

# Clownfish metapopulation persistence draft

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;

b. Oregon State University

c. Visayas State University

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

*(Author order not yet determined)*

## Introduction

Metapopulation dynamics and persistence depend on the demographic rates at each

<sup>3</sup> patch and the connectivity among patches (e.g. Hastings and Botsford, 2006; Hanski,

1998). Many marine species exist in metapopulations, consisting of patch populations connected through dispersal where connectivity affects patch dynamics though  
6 individual patches are unlikely to go extinct (Kritzer and Sale, 2006). Assessing demographic parameters and levels of connectivity has been particularly challenging for marine species, where much of the mortality and movement happens at larval and  
9 juvenile stages when individuals are hard to track and have the potential to travel long distances with ocean currents (e.g. Kritzer and Sale, 2006; Cowen and Sponaugle, 2009; Roughgarden et al., 1988). A need to understand metapopulations for  
12 conservation and management, 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.

15 For any population to persist, individuals must on average replace themselves during their lifetimes. In non-spatially structured populations, we use criteria such as the average number of recruiting offspring contributed by each individual during  
18 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). To assess replacement,  
21 metrics must take into account the demographic processes across the whole life cycle, including how likely individuals are to survive to the next age or stage, their expected fecundity at each stage, and the survival of any offspring produced to recruitment.

24 In a spatially-structured population, persistence still requires replacement but in addition to assessing whether the reproductive output and survival of a population is sufficient, we must also consider how the offspring are distributed across space.

<sup>27</sup> The spread of offspring is often described using dispersal kernels, probability density functions that give the relative number of settlers with increasing distance from the origin patch (e.g. Bode et al., 2018). Connectivity can also be described using a  
<sup>30</sup> connectivity matrix, where entries give the probability of dispersing from one patch to another, either found by discretizing the dispersal kernel or through direct estimates of pairwise exchange among patches (choose some examples to cite). A long-held  
<sup>33</sup> paradigm suggested that marine larvae were well-mixed and dispersed far on ocean currents (Roughgarden et al., 1988), suggesting widespread connectivity. With the ability to estimate connectivity through natural tags such as otolith microchemistry  
<sup>36</sup> or genetics and the realization that larvae can alter their dispersal through behavior (e.g. Morgan et al., 2009), however, the paradigm is shifting and local persistence of marine populations is seeming more possible.

<sup>39</sup> When we consider both the demographic processes within patches and the connectivity 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 loops of connectivity with other patches in the metapopulation (Hastings and Botsford, 2006; Burgess et al., 2014). In the first case (termed self-persistence), enough of the reproductive output produced at a patch is  
<sup>45</sup> retained at the patch for the patch, and therefore the metapopulation as a whole, to persist. In the second (network persistence), closed loops of connectivity among at least some of the patches - where individuals from one patch settle at another and  
<sup>48</sup> eventually send offspring back to the first in a future generation - provide the patch with enough recruitment to persist in the network. Though it has been challeng-

ing to estimate the parameters necessary to understand how actual metapopulations  
51 persist, a large work of theory developed to guide marine protected areas helps pre-  
dict when each type of persistence is likely to occur (i.e. large patches relative to the  
mean dispersal distance are likely to be self-persistent, Botsford et al., 2001).

54      New ways of identifying individuals and determining their origins, such as otolith  
microchemistry and genetic parentage analysis (e.g. Wang, 2004, 2014), however, are  
making it increasingly possible to estimate both the demographic [ADD EXAMPLE  
57 CITATIONS] and the connectivity (e.g. Hameed et al., 2016; Almany et al., 2017)  
parameters necessary to assess persistence in real metapopulations. We might expect  
that populations on isolated islands are the most likely to be self-persistent and  
60 Salles et al. (2015) find that the population of orange clownfish (*Amphiprion percula*)  
at isolated Kimbe Island in Papua New Guinea can likely persist without outside  
immigration. In contrast, a metapopulation of bicolor damselfish (*Stegastes partitus*)  
63 across four isolated islands in the Bahamas does not appear able to persist without  
outside input (Johnson et al., 2018). For populations that exist in patches along  
a continuous coastline, rather than on separate islands, however, it is still an open  
66 question of how patches interact and what is the scale of metapopulation persistence.

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  
69 Philippines. We assess persistence for all patches of habitat within a 30 km stretch  
of coastline, which exceeds estimates of the dispersal spread for this species (Pinsky  
et al., 2010), suggesting the network is likely to operate as a contained metapopula-  
72 tion. With seven years of sampling data, we are able to estimate persistence metrics

and replacement over the longer term, rather than just capturing a snapshot of the population dynamics. Our annual sampling also enables us to estimate abundance  
75 each year and investigate dynamics through time to compare with the replacement-based persistence metrics. Here, we use a long-term data set from habitat patches on a continuous set of coastline to understand persistence within a local network.

## 78 Methods

### Study system

We focus on a tropical metapopulation of yellowtail clownfish (*Ampiprion clarkii*) in  
81 the Philippines. Like many clownfish species, yellowtail clownfish have a mutualistic relationship with anemones, where small colonies of fish live (Buston, 2003; Fautin et al., 1992). Yellowtail clownfish are protandrous hermaphrodites and maintain a  
84 size-structured 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 on an anemone maintain a strict social and size hierarchy (Buston, 2003),  
87 with fish moving up in rank to become breeders only after the larger fish have died or left. In the tropical patch reef habitat of the Philippines, yellowtail clownfish spawn once per lunar month from November to May, laying clutches of benthic eggs that  
90 the parents protect and tend (Ochi, 1989). Larvae hatch after about six days and spend 7-10 days in the water column before returning to reef habitat to settle in an anemone (Fautin et al., 1992).

93 Clownfish are particularly well-suited to metapopulation studies due to their lim-

ited movement as adults and clearly patchy habitat. Once fish have settled, they tend to stay within close proximity of their anemones [XX meters, CITATION]. This  
96 makes fish easier to relocate for mark-recapture studies and simiplifies the exchange between patches to only the dispersal during the larval phase. Patches, whether considered to be the reef patch or the anemone territory of the fish, are clearly discrete  
99 and easily delineated, 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 characterstics in habitat and fish behavior  
102 make clownfish and other similarly territory-based reef fish useful model systems for studies of metapopulation dynamics and persistence (e.g. Buston and D'Aloia, 2013; Salles et al., 2015; Johnson et al., 2018).

## 105 **Field data collection**

We focus on a set of seventeen patch reef sites spanning approximately 30 km along the western coast of Leyte island in the Philippines (MAP FIGURE). The sites  
108 consist of rocky patches of coral reef and are separated by sand flats. Previous work using genetic isolation by distances estimated that yellowtail clownfish larvae have a dispersal spread of about 10 km (range 4-27 km, Pinsky et al., 2010), so our sites were  
111 selected to cover and exceed that range. On the north edge, the sites are isolated from nearby habitat with no additional reef habitat for at least 20 km.

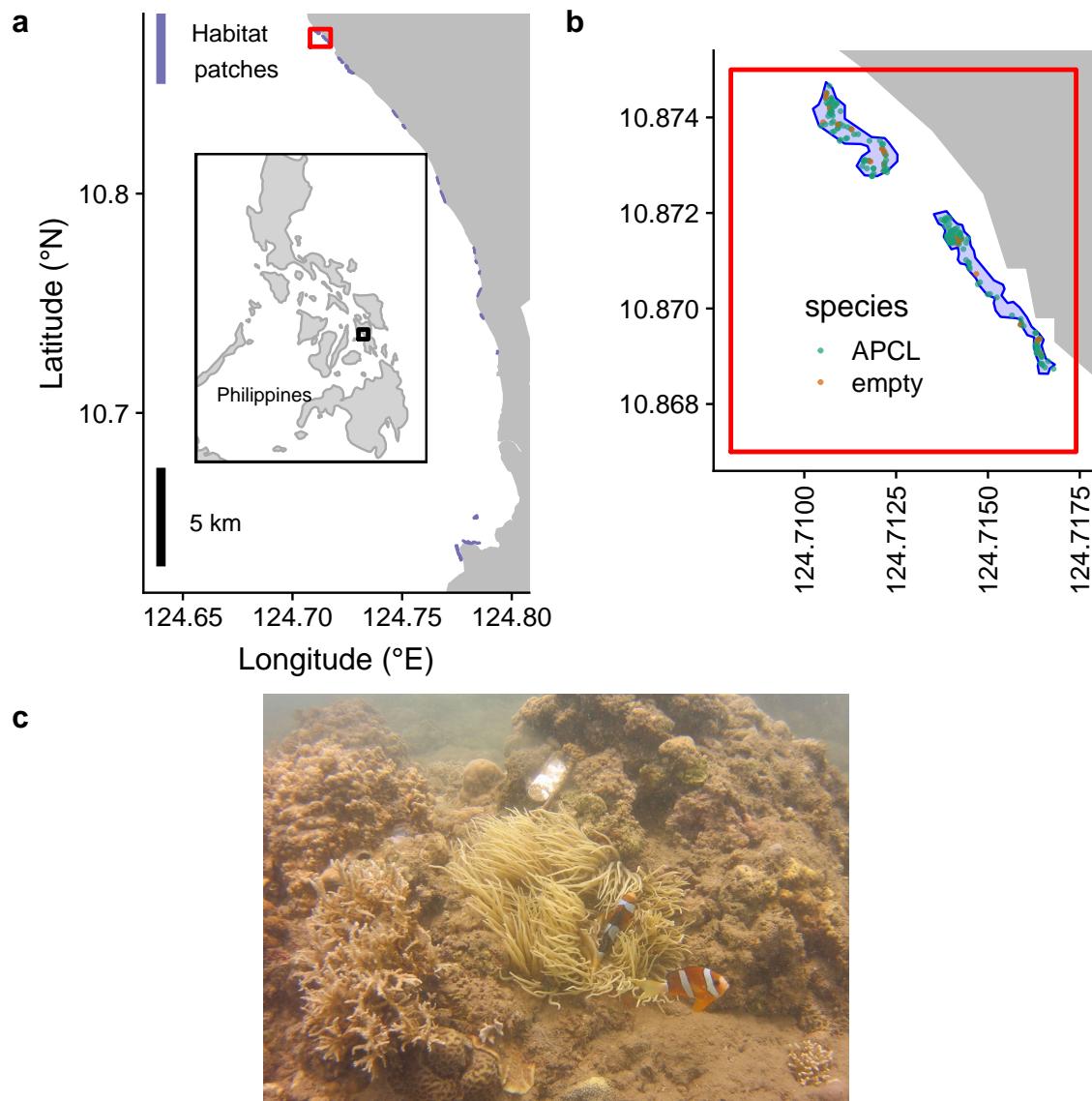


Figure 1: a) A map of the sites along the coast of Leyte in the Philippines. b) A zoomed-in map of Palanas and Wangag to show anemone arrangement for anemones occupied by *A. clarkii* (green) or unoccupied by clownfish (green). c) An example anemone occupied by *A. clarkii* in a typical habitat at the sites.

