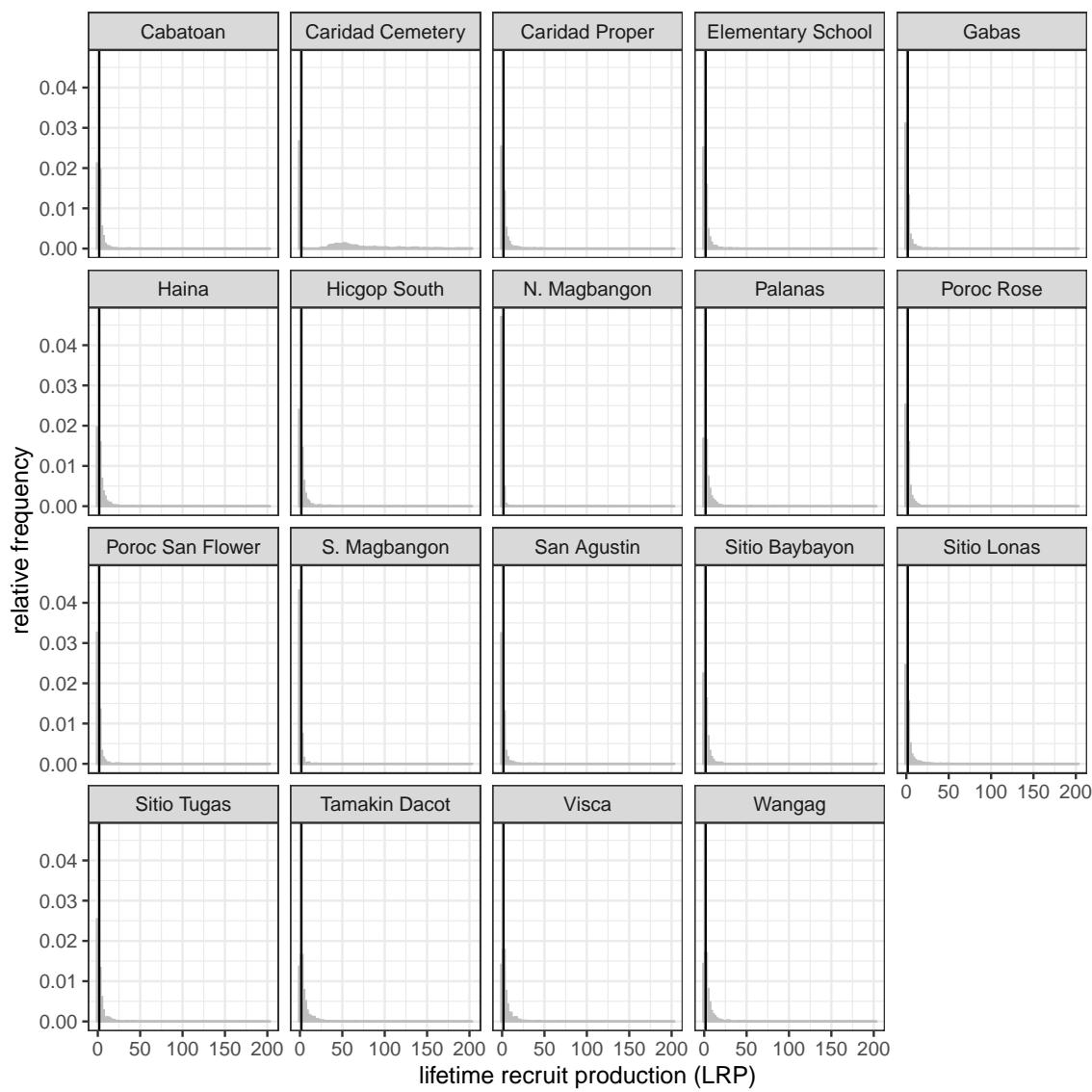


Figure C.5: Write a caption.

