Clownfish metapopulation persistence draft

Introduction

Paragraph 1: Long been a challenge to understand how marine metapopulations - patchy, interconnected populations - are connected and how they persist. Particularly challenging in the marine environment, where much movement happens at the larval stage - high mortality, hard to tag/track, can go far b/c of currents. Of key interest, though - both just because we want to understand population dynamics, persistence, and connectivity for basic ecological reasons, and for more applied reasons, like conservation (establishment of marine protected areas), understanding where to fish.

Lots of theory has been developed about how we would expect populations largely connected by larval movement to persist, which essentially boils down to replacement. To understand persistence, we need to know 1) whether we see replacement and 2) the origins of that replacement - self or other patches? Define and distinguish self persistence vs network persistence. Hastings and Botsford 2006 papers, Bostford 2009, some of the White papers (2010?), Burgess et al. 2014. Walk through the theory, our expectations for persistence, and what we need to know. Might take more than one paragraph.

Testing/validating/demonstrating that theory empirically has been challenging but new technologies in tagging and genetics are making it more feasible. There have been some recent successes. Cite Johnson et al., Salles et al. the recent paper by someone Will works with (Garavelli or something?). Go through (briefly) what they did and found.

We continue this work with our long-ish term data set of clownfish. Like Salles (check) and Johnson (check), we use a tropical reef species blablalbah (maybe study system description goes in next paragraph). Talk about how we have longer-term data set (check that this is actually true...) so we can capture longer-term replacement, smooth out some of the year-to-year variability. Many years of sampling lets us assess pop persistence in two ways: 1) from estimated abundance trends and 2) estimating replacement and persistence metrics. Also, have a portion of the metapopulation that is continuous (though might not be quite enough of the metapopulation to capture the full network). Here, essentially, distinguish what we are doing from the other studies. Not exactly sure of the flow here - maybe 1 paragraph or maybe 2?

Overview of clownfish and why they are a great system. Here we.....

Methods

Study system

We focus on a tropical metapopulation of yellowtail clownfish (*Ampiprion clarkii* in the Philippines. Like many clownfish species, yellowtail clownfish have a mutualistic relationship with anemones, where small colonies of fish live (CITATION). Yellowtail clownfish are protandrous hermaphrodites and maintain a size-strutured 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 (CITATION). The fish on an anemone maintain a strict social and size hierarchy (Buston, 2003), with fish moving up in rank to become breeders only after the larger fish have died. In the tropical patch reef habitat of the Philippines, yellowtail clownfish spawn once per lunar month from November to May, laying clutches benthic eggs that the parents protect and tend (Ochi, 1989). Larvae hatch after about six days (check this) and spend 7-10 days as pelagic larvae 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 clearly patchy habitat. Once fish have settled, they tend to stay within close proximity of their anemones (XX meters, CITATION). This 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 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 are SOMETHING ABOUT CLOWNFISH AND OTHER SIMILAR SPECIES AS A MODEL SYSTEM FOR METAPOPULATION PERSISTENCE WORK, CITATIONS.

Field data collection

We focus on a set of seventeen patch reef sites spanning approximately 30km along the western coast of Leyte island in the Philippines (MAP FIGURE). The sites consist of rocky patches of coral reef and are separated by sand flats. Previous work using genetic isolation by distances estimated that yellowtail clowfish larvae have a dispersal spread of about 10km (Pinsky et al., 2010), so our sites were 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 20km.

TO ADD: Figure 1: map of study sites, picture of clownfish *(Any figure summarizing the data? How many fish captured, sequenced, etc? Could go in the appendix?)*

Since 2012, members of the team have sampled fish and habitat at most of the 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 anemone, the divers attempt to catch all of the yellowtail clownfish 3.5cm 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 the 2015 field season, fish 6.0cm and larger are tagged 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.

Processing genetic samples

TALK TO KATRINA AND MICHELLE, ADD IN HERE. BRIEF OVERVIEW, WITH CITATIONS TO PAPERS WITH RELEVANT METHODS AND TO KATRINA'S CONNECTIVITY PAPER.