Since 2012, members of the team have sampled fish and habitat at most of the  
114 sites annually. During sampling, divers using SCUBA and tethered to GPS readers  
swim the extent of each site. Divers visit each anemone inhabited by yellowtail  
clownfish, tagging the anemone to be able to track anemones through time. At each  
117 anemone, the divers attempt to catch all of the yellowtail clownfish 3.5 cm and larger,  
taking a non-lethal tail fin-clip from each for use in genetic analysis, measuring the  
fork length, and noting the tail color (as an indicator of life stage). Starting in  
120 the 2015 field season, fish 6.0 cm and larger are tagged with a passive integrated  
transponder (PIT) tag, unless already tagged. Divers also look for eggs around each  
anemone and measure and photograph any clutches found. In total, we took fin clips  
123 from XX fish and PIT-tagged XX fish across all years and sites combined, with an  
average of XX fish clipped and XX fish tagged per year.

### **Genotyping and parentage analysis**

126 [Add in brief overview of genetic methods, with citations to papers with relevant  
methods and to Katrina's connectivity paper. Include number of fish genotyped.]

## Estimating inputs from empirical data

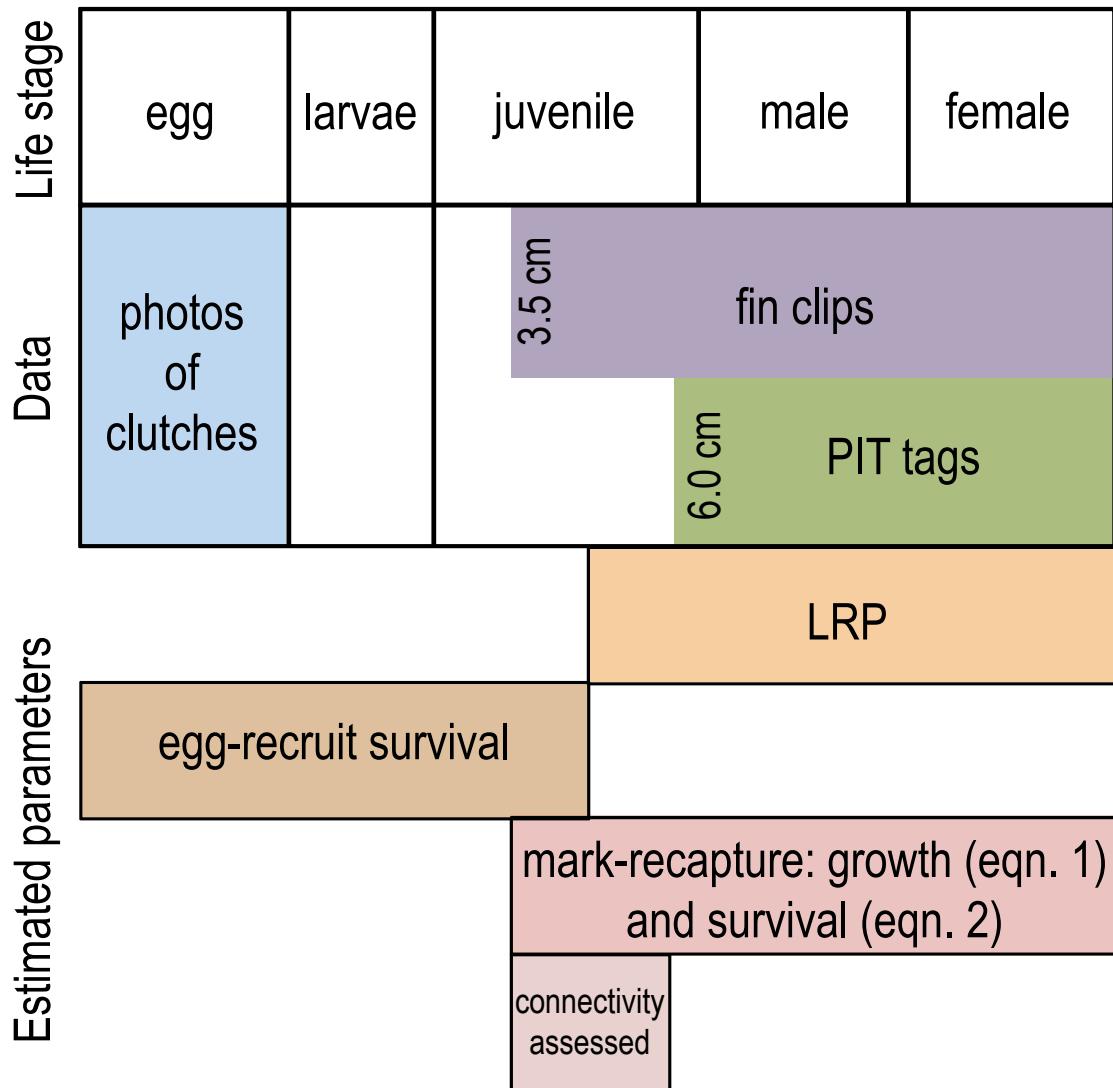


Figure 2: Here, we show the data collected for fish at each life stage (life stage boxes are not scaled by length of stage) and how the empirical data fit into the metric calculations.

<sup>129</sup> **Growth and survival: mark-recapture analyses**

We mark fish through both genetic samples and PIT tags, allowing us to estimate growth and survival through mark-recapture. After matching up recaptures of the <sup>132</sup> same fish identified by genotype or tag, we have a set of encounters of XX marked fish that includes size and stage at each capture time.

For growth, we estimate the parameters of a von Bertalanffy growth curve (Fabens, <sup>135</sup> 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:

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

<sup>138</sup> We see from eqn. 1 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 can estimate the von Bertalanffy parameters, such <sup>141</sup> that  $K = -\ln m$  and  $L_\infty = \frac{b}{(1-m)}$ . We use 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 multiple times so we randomly select only <sup>144</sup> one pair of recaptures from each to use in estimating the parameters, then repeat this process 1000 times to generate a distribution (Fig. 4b, B.1d).

We use the full set of fish encountered multiple times to estimate annual survival <sup>147</sup>  $\phi$  and probability of recapture  $p_r$  using the mark-recapture program MARK

implemented in R (Laake, 2013). We consider several models with year, size, and site effects on the probability of survival and year and size effects on the probability of recapturing a fish on a log-odds scale (see full list in Table A1). For fish that are not recaptured in particular year, we estimate their size using our growth model (eqn. 1) and the size recorded or estimated in the previous year. Because fish are not well-mixed at our sites and instead stay quite close to their home anemones, we need to swim near an anemone to have a reasonable chance of capturing the fish on it. Therefore, we also consider a distance effect on recapture probability; we use the GPS tracks of divers to estimate the minimum distance between a diver and the anemone for each tagged fish in each sample year and include it as a factor in some of the models.

The best-fit model using model selection with AICc has an effect  $b_a$  of fish size on survival, and additive effects  $b_1$  and  $b_2$  of fish size and shortest distance to anemone on the probability of recapture:

$$\log\left(\frac{\phi}{1-\phi}\right) = b_\phi + b_a \text{size} \quad (2)$$

$$\log\left(\frac{p_r}{1-p_r}\right) = b_{pr} + b_1 \text{size} + b_2 d. \quad (3)$$

## 162 Fecundity

We use a size-dependent fecundity relationship, determined using photos of egg clutches and females (Yawdoszyn, in prep), where the number of eggs per clutch (165  $E_c$ ) 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:

$$\ln(E_c) = \beta_l \ln(L) + \beta_e[\text{eyed}] + b. \quad (4)$$

<sup>168</sup> To get total annual fecundity  $f$ , we multiply the number of eggs per clutch by the number of clutches per year  $c_e = 11.9$ , using the estimate from Holtswarth et al. (2017).

<sup>171</sup> We only consider reproductive effort once the fish has reached the female stage. Though the size at which a fish transitions to become a breeding female  $L_f$  will depend on the size hierarchy in each particular colony [CITATION], we use the <sup>174</sup> average size recaptured fish were first observed as female.

## Lifetime egg production

We use an integral projection model (IPM) (e.g. Rees et al., 2014) to estimate the <sup>177</sup> total number of eggs produced by one individual (lifetime egg production: LEP), starting at the recruit stage, when individuals have settled and survived to a size we can sample.

<sup>180</sup> In an IPM, the state of the population at time  $t$  is described by the distribution of the population over a continuous trait  $z$ , for which we use size:  $n(z, t)$ . The total number of individuals in the population at time  $t$  is the integral of the size distribution <sup>183</sup> over size from the lower size bound  $L$  to the upper size bound  $U$ :  $\int_L^U n(z, t) dz$ . The population is projected forward with probability density functions, called the kernel, that describe the survival, growth, and reproductive output of existing individuals

186 into the next time step.

We initialize the IPM with one recruit-sized individual ( $\text{size}_{\text{recruit}}$ ):  $n(t = 0) = n(\text{size}_{\text{recruit}}, 0)$ , then use a kernel with the size-dependent survival and growth functions described above to project forward for 100 time steps. This gives us the size distribution at each time step, which represents the probability that the individual has survived and grown into each of the possible size categories. 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.

195 We then multiply each size-distribution vector  $v_z$  in the matrix by the size-dependent fecundity function described above (eqn. 4) to get the total number of eggs produced at each time step. To get the total number of eggs one individual is likely to produce in its lifetime, we then sum across all time steps in the individual's potential life.

$$\text{LEP} = \sum_{t=0}^{t=100} \sum_{z=L}^{z=U} v_z, t f_z. \quad (5)$$

## Survival from egg to recruit

201 We estimate survival  $S_e$  from egg to recruit using the number of recruited offspring we can match back to genotyped parents as surviving individuals from genetically "tagged" eggs in a method similar to that in Johnson et al. (2018). We estimate 204 the number of eggs produced by genotyped parent fish by multiplying the number of genotyped parents ( $N_g = 913$ ) by the expected lifetime egg production of a parent

fish  $LEP_p$ , using LEP calculated starting with an individual of 6 cm. We make the  
 207 assumption that all recruited offspring originating from the genotyped parents end  
 up in one of the sites we sample and estimate the total number of offspring that  
 survive to recruit  $R_t$  by dividing the number of offspring matches we find ( $R_m = 90$ )  
 210 by the proportion of our site habitat we sample cumulatively across all sampling  
 years ( $P_h = 0.34$ ) and the probability of capturing a fish if we sample an anemone  
 $P_c$  (see A.1, A.2 for details on  $P_h$  and  $P_c$  estimates, respectively). Our estimated  
 213 survival from egg to recruit is the number of tagged recruits divided by the number  
 of tagged eggs produced:

$$S_e = \frac{\frac{R_m}{P_h P_c}}{N_g LEP_p}. \quad (6)$$

### Defining recruit and census stage

216 When assessing persistence, it is important to consider mortality and reproduction  
 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  
 219 define a recruit to be a juvenile individual that has settled on the reef within the  
 previous year; lifetime egg production assesses how many offspring an individual  
 recruit is likely to produce in its lifetime from that point forward and egg-recruit  
 222 survival gives us the fraction of those eggs that will survive to reach the recruit  
 stage. In theory, it should not matter exactly how we define recruit so long we use  
 that definition in our calculations of both egg-recruit survival and LEP. In our system

<sup>225</sup> it is straightforward to calculate LEP from any point but it is not possible to change our estimate of egg-recruit survival to allow different definitions of recruit: we do not have enough tagged recruits to reliably estimate survival to different recruit sizes.  
<sup>228</sup> Instead, we choose the mean size of offspring matched in the parentage study as our best estimate of the size of a recruit ( $\text{size}_{\text{recruit}}$ ) and test sensitivity to different sizes within the range of sizes that the recruit stage covers (Table 1).