#### C.4 Sensitivity to parameters

EXPLAIN THAT THESE ARE THE REST OF THE PARAMETERS, NOT SHOWN

684 IN THE MAIN TEXT

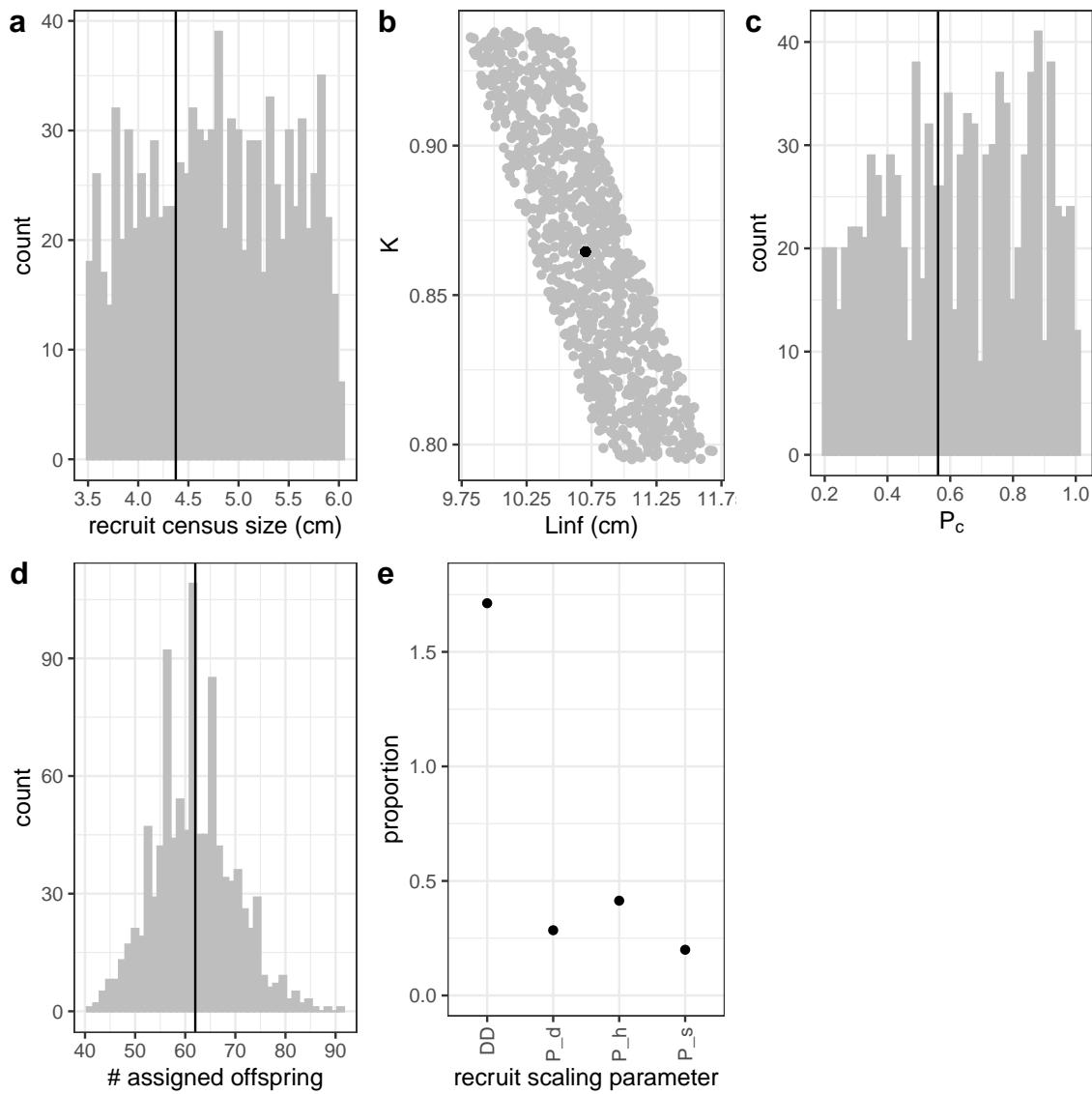


Figure C.6: Range of parameter inputs for uncertainty runs with all uncertainty included, where the black line or point is the value used for the best estimate: a)  $\text{size}_{\text{recruit}}$ , the census size at which fish are considered to have recruited after egg-recruit survival occurs; b) the parameters  $L_\infty$  and  $K$  of the von Bertalanffy growth model; c)  $P_c$ , the probability of capturing a fish; d) number of offspring assigned back to parents in the parentage analysis; e) factors that scale the number of estimated recruits from our site based on density-dependence in settler success (DD), proportion of the dispersal kernel captured by our sampling region ( $P_d$ ), the cumulative proportion of our sites we sampled over time ( $P_h$ ), and the proportion of our sampling area that is habitat ( $P_s$ ). 60