<sup>231</sup> **Probability of dispersal**

We use a distance-based dispersal kernel, estimated in other work using parent-offspring matches from our genetic data (Catalano et al., in prep) using the method described in Bode et al. (2018). The relative dispersal is a function of distance  $d$  as measured in kilometers and parameters  $\theta$  and  $k_d$ , which control the shape and scale of the kernel:

$$p(d) = e^k e^{-(e^k d)^\theta}. \quad (7)$$

<sup>237</sup> We use a Laplacian dispersal kernel with shape parameters  $\theta = 1$  and scale parameter  $k_d = -1.84$  (Fig. 4a, estimated in (Catalano et al., in prep)).

The dispersal kernel is estimated using fish that have already recruited to a population and survived to be sampled so it gives the relative amount of dispersal given that a fish recruits somewhere, not the probability that a released larva will travel a particular distance. To find the probability of fish dispersing among our sites, we calculate the distance between the middle of each site to the closest and farthest edge of each other site, then use the distances as upper and lower bounds

when integrating eqn. 7, which we do numerically. For example, the probability of  
246 dispersal from site A to B, where  $d_1$  is the distance from the middle of A to the  
closest edge of B and  $d_2$  is the distance from the middle of A to the far edge of B, is:

$$p_{A,B}(d) = \int_{d_1}^{d_2} e^k e^{-(e^k d)^\theta} dd. \quad (8)$$

## Persistence metrics

249 For a metapopulation to persist, at least one patch needs to achieve replacement,  
where the number of individuals entering the population balances those lost to mor-  
tality or emigration (Burgess et al., 2014; Hastings and Botsford, 2006). In our  
252 focal system, adults do not move among patches so we do not need to consider emmi-  
gration and only need to assess whether fish produce enough offspring that survive to  
recruitment to be able to replace themselves and where those offspring travel within  
255 the metapopulation. We consider three primary metrics to assess whether and how  
the population is persistent: 1) lifetime production of recruits, to assess whether the  
population has enough surviving offspring to achieve replacement 2) self-persistence,  
258 to assess whether any individual patches would be able to persist in isolation with-  
out any input from other patches, and 3) network persistence, to assess whether the  
metapopulation is persistent as a connected unit. We explain each metric below in  
261 detail.

## Estimated abundance over time

[Add brief section here.]

<sup>264</sup> **Lifetime production of recruits**

To assess whether individuals at our focal patches produce enough offspring that survive to become recruits themselves, we find the estimated number of recruits an individual recruit will produce over its lifetime (lifetime recruit production: LRP) by multiplying LEP by the estimated survival from egg to recruit  $S_e$ :

$$\text{LRP} = \text{LEP} * S_e. \quad (9)$$

If  $\text{LRP} \geq 1$ , the population has the possibility for replacement; individuals produce enough surviving offspring, before taking into account the probability of dispersal. If  $\text{LRP} < 1$ , the individuals are not replacing themselves and the population cannot persist without input from outside patches.

<sup>273</sup> **Self-persistence**

A patch is able to persist in isolation (self-persistent) if individuals produce enough offspring (LEP) that disperse back to the natal patch and survive to recruitment to be able to replace themselves (LR):  $\text{LEP} * \text{LR} \geq 1$  (Burgess et al., 2014). Our dispersal kernel represents the probability that a recruit disperses a distance given that it recruits somewhere, rather than the probability of a larva dispersing and recruiting to a particular patch, which implicitly encompasses mortality from egg to recruitment. We modify the equation to fit our data and include survival from egg

to recruit to assess whether a particular patch  $i$  is self-persistent:

$$SP_i = LEP \times \frac{\text{recruits}}{\text{egg}} \times \frac{p_{i,i} \times \# \text{ recruits from patch } i}{\frac{\text{recruits}}{\text{egg}} \times \# \text{ eggs produced by patch } i} \quad (10)$$
$$SP_i = LEP \times S_e \times p_{i,i}.$$

<sup>282</sup> A patch is self-persistent if  $SP \geq 1$ . 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).

<sup>285</sup> **Realized connectivity matrix and network persistence**

We find the probabilities of a recruit dispersing between each set of sites ( $p_{i,j}$ ) by integrating the dispersal kernel (eqn. 7) over the distance between each set of sites.

<sup>288</sup> We then create a realized connectivity matrix  $C$  by multiplying the dispersal probabilities by the expected number of recruits an individual produces:  $C_{i,j} = LRP \times p_{i,j}$  (Burgess et al., 2014). The diagonal entries of  $C$ , where the origin and destination

<sup>291</sup> are the same sites, are the values of self-persistence we calculate above.

Network persistence requires that the largest real eigenvalue of the realized connectivity matrix  $\lambda_C$  be greater than 1:  $NP = \lambda_C > 1$  (e.g. Hastings and Botsford, 2006; White et al., 2010; Burgess et al., 2014).

### Incorporating uncertainty

To represent the uncertainty in our estimates of the parameters that go into calculating our persistence metrics, we calculate each metric 1000 times, pulling each parameter from a distribution or range. In our results, we show the range of values of

each persistence metric as well as the value with our best estimate of each parameter.

300 For the dispersal kernel, we keep the shape parameter  $\theta$  constant and pull the scale  
parameter  $k_d$  from a set capturing the 95% confidence intervals, which was produced  
during kernel estimation in Catalano et al. (in prep). To capture uncertainty in the  
303 size of a recruit size<sub>recruit</sub>, and therefore the transition of mortality being captured by  
egg-recruit survival to being captured by LEP, we pull from a uniform distribution  
over the range of fish sizes (3.5 - 6.0 cm) considered as offspring in the parentage  
306 analyses (Catalano et al., in prep). We include uncertainty in the size of transition to  
a breeding female  $L_F$  by pulling from the set of sizes observed in the data for fish at  
their first recapture as a female. For the von Bertalanffy growth parameters  $L_\infty$  and  
309  $K$ , we pull from the full set of estimates using different combinations of recapture  
pairs for fish recaptured more than twice. For uncertainty in adult survival, we  
pull from a normal distribution generated using the uncertainty estimated in the  
312 mark-recapture analysis for both the intercept  $b_\phi$  and the size effect  $b_a$ .

To incorporate uncertainty in egg-recruit survival, we consider uncertainty in  
both the number of offspring assigned to parents  $R_m$  during the parentage analysis  
315 and the probability of capturing a fish  $P_c$ , which affects how the captured assigned  
offspring are scaled up to account for fish uncaught. For the number of assigned  
offspring, we generate a set of values of number of assigned offspring using a random  
318 binomial, where the number of trials is the number of genotyped offspring (XX)  
and the probability of success on each trial is the assignment rate XX of offspring  
from the parentage analysis (Catalano et al., in prep). To represent uncertainty  
321 in the probability of capturing a fish, we pull values from a beta distribution with

parameters  $\alpha_{P_c}$  and  $\beta_{P_c}$ , found using the mean and variance of capture probabilities estimated from recapture dives across sites and sampling seasons (details in A.2).

<sup>324</sup> **Results**

Our estimated abundance of females at each site over time is relatively constant [*add some sort of actual analysis here*] (Fig. 3), suggesting that our sample populations  
<sup>327</sup> are stable over time.

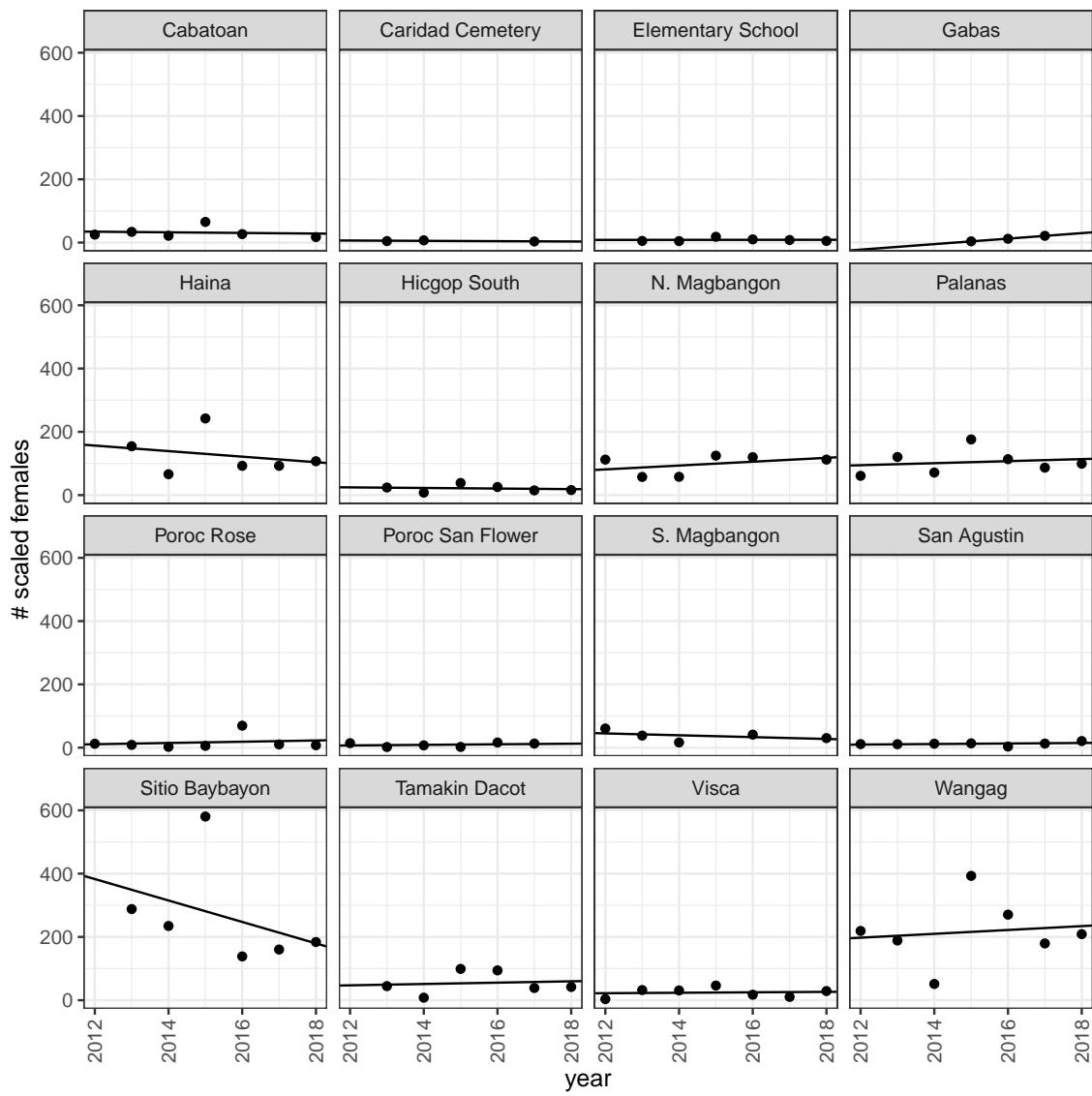


Figure 3: The estimated number of females at each site over the sampling years. The total number of females at each site was estimated by taking the number of females (fish  $> 5$  cm with the yellow pointed tail indicating female) captured at each site in each year and scaling up by the proportion of habitat sampled at that site that season (see A.1 for details) and by the average probability of capturing a fish (see A.2).

From the mark-recapture analysis of tagged and genotyped fish, we estimate mean values of  $L_\infty = 10.58\text{cm}$  (range of estimates 10.39 - 10.75 cm) and  $K = 0.928$  (range of estimates 0.854 - 1.025) for the von Bertalanffy growth curve parameters (Fig. 4b, Table 1). For juvenile and adult (post-recruitment) survival on a log-odds scale, the best-fit model has a coefficient  $b_a = 0.74 \pm 0.060$  SE for the effect of size and an intercept  $b_\phi = -4.83 \pm 0.340$  SE. These results suggest that larger fish have higher annual survival, which is similar to survival estimates in other clownfish species (check Buston paper). The accompanying best-fit model for log-odds recapture probability has intercept  $b_{pr} = 17.93 \pm 0.858$  SE, size effect  $b_1 = -1.816 \pm 0.080$  SE, and effect of diver distance from the anemone  $b_2 = -0.171 \pm 0.021$  SE. The negative effect of both size and distance suggest that divers are less likely to recapture larger fish and those at anemones far from areas sampled, with the chance of recapturing an average-sized fish falling below 5% if a diver stays farther than XX from its home anemone [add the recapture probability plots, like the survival one in Fig. 4, to the appendix and reference here.]

We set the transition size to breeding female  $L_f$  at 9.32 cm, the mean size of first female capture of recaptured fish (Fig. 4d). [*Contextualize these values.*]

[*Not sure where to put this table - kind of a methods/results hybrid, or if it should exist, but seems like it might be helpful. Need to clarify somewhere what kind of distributions are going into the uncertainty runs (drawn from data, uniform across a range, 95% confidence bounds, etc.)*]

---

Parameter	Description	Best estimate	Range in uncertainty runs	Notes
$k_d$	scale parameter in dispersal kernel	-1.36	-2.03 to -0.96	estimated using methods in Bode et al. (2018) in Catalano et al. (in prep)
$\theta$	shape parameter in dispersal kernel	0.5	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	10.58 cm	10.39 to 10.75 cm	
$K$	growth coefficient in von Bertalanffy growth curve	0.928	0.854 to 1.025	

$b_\phi$	intercept for adult survival	-4.83	$\pm 0.340$ standard error	
$b_a$	size effect for adult survival	0.74	$\pm 0.060$ standard error	
$b_{p_r}$	intercept for recapture probability from mark-recapture analysis	17.93	$\pm 0.858$ standard error	not used in persistence estimates
$b_1$	size effect for re-capture	-1.816	$\pm 0.080$ standard error	not used in persistence estimates
$b_2$	distance effect for recapture	-0.171	$\pm 0.021$ standard error	not used in persistence estimates
$\text{size}_{\text{recruit}}$	size (cm) of recruited offspring	mean of size of offspring in parentage analysis = 4.4 cm	3.5 - 6.0 cm	
$b_e$	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)
$L_f$	size at transition to female	9.32cm	5.2 - 12.7cm	
$P_c$	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 A.2

Table 1:

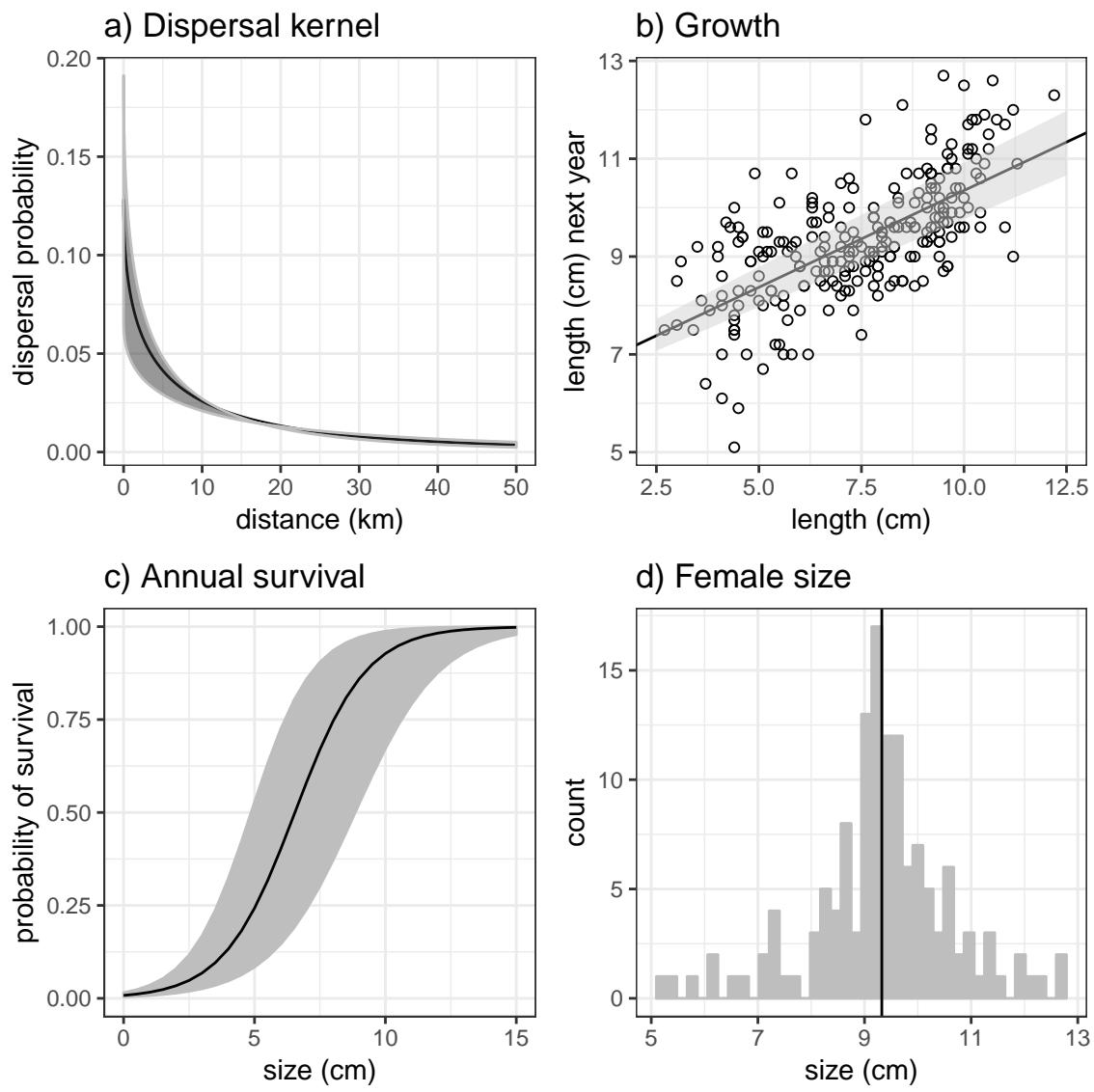


Figure 4: Best estimates (solid black line) and range included for uncertainty (gray) for dispersal (a), growth (b), post-recruit survival (c), and size at female transition (d) parameters.

Using our best estimates for growth, survival, and fecundity, we calculate a value

of LEP for 10876, ranging from XX to XX when we consider uncertainty in the inputs  
351 (Fig. 5a). The size at recruitment - the census point between egg-recruit survival  
and LEP - has the most effect on the value of LEP (Fig. B.3), with higher values of  
LEP the higher the size of recruitment as less mortality is included before reaching  
354 reproductive sizes.

We estimate egg-recruit survival  $S_e$  to be 1.82e-05, ranging from XX to XX  
when we include uncertainty in the number of offspring assigned to parents and  
357 the probability of catching a fish (Fig. 5b). Uncertainty in the size of transition  
to breeding female  $L_f$  has the largest effect on egg-recruit survival (Fig. B.5); we  
only consider reproduction from females, to avoid double-counting, so the larger  
360 the transition size to female, the fewer tagged eggs we estimate were produced by  
genotyped parents and the higher egg-recruit survival.

We estimate lifetime recruit production, the product of LEP and  $S_e$ , to be 0.20,  
363 below the value of 1 necessary for replacement. This suggests that even without  
considering connectivity, the individuals at our sample populations do not produce  
enough offspring that survive to recruitment to replace themselves. When we consider  
366 uncertainty in our parameter estimates, we do see a few cases where  $LRP > 1$ , but  
the majority are well below the threshold for replacement.

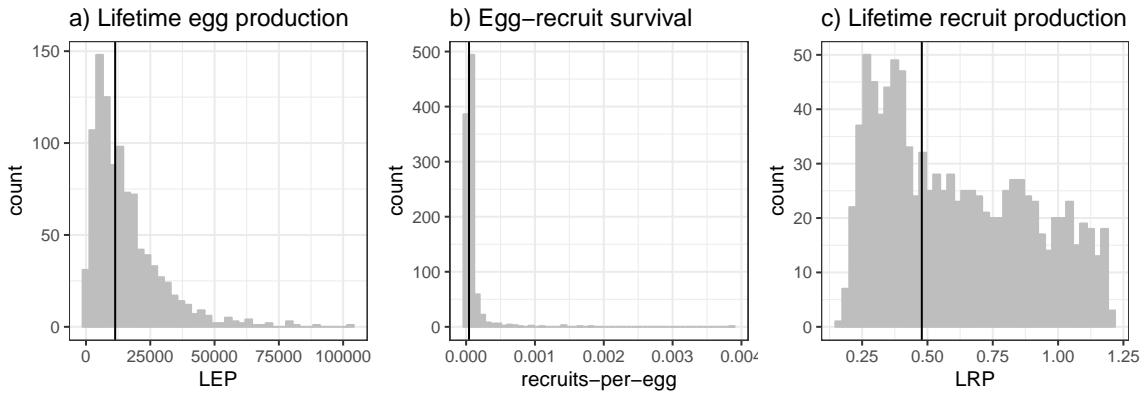


Figure 5: Estimates of a) LEP, b) egg-recruit survival, and c) LRP, showing the best estimate (black solid line) and range of estimates considering uncertainty in the inputs.