## C.5 Effects of different types of uncertainty on metrics

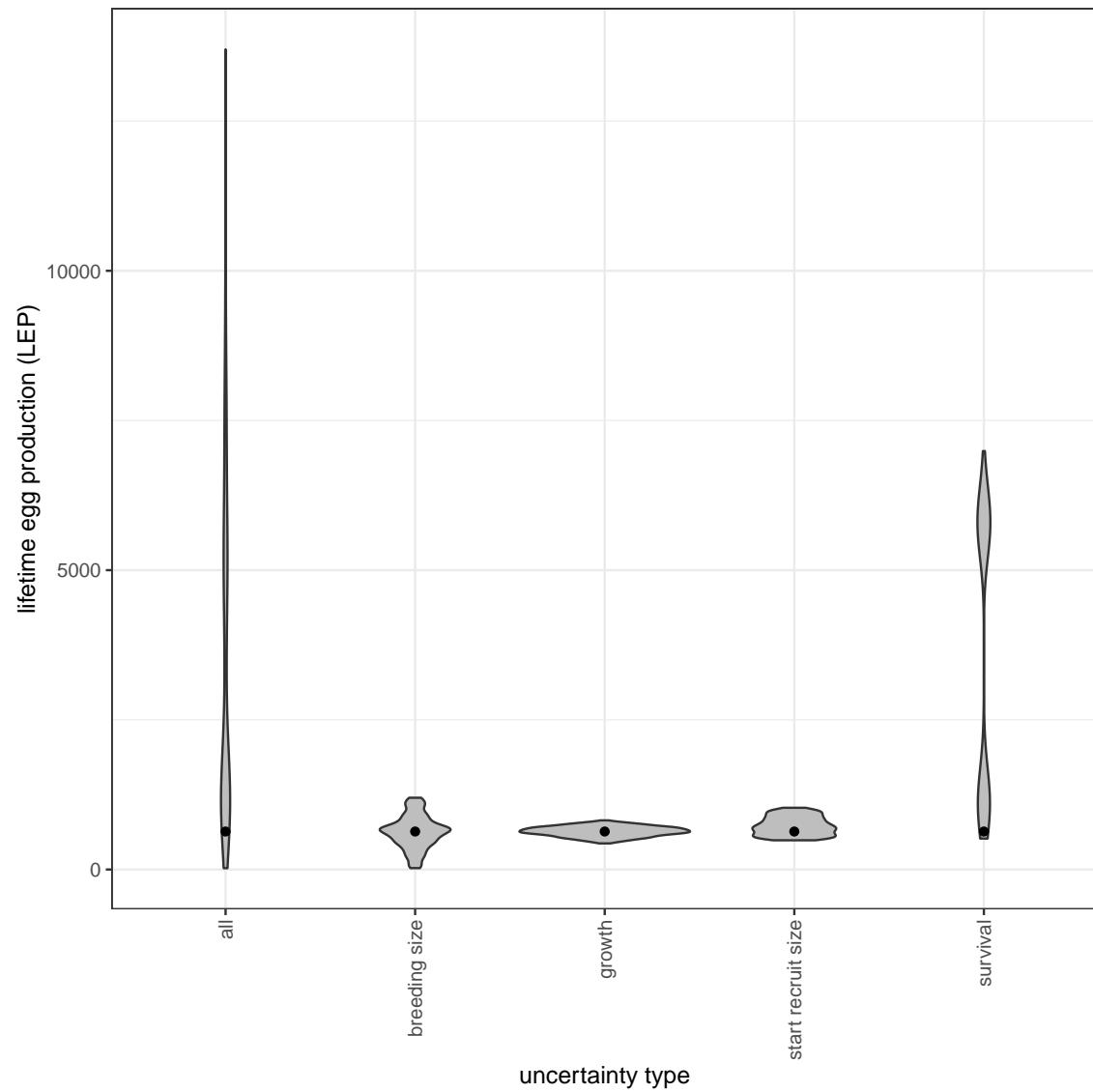


Figure C.7: The contribution of different sources of uncertainty in LEP.