We do not find any sites with self-persistence values  $> 1$ , indicating that the site could persist in isolation. Given that our estimate of LRP does not suggest replacement and only a fraction of that recruitment stays at the natal site, this makes sense. We see the highest values of self-persistence at Haina (SP = 0.024) and Wangag (SP = 0.010), our two widest sites.

Self-persistence by site

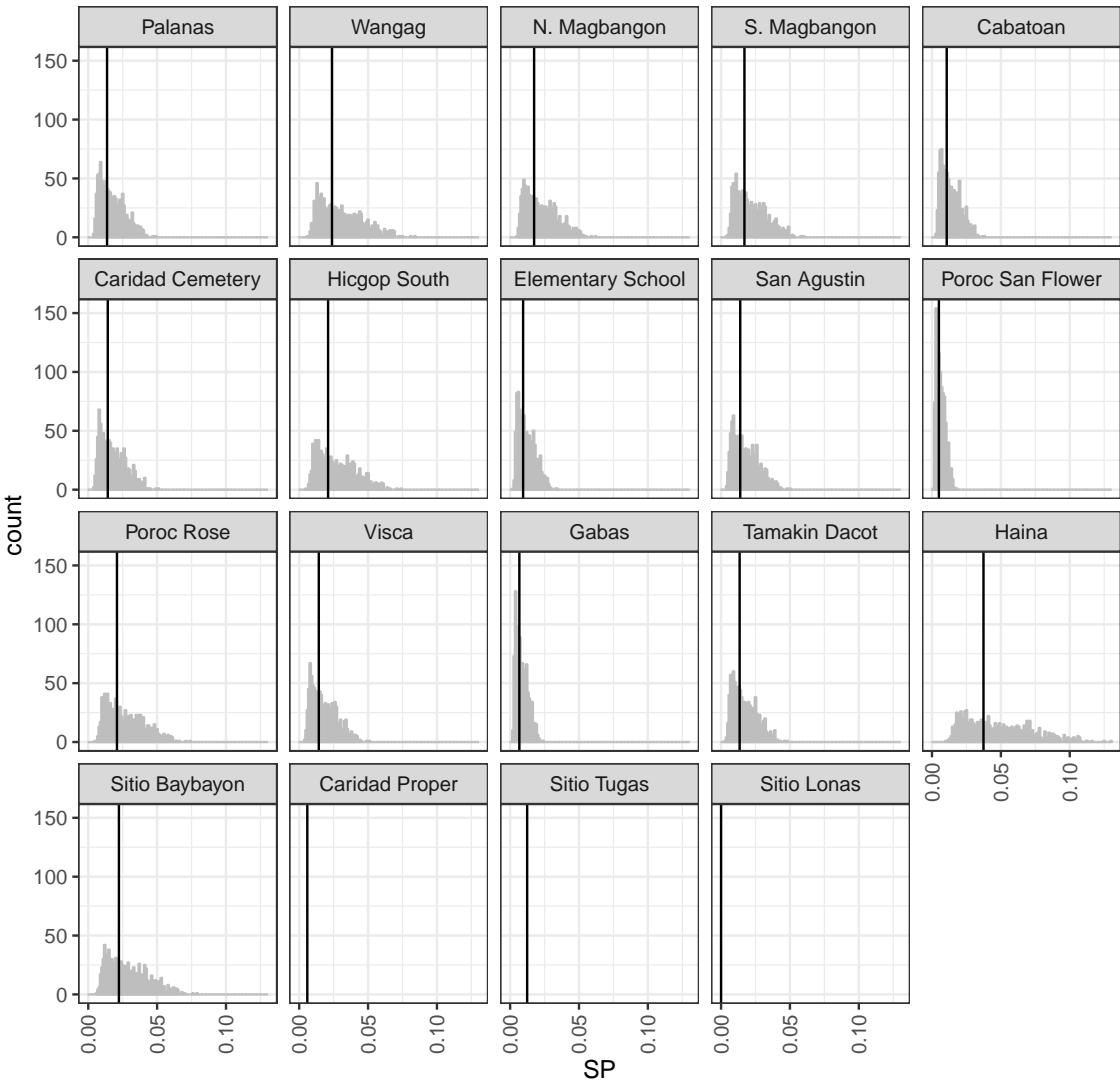


Figure 6: Values of self-persistence at each site, showing the best estimate (black line) and range of estimates considering uncertainty in the input parameters. No site reaches a value of  $SP > 1$ , necessary to be self-persistent.

We also do not find evidence of network persistence; the dominant eigenvalue of

the realized connectivity matrix  $\lambda_c$  is 0.034, well below the value of 1 that indicates network persistence (Fig. 7a). We see that most of the connectivity occurs among the sites in the northern part of our sample area, from Palanas to Caridad Cemetery.

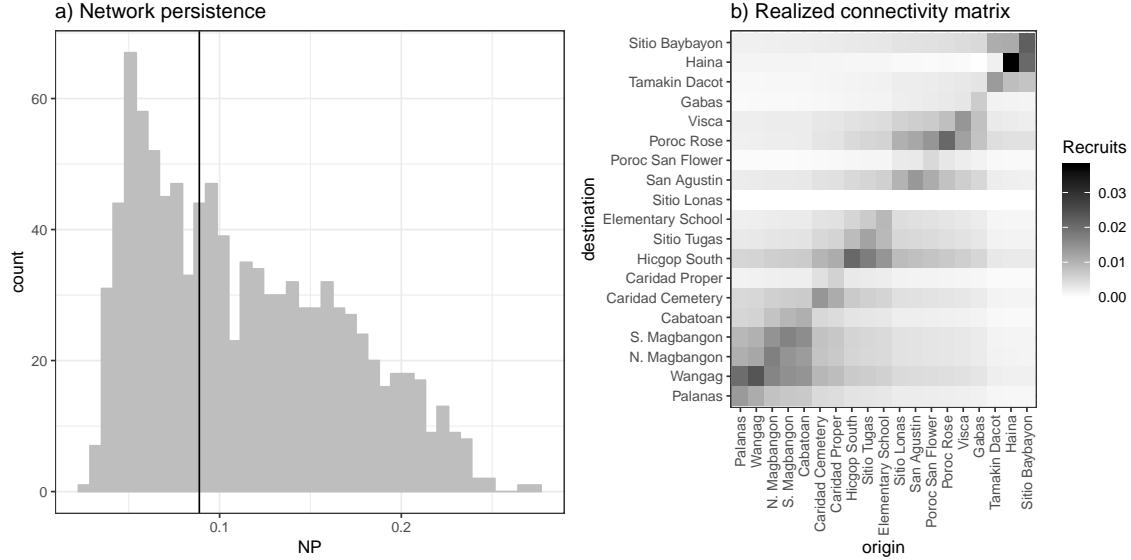


Figure 7: a) Network persistence values, showing the best estimate (black solid line) and range of estimates considering uncertainty. b) The realized connectivity matrix  $C$ , with sites arranged from north (Palanas) to south (Sitio Baybayon).

Based on our estimates of LRP, SP, and NP, we do not expect that our set of sites  
 378 is able to persist in isolation as a closed system. To explore what would be required  
 for persistence, we consider a hypothetical scenario in which we consider the system  
 closed and assume that all of the recruits arriving at our sites came from adults at  
 381 our sites. In this case, we find a value of  $LRP = 1.21$ , above the value of 1 necessary  
 for replacement (Fig. 8a). When we add in the connectivity, we see a higher value

of  $\lambda_c$  in our best estimate ( $NP = 0.20$ ) but still not high enough to indicate network persistence (Fig. 8b). We see more of the distribution of estimates above 1, however, suggesting that network persistence is within our range of uncertainty in this case, though not likely. With our site configuration and dispersal kernel estimate, we would need a value of LRP of XX (an egg-recruit survival of XX with our estimated value of LEP or a value of LEP of XX with our estimated value of egg-recruit survival), to  $\lambda_c = 1$  and network persistence.

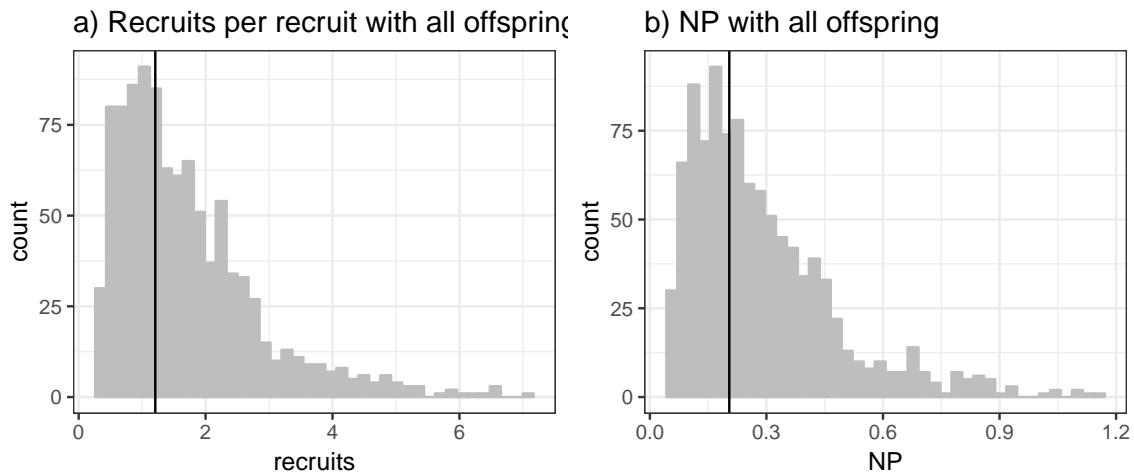


Figure 8: a) Recruits per recruit when we consider all arriving recruits to have originated from our sites. b) Range of values of NP considering all arriving recruits to be offspring from our sites, with the best estimate in a black solid line.

## 390 Discussion

Big picture: What do our results mean for persistence in this system and our understanding of metapopulations generally?

- 393     ● So we don't see persistence in our metrics, either self-persistence or network  
394       persistence but our abundances don't seem to be changing. Suggests that  
395       this is just a portion of a larger metapopulation, rather than a self-contained  
396       metapopulation. Maybe it is a sink? Persistent in terms of constant abundance  
397       but relies on outside immigration to persist.
- 399     ● How does dispersal spread (estimated to be within our sites) interact with scale  
400       of a self-contained metapopulation? How do we reconcile this in our system,  
401       where we don't estimate dispersal that far but don't see network persistence  
402       in an area range that spans the estimated spread? (This point might change,  
403       depending on mean dispersal distance from the new kernels)
- 405     ● Sensitivity - how would our parameters need to change to see persistence?  
406       Egg-recruit survival is a big one. Discuss limitations of how we calculated it  
407       (offspring going outside our pops not included - though we might change this),  
408       what we see for persistence when estimate recruits/recruits instead. Contextualize this with what other studies have found for these parameters, how  
409       reasonable it would be to get better estimates in the field.

More detailed discussion of our estimates, limitations, ways to move forward:

- 411     ● Discuss density-dependence: not explicitly accounting for it, included in our  
412       egg-recruit survival estimate. But it's these metrics at low abundance, when  
413       DD isn't happening, that really matter for persistence. Egg-recruit-survival  
414       is probably higher in that case than our estimate of it here (b/c larvae able

414 to settle without being chased off by already-settled recruits). But is it high enough?

- Discuss site-specific demographic rates, why we don't estimate them in our system, the importance they play in other studies, what we might need to go about resolving them, whether we would expect to see them.
- Contextualize our parameter estimates with those from other studies (esp. survival, growth, fecundity).

Broadening back out:

- What does this mean for moving forward in understanding metapopulation persistence more broadly? Stability in abundance doesn't mean the population would be able to persist in isolation. Area required seems to be much wider than dispersal kernel spread (particularly if LRP production is right around replacement). Even areas of habitat along a linear coastline seem to be drawing much of their recruitment from a larger surrounding area - even though we see some local retention, maybe broader connectivity is still the story in terms of receiving enough recruitment to persist.

# Appendix

## A Method details

### <sup>432</sup> A.1 Proportion of habitat sampled

[Need to add in the details here]

## A.2 Probability of capturing a fish, from recapture dives

435 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 (XX),  
portions of the sites were sampled again XX-XX weeks after 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  
change their behavior or response to divers, so therefore assume that the probability  
438 of recapturing a fish is the same as the probability of capturing a fish on a sample dive.  
For each recapture dive, we use GPS tracks of the divers to identify the anemones  
covered in the recapture dive and the set of PIT-tagged fish encountered on those  
441 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  
total number of fish caught on the recapture dive  $n_2$ :  $P_c = \frac{m_2}{n_2}$ .

447 We use the mean  $P_c$  across all 14 recapture dives, covering XX sites in 3 sampling  
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  
450 of uncertainty, we represent the probability of capture as a distribution, rather than  
pulling 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  
453 beta distribution, using the mean  $\mu_{P_c}$  and variance  $V_{P_c}$  of the set of 14 values to find  
the appropriate  $\alpha_{P_c}$  and  $\beta_{P_c}$  parameters, where

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

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

The mean of the individual capture probability values is  $\mu_{P_c} = 0.56$ , with variance  
 456  $V_{P_c} = 0.069$ , which gives beta distribution parameters  $\alpha_{P_c} = 1.44$  and  $\beta_{P_c} = 1.13$ .  
 We 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  
 459 avoid extremely low values that are sometimes sampled 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.

<sup>462</sup> **A.3 Full set of MARK models**

We consider the following set of models in MARK [*Need to add in models*]:

Model	Model description	AICc	dAICc
	survival size, recapture size+distance	3348.861	0
	survival size, recapture distance	3359.998	-11.1371
	survival constant, recapture dis- tance	3383.175	34.3141
	survival constant, recapture size+distance	3384.959	36.0981
	survival time, recapture constant	3408.342	59.4816
	survival site, recapture constant	3440.842	91.98112
	survival site, recapture size+distance	3440.842	91.98112
	survival constant, recapture time	3453.609	104.74839
	survival size, recapture size	3527.710	178.84940
	survival constant, recapture con- stant	3570.908	222.04690

Table A1:

## B Uncertainty details

### <sup>465</sup> B.1 Sensitivity to parameters

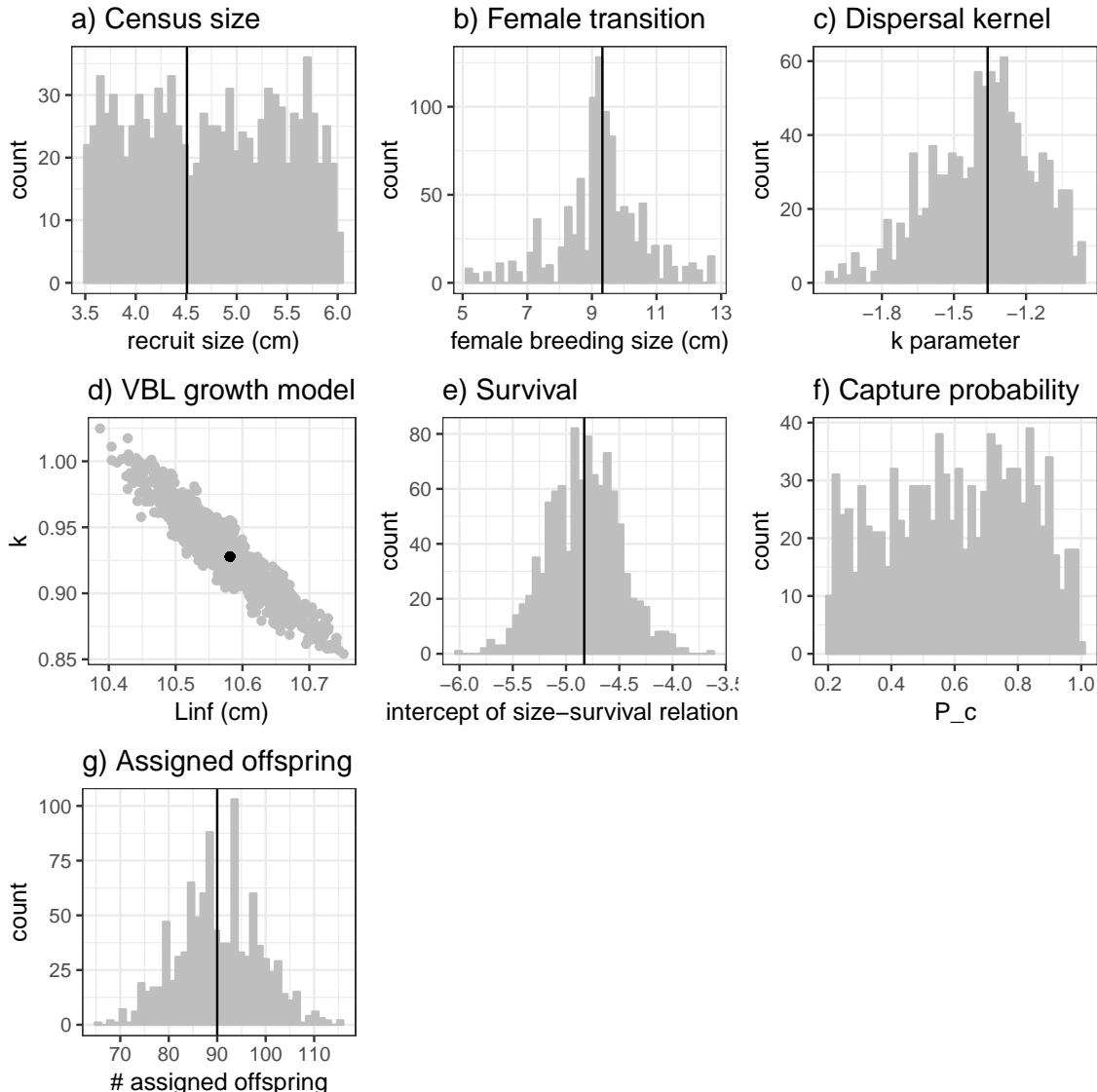


Figure B.1: Range of parameter inputs for uncertainty runs with all uncertainty included: a)  $\text{size}_{\text{recruit}}$ , the census size at which fish are considered to have recruited after egg-recruit survival occurs; b)  $L_f$ , the size at which fish transition from male to female and their reproductive output is included in the estimate of lifetime egg production (LEP); c)  $k_d$ , the scale parameter in the dispersal kernel; d) the parameters  $L_\infty$  and  $K$  of the von Bertalanffy growth model; e) the intercept  $b_\phi$  of the adult size-dependent survival relationship; f)  $P_c$ , the probability of capturing a fish;

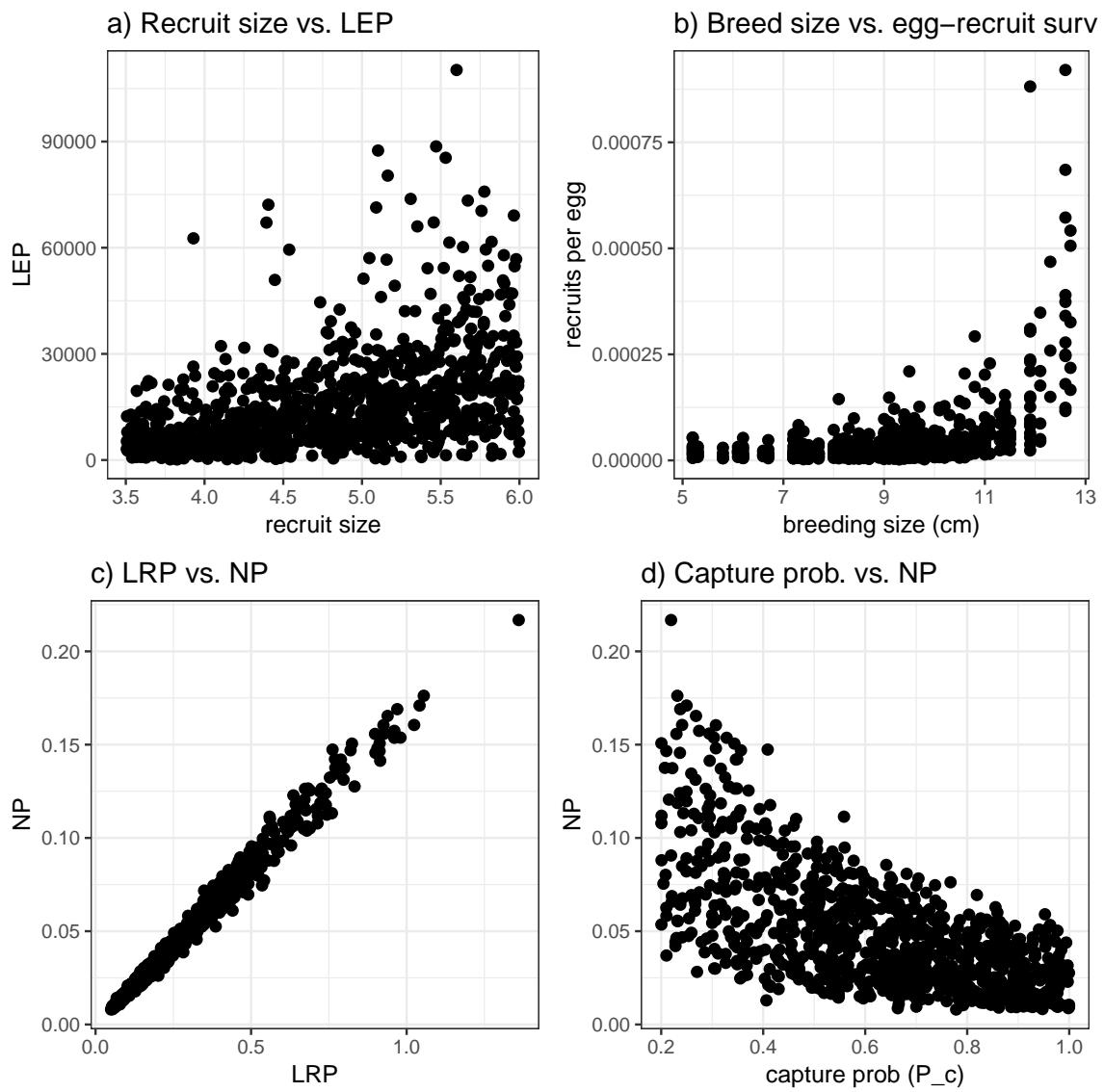


Figure B.2: Relationships among parameters and metrics.