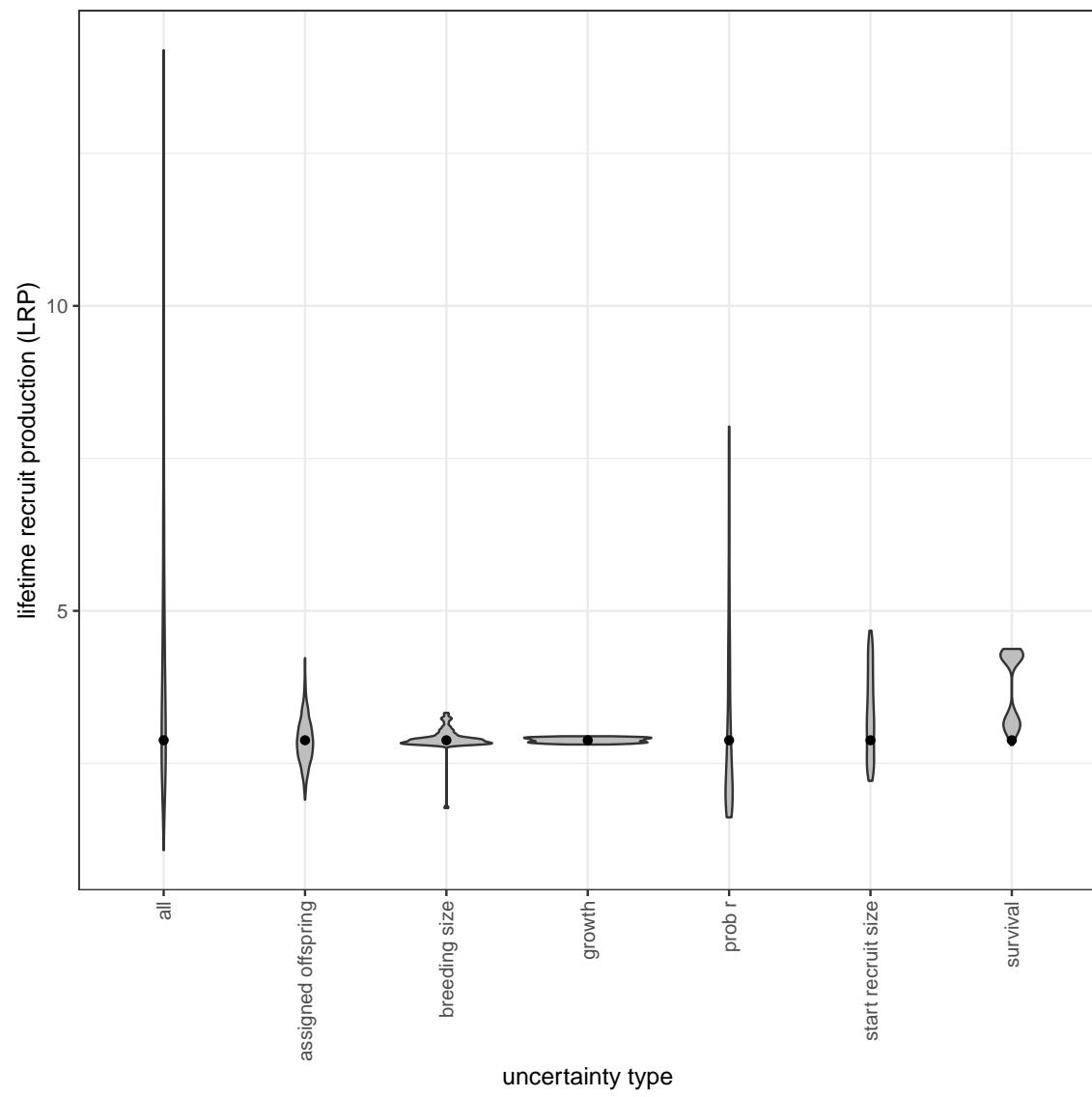


Figure C.8: The contribution of different sources of uncertainty in LRP.

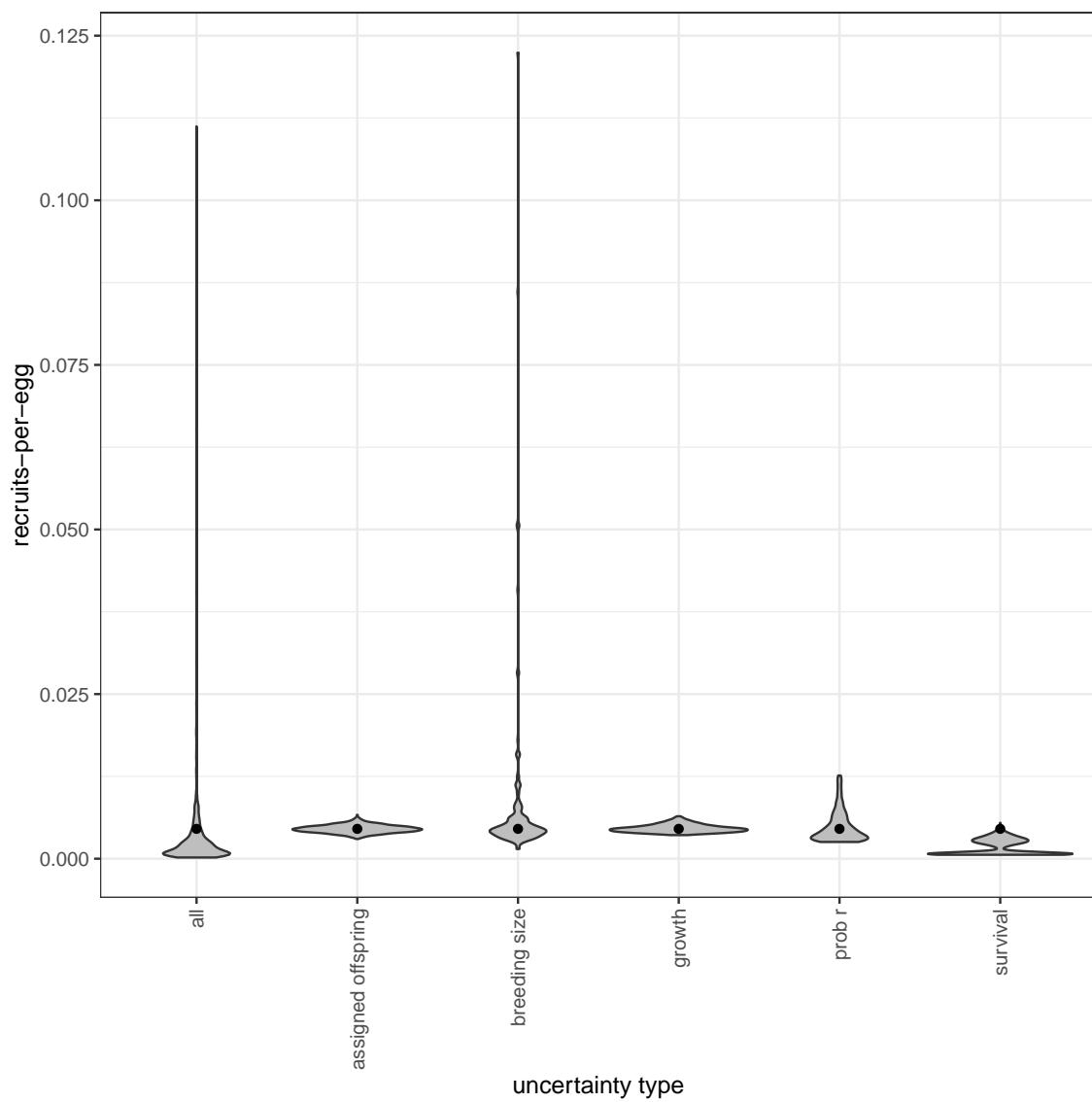


Figure C.9: The contribution of different sources of uncertainty in egg-recruit survival.