## B.2 Effects of different types of uncertainty on metrics

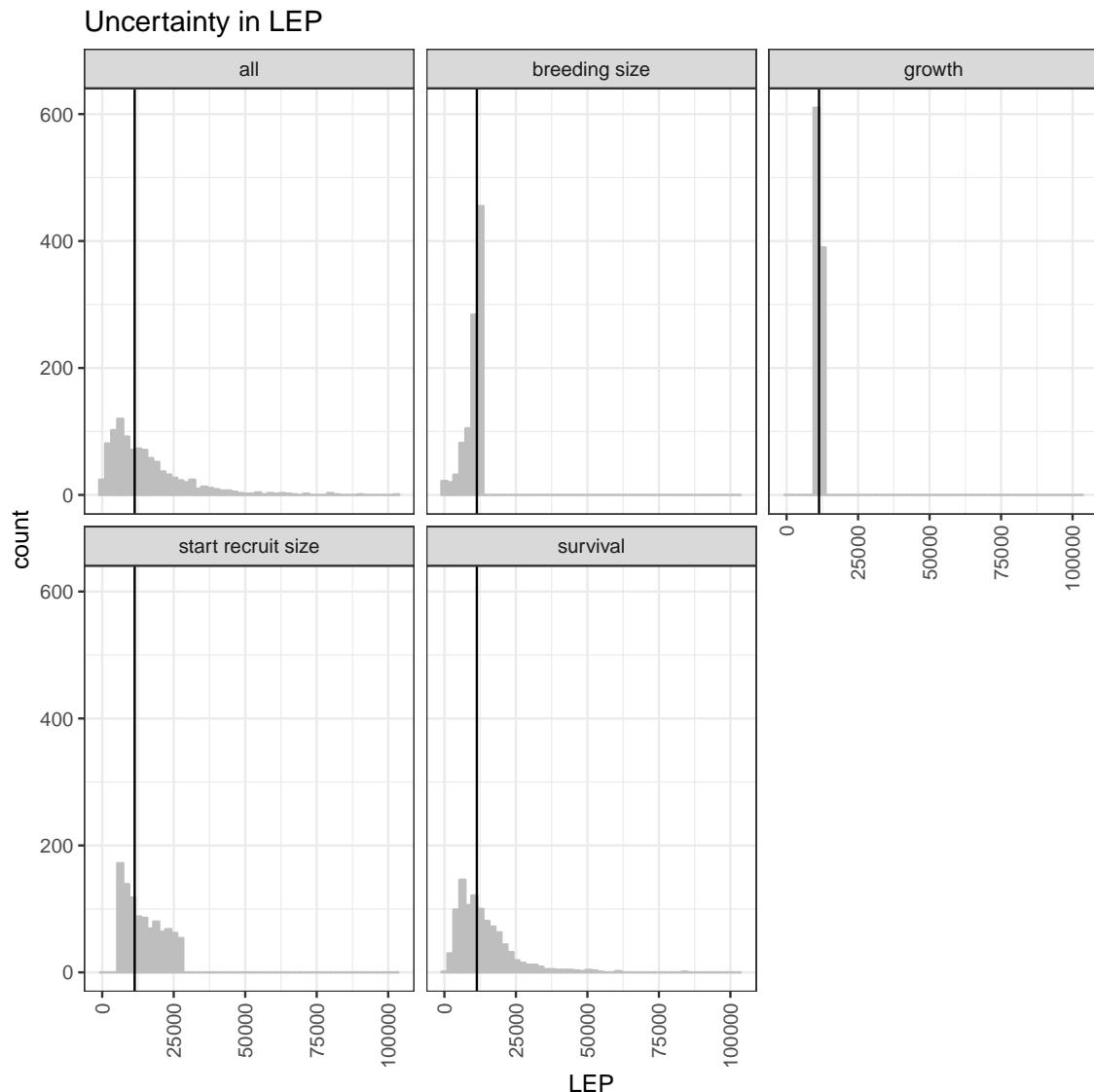


Figure B.3: The contribution of different sources of uncertainty in LEP.

### Uncertainty in LRP

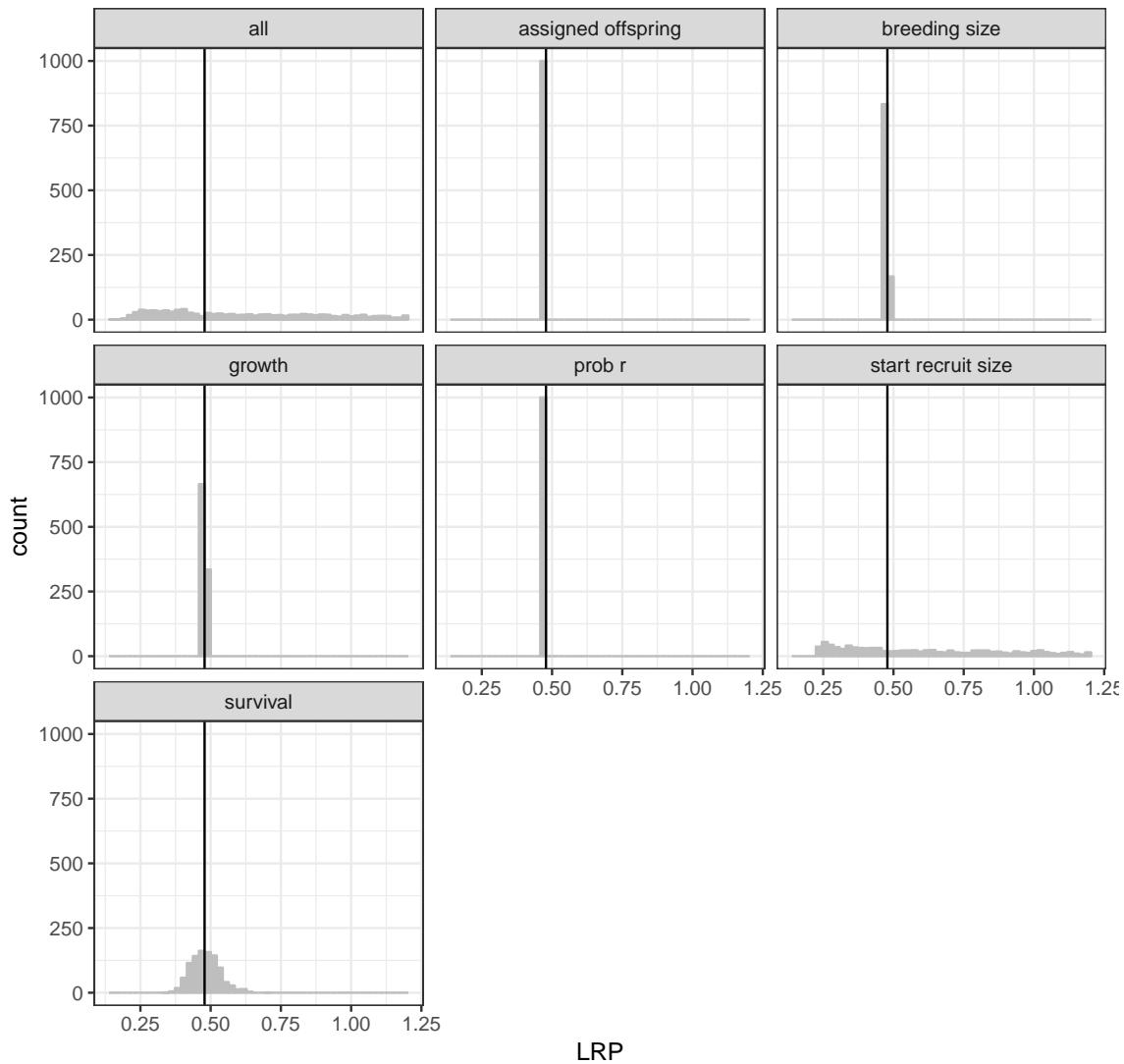


Figure B.4: The contribution of different sources of uncertainty in LRP.