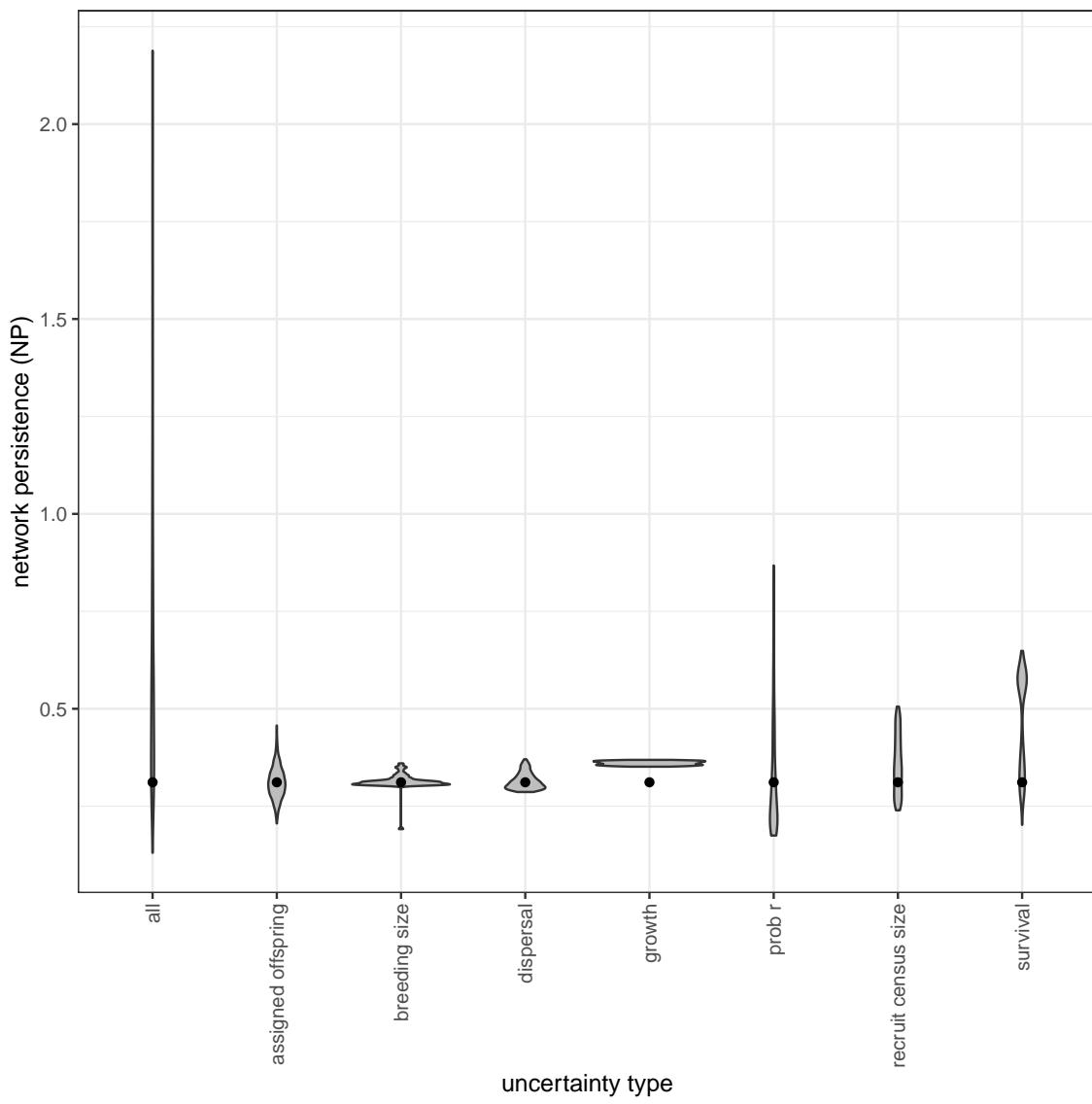


Figure C.10: The contribution of different sources of uncertainty in NP.

## References

- 687 Glenn R Almany, Serge Planes, Simon R Thorrold, Michael L Berumen, Michael  
Bode, Pablo Saenz-Agudelo, Mary C Bonin, Ashley J Frisch, Hugo B Harrison,  
Vanessa Messmer, et al. Larval fish dispersal in a coral-reef seascape. *Nature  
Ecology & Evolution*, 1:0148, 2017.
- 690 Michael Arvedlund, Mark I McCormick, Daphne G Fautin, and Mogens Bildsøe. Host  
recognition and possible imprinting in the anemonefish *amphiprion melanopus*  
693 (pisces: Pomacentridae). *Marine Ecology Progress Series*, 188:207–218, 1999.
- Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. Fitting linear mixed-  
effects models using lme4. *Journal of Statistical Software*, 67(1):1–48, 2015. doi:  
696 10.18637/jss.v067.i01.
- David R Bellwood and Rebecca Fisher. Relative swimming speeds in reef fish larvae.  
*Marine Ecology Progress Series*, 211:299–303, 2001.
- 699 Michael Bode, David H Williamson, Hugo B Harrison, Nick Outram, and Geoffrey P  
Jones. Estimating dispersal kernels using genetic parentage data. *Methods in  
Ecology and Evolution*, 9(3):490–501, 2018.
- 702 Louis W. Botsford, Alan Hastings, and Steven D. Gaines. Dependence of sustainability  
on the configuration of marine reserves and larval dispersal distance. *Ecology  
Letters*, 4:144–150, 2001.

705 Louis W Botsford, J Wilson White, and Alan Hastings. *Population Dynamics for  
Conservation*. Oxford University Press, 2019.

708 Scott C Burgess, Kerry J Nickols, Chris D Griesemer, Lewis AK Barnett, Alli-  
son G Dedrick, Erin V Satterthwaite, Lauren Yamane, Steven G Morgan, J Wilson  
White, and Louis W Botsford. Beyond connectivity: how empirical methods can  
711 quantify population persistence to improve marine protected-area design. *Ecolog-  
ical Applications*, 24(2):257–270, 2014.

Peter Buston. Forcible eviction and prevention of recruitment in the clown anemone-  
fish. *Behavioral Ecology*, 14(4):576–582, 2003a.

714 Peter Buston. Social hierarchies: size and growth modification in clownfish. *Nature*,  
424(6945):145–146, 2003b.

717 Peter M Buston and Cassidy C DAloia. Marine ecology: reaping the benefits of local  
dispersal. *Current Biology*, 23(9):R351–R353, 2013.