### Uncertainty in egg–recruit survival

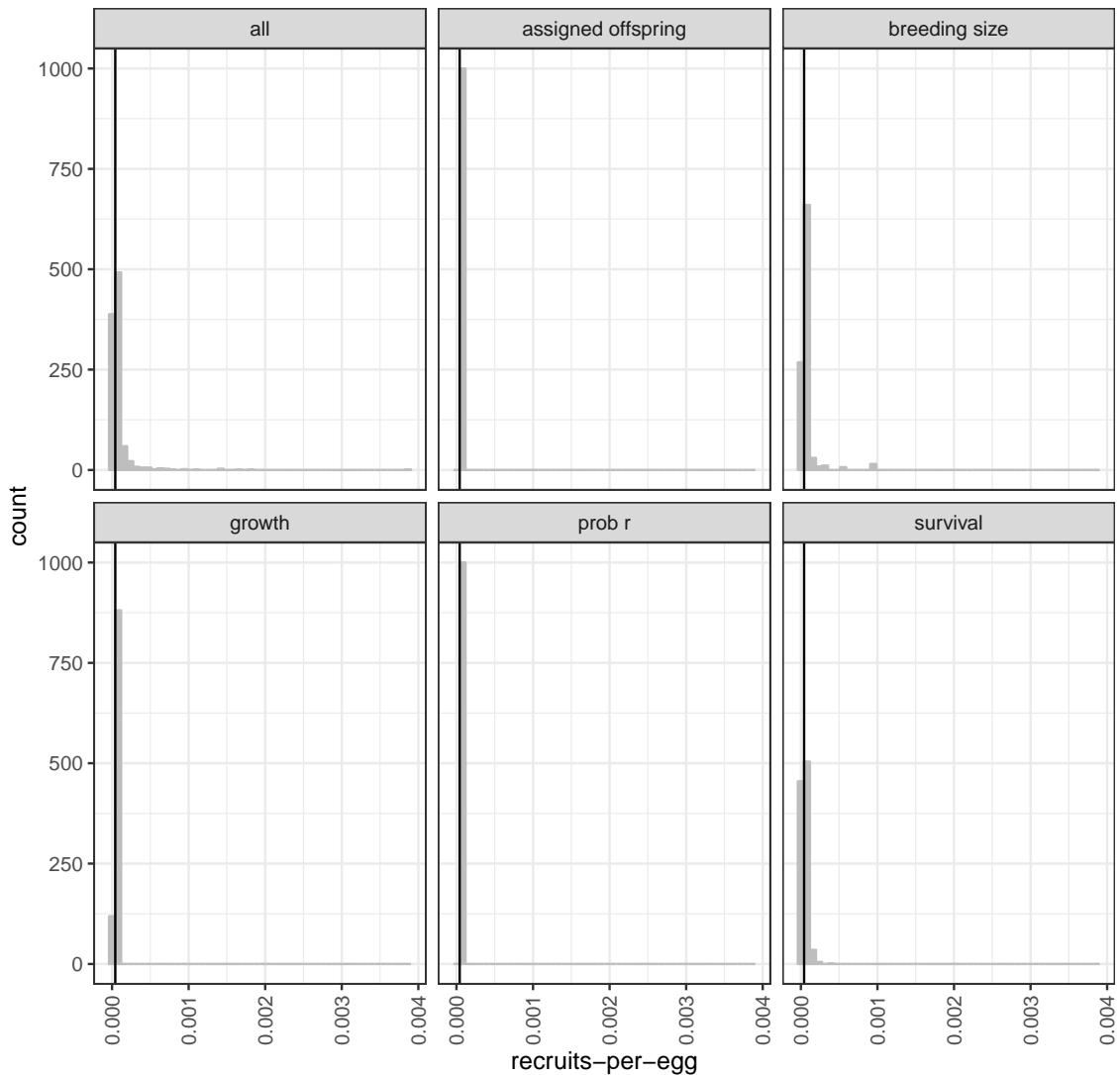


Figure B.5: The contribution of different sources of uncertainty in egg-recruit survival.

### Uncertainty in network persistence

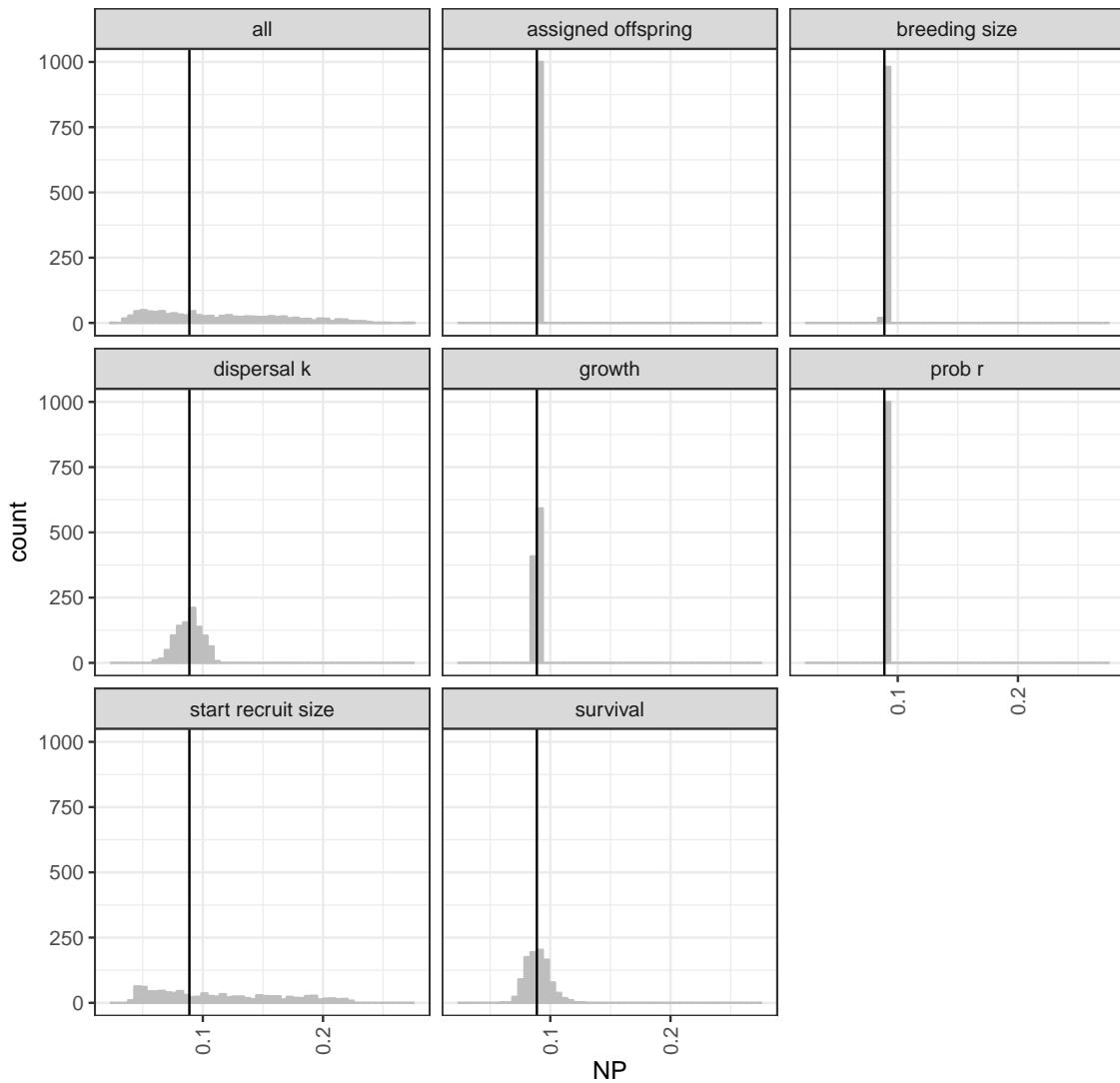


Figure B.6: The contribution of different sources of uncertainty in NP.

## References

- 468 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.
- 471 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.
- 474 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.
- 477 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  
quantify population persistence to improve marine protected-area design. *Ecolog-  
ical Applications*, 24(2):257–270, 2014.
- 480 Peter Buston. Social hierarchies: size and growth modification in clownfish. *Nature*,  
424(6945):145–146, 2003.
- 483 Peter M Buston and Cassidy C DAloia. Marine ecology: reaping the benefits of local  
dispersal. *Current Biology*, 23(9):R351–R353, 2013.
- 486

Hal Caswell. *Matrix population models: construction, analysis, and interpretation*.  
Sinauer Associates Inc., Sunderland, Massachusetts, 2nd edition, 2001.

<sup>489</sup> Katrina A Catalano, Allison G Dedrick, Michelle Stuart, Jonathan Purtiz, Humberto Montes, Jr., and Malin Pinsky. Interannual variability of genetic connectivity in a coral reef fish *Amphiprion clarkii*. in prep.

<sup>492</sup> Robert K. Cowen and Su Sponaugle. Larval Dispersal and Marine Population Connectivity. *Annual Review of Marine Science*, 1(1):443–466, January 2009. ISSN 1941-1405, 1941-0611. doi: 10.1146/annurev.marine.010908.163757. URL <http://www.annualreviews.org/doi/abs/10.1146/annurev.marine.010908.163757>.

Augustus J. Fabens. Properties and fitting of the von bertalanffy growth curve.  
*Growth*, 29:265–289, 1965.

<sup>498</sup> Daphne Gail Fautin, Gerald R Allen, Gerald Robert Allen, Australia Naturalist, Gerald Robert Allen, and Australie Naturaliste. Field guide to anemonefishes and their host sea anemones. 1992.

<sup>501</sup> Sarah O Hameed, J Wilson White, Seth H Miller, Kerry J Nickols, and Steven G 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.

Deborah R Hart and Antonie S Chute. Estimating von bertalanffy growth parameters

507 from growth increment data using a linear mixed-effects model, with an application  
to the sea scallop *Placopecten magellanicus*. *ICES Journal of Marine Science*, 66  
(10):2165–2175, 2009.

510 Alan Hastings and Louis W. Botsford. Persistence of spatial populations depends on  
returning home. *Proceedings of the National Academy of Sciences*, 103:6067–6072,  
2006.

513 Jordan N. Holtswarth, Shem B. San Jose, Humberto R. Montes Jr., James W. Morley,  
and Malin. L Pinsky. The reproductive seasonality and fecundity of yellowtail  
clownfish (*Amphiprion clarkii*) off the Philippines. *Bulletin of Marine Science*, 93,  
516 2017.

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

Jacob P Kritzer and Peter F Sale. *Marine metapopulations*. Elsevier Academic Press,  
2006.

522 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/  
525 Publications/ProcRpt/PR2013-01.pdf](http://www.afsc.noaa.gov/Publications/ProcRpt/PR2013-01.pdf).

Steven G. Morgan, Jennifer L. Fisher, Seth H. Miller, Skyli T. McAfee, and John L.  
Largier. Nearshore larval retention in a region of strong upwelling and recruit-

- 528      ment limitation. *Ecology*, 90(12):3489–3502, December 2009. ISSN 0012-9658.  
doi: 10.1890/08-1550.1. URL <http://www.esajournals.org/doi/abs/10.1890/08-1550.1>.
- 531      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.
- 534      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.
- 537      Mark Rees, Dylan Z Childs, and Stephen P Ellner. Building integral projection  
models: a user’s guide. *Journal of Animal Ecology*, 83(3):528–545, 2014.
- 540      J Roughgarden, S Gaines, and H Possingham. Recruitment dynamics in complex  
life cycles. *Science*, 241(4872):1460–1466, September 1988. ISSN 0036-8075, 1095-  
9203. doi: 10.1126/science.11538249. URL <http://www.sciencemag.org/cgi/doi/10.1126/science.11538249>.
- 543      Ocane C. Salles, Jeffrey A. Maynard, Marc Joannides, Corentin M. Barbu, Pablo  
Saenz-Agudelo, Glenn R. Almany, Michael L. Berumen, Simon R. Thorrold, Ge-  
offrey P. Jones, and Serge Planes. Coral reef fish populations can persist without  
546      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.

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/>.

Adam Yawdoszyn. Fecundity in clownfish. in prep.