720 Peter M Buston and Jane Elith. Determinants of reproductive success in dominant  
pairs of clownfish: a boosted regression tree analysis. *Journal of Animal Ecology*,  
80(3):528–538, 2011.

723 Peter M Buston, Geoffrey P Jones, Serge Planes, and Simon R Thorrold. Probability  
of successful larval dispersal declines fivefold over 1 km in a coral reef fish. *Pro-  
ceedings of the Royal Society of London B: Biological Sciences*, page rspb20112041,  
2011.

- Henry S Carson, Geoffrey S Cook, Paola C López-Duarte, and Lisa A Levin. Evaluating the importance of demographic connectivity in a marine metapopulation.  
726      *Ecology*, 92(10):1972–1984, 2011.
- Hal Caswell. *Matrix population models: construction, analysis, and interpretation*.  
729      Sinauer Associates Inc., Sunderland, Massachusetts, 2nd edition, 2001.
- Katrina A Catalano, Allison G Dedrick, Michelle Stuart, Jonathan Purtiz, Humberto Montes, Jr., and Malin Pinsky. Interannual variability of genetic connectivity in  
732      a coral reef fish *Amphiprion clarkii*. in prep.
- C. C. D'Aloia, S. M. Bogdanowicz, J. E. Majoris, R. G. Harrison, and P. M. Buston. Self-recruitment in a Caribbean reef fish: a method for approximating dispersal  
735      kernels accounting for seascape. *Molecular Ecology*, 22(9):2563–2572, May 2013. ISSN 09621083. doi: 10.1111/mec.12274. URL <http://doi.wiley.com/10.1111/mec.12274>.
- JK Elliott, JM Elliott, and RN Mariscal. Host selection, location, and association behaviors of anemonefishes in field settlement experiments. *Marine Biology*, 122  
738      (3):377–389, 1995.
- Stephen P Ellner, Dylan Z Childs, Mark Rees, et al. Data-driven modelling of structured populations. *A practical guide to the Integral Projection Model*. Cham:  
741      Springer, 2016.
- Augustus J. Fabens. Properties and fitting of the von bertalanffy growth curve.  
744      *Growth*, 29:265–289, 1965.

- Daphne Gail Fautin, Gerald R Allen, Gerald Robert Allen, Australia Naturalist,  
747 Gerald Robert Allen, and Australie Naturaliste. Field guide to anemonefishes and  
their host sea anemones. 1992.
- Rebecca Fisher. Swimming speeds of larval coral reef fishes: impacts on self-  
750 recruitment and dispersal. *Marine Ecology Progress Series*, 285:223–232, 2005.
- Lysel Garavelli, J Wilson White, Iliana Chollett, and Laurent Marcel Chérubin. Population  
models reveal unexpected patterns of local persistence despite widespread  
753 larval dispersal in a highly exploited species. *Conservation Letters*, 11(5):e12567,  
2018.
- Sarah O Hameed, J Wilson White, Seth H Miller, Kerry J Nickols, and Steven G  
756 Morgan. Inverse approach to estimating larval dispersal reveals limited population  
connectivity along 700 km of wave-swept open coast. *Proceedings of the Royal  
Society B: Biological Sciences*, 283(1833):20160370, 2016.
- Ilkka Hanski. Metapopulation dynamics. *Nature*, 396(6706):41–49, 1998.  
759
- Deborah R Hart and Antonie S Chute. Estimating von bertalanffy growth parameters  
from growth increment data using a linear mixed-effects model, with an application  
762 to the sea scallop *placopecten magellanicus*. *ICES Journal of Marine Science*, 66  
(10):2165–2175, 2009.
- Alan Hastings and Louis W. Botsford. Persistence of spatial populations depends on  
765 returning home. *Proceedings of the National Academy of Sciences*, 103:6067–6072,  
2006.

Akihisa Hattori and Yasunobu Yanagisawa. Life-history pathways in relation to  
768 gonadal sex differentiation in the anemonefish, *amphiprion clarkii*, in temperate  
waters of japan. *Environmental Biology of Fishes*, 31(2):139–155, 1991.

Kina Hayashi, Katsunori Tachihara, and James Davis Reimer. Low density popu-  
771 lations of anemonefish with low replenishment rates on a reef edge with anthro-  
pogenic impacts. *Environmental Biology of Fishes*, 102(1):41–54, 2019.

Jordan N. Holtswarth, Shem B. San Jose, Humberto R. Montes Jr., James W. Morley,  
774 and Malin. L Pinsky. The reproductive seasonality and fecundity of yellowtail  
clownfish (*amphiprion clarkii*) off the philippines. *Bulletin of Marine Science*, 93,  
2017.

777 Darren W Johnson, Mark R Christie, Timothy J Pusack, Christopher D Stallings,  
and Mark A Hixon. Integrating larval connectivity with local demography reveals  
regional dynamics of a marine metapopulation. *Ecology*, 99(6):1419–1429, 2018.

780 J.L. Laake. RMark: An r interface for analysis of capture-recapture data with  
MARK. AFSC Processed Rep. 2013-01, Alaska Fish. Sci. Cent., NOAA,  
Natl. Mar. Fish. Serv., Seattle, WA, 2013. URL [http://www.afsc.noaa.gov/  
783 Publications/ProcRpt/PR2013-01.pdf](http://www.afsc.noaa.gov/Publications/ProcRpt/PR2013-01.pdf).

David Lecchini, Serge Planes, and René Galzin. Experimental assessment of sensory  
modalities of coral-reef fish larvae in the recognition of their settlement habitat.  
786 *Behavioral Ecology and Sociobiology*, 58(1):18–26, 2005.

- Dale R Lockwood, Alan Hastings, and Louis W Botsford. The effects of dispersal patterns on marine reserves: does the tail wag the dog? *Theoretical population biology*, 61(3):297–309, 2002.  
789
- Anna Metaxas and Megan Saunders. Quantifying the "Bio-" Components in Biophysical Models of Larval Transport in Marine Benthic Invertebrates: Advances and Pitfalls. *Biological Bulletin*, 216:257–272, 2009.  
792
- Haruki Ochi. Mating behavior and sex change of the anemonefish, *amphiprion clarkii*, in the temperate waters of southern japan. *Environmental Biology of Fishes*, 26(4):257–275, 1989.  
795
- Malin L Pinsky, Humberto R Montes Jr, and Stephen R Palumbi. Using isolation by distance and effective density to estimate dispersal scales in anemonefish. *Evolution*, 64(9):2688–2700, 2010.  
798
- BJ Puckett and DB Eggleston. Metapopulation dynamics guide marine reserve design: importance of connectivity, demographics, and stock enhancement. *Ecosphere*, 7(6), 2016.  
801
- H Ronald Pulliam. Sources, sinks, and population regulation. *The American Naturalist*, 132(5):652–661, 1988.
- 804 Steven S. Rumrill. Natural mortality of marine invertebrate larvae. *Ophelia*, 32(1-2):163–198, October 1990. ISSN 0078-5326. doi: 10.1080/00785236.1990.10422030. URL <http://www.tandfonline.com/doi/abs/10.1080/00785236.1990.10422030>.  
807 1990.10422030.

Ocane C. Salles, Jeffrey A. Maynard, Marc Joannides, Corentin M. Barbu, Pablo Saenz-Agudelo, Glenn R. Almany, Michael L. Berumen, Simon R. Thorrold, Geoffrey P. Jones, and Serge Planes. Coral reef fish populations can persist without immigration. *Proceedings of the Royal Society B: Biological Sciences*, 282(1819): 20151311, November 2015. ISSN 0962-8452, 1471-2954. doi: 10.1098/rspb.2015.1311. URL <http://rspb.royalsocietypublishing.org/lookup/doi/10.1098/rspb.2015.1311>.

Jinliang Wang. Sibship reconstruction from genetic data with typing errors. *Genetics*, 166(4):1963–1979, 2004.

Jinliang Wang. Estimation of migration rates from marker-based parentage analysis. *Molecular ecology*, 23(13):3191–3213, 2014.

J Wilson White, Steven G Morgan, and Jennifer L Fisher. Planktonic larval mortality rates are lower than widely expected. *Ecology*, 95(12):3344–3353, 2014.

J. Wilson White, Mark H Carr, Jennifer E Caselle, Libe Washburn, C. Brock Woodson, Stephen R Palumbi, Peter M Carlson, Robert R Warner, Bruce A Menge, John A Barth, Carol A Blanchette, Peter T Raimondi, and Kristen Milligan. Connectivity, dispersal, and recruitment: Connecting benthic communities and the coastal ocean. *Oceanography*, 32(3):50–59, 2019.

Jw White, Lw Botsford, A Hastings, and Jl Largier. Population persistence in marine reserve networks: incorporating spatial heterogeneities in larval dispersal. *Marine Ecology Progress Series*, 398:49–67, January 2010. ISSN 0171-8630, 1616-

1599. doi: 10.3354/meps08327. URL <http://www.int-res.com/abstracts/meps/v398/p49-67/>.

<sup>831</sup> Adam Yawdoszyn. Fecundity in clownfish. in prep.