Estimating inputs from empirical data

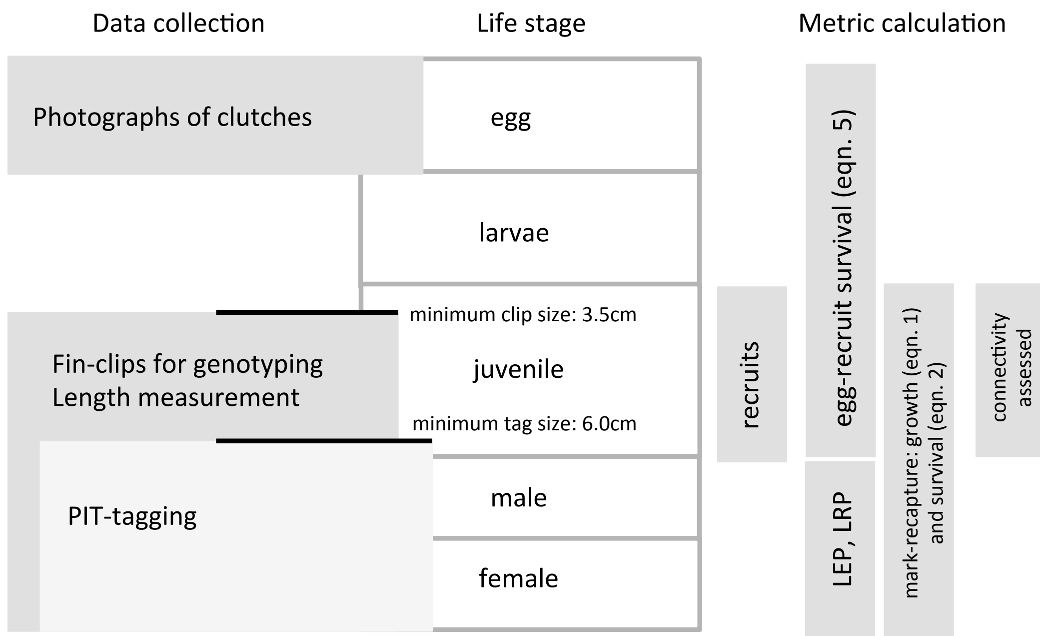

Figure 1: Here, we show the data collected for fish at each life stage (life stage boxes are not scaled by length of stage). We also show how the empirical data fit with the metric calculations.

87 ### 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 same fish identified by genotype or tag, we have a set of encounters of each recaptured fish that includes size and stage at each capture time.

For growth, we estimate 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_∞ is the average asymptotic size across the population and K controls the rate of growth:

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

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 that $K = -\ln m$ and $L_\infty = \frac{b}{(1-m)}$. We use the first and second capture lengths for fish that were recaught after a year (within 345 to 385 days) to estimate L_∞ and K. We have some fish that were recaptured multiple times so we randomly select only one pair of recaptures from each to use in estimating the parameters, and repeat this process 1000 times to generate a distribution (Fig. 2b, B.1d).

We use the full set of fish encountered multiple times to estimate annual survival ϕ 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 at a particular time point, 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 shortest distance a diver got to 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 of fish size on survival and additive effects of fish size and shortest distance to anemone on the probability of recapture:

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

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

Fecundity

We use a size-dependent fecundity relationship, determined using photos of egg clutches and females (Yawdoszyn et al. in prep), where the number of eggs per clutch (E_c) is exponentially related to the length of the female (L) with slope $beta_l$, intercept b, and effect β_e dependent on if the eggs are old enough to have visible eyes:

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

To get total annual fecundity, we multiply the number of eggs per clutch by the number of clutches per year c_e, using the estimate from Holtswarth et al. (2017).

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 (CTIATION), we use the average size recaptured fish were first observed as female for the best estimate.

Lifetime egg production

We use an integral projection model (IPM) (e.g. Rees et al., 2014) to estimate the 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.

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 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 into the next time step.

We initalize 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.

We then multiply each size-distribution vector 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 of eggs one individual is likely to produce in its lifetime, we then sum across all time steps in the individual's potential life.

$$\mathrm{LEP}z = IPM * Fec. \tag{5}$$

Survival from egg to recruit

We estimate survival from egg to recruit S_e 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 the number of eggs produced by genotyped parent fish by multiplying the number of genotyped parents N_g by the expected lifetime egg production of a parent fish LEP_p, using LEP calculated starting with an individual of 6cm. We make the 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 by the proportion of our site habitat we sample P_h and the probability of capturing a fish if we sample an anemone P_c (see B.4 for details on P_h and P_c). Our estimated survival from egg to recruit is the number of tagged recruits divided by the number of tagged eggs produced:

$$S_e = \frac{\frac{R_m}{P_h P_c}}{N_g \mathrm{LEP}_p}. \tag{6}$$

Defining recruit and census stage

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 its same life stage (CITATION?). We define recruit to be a juvenile individual that has settled on the reef 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 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 defition in our calculations of both egg-recruit survival and LEP. In our system it is straighforward 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. Instead, we choose the mean size of offspring matched in the parentage study as our best estimate of the size of a recruit ($\mathrm{size}_{\mathrm{recruit}}$) and test sensitivity to different sizes within the range of sizes that the recruit stage covers (Table 1).

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 θ and k_d, which control the shape and scale of the kernel:

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

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 larvae 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 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:

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

Persistence metrics

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 mortality or emmigration (CITATION). In our focal system, adults do not move among patches so we do not need to consider emmigration 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 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 acheive replacement 2) self-persistence, to assess whether any individual patches would be able to persist in isolation without 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 detail.

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:

$$\mathrm{LRP} = \mathrm{LEP} * S_e. \tag{9}$$

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

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): LEP $\times$ LR ≥ 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 larvae 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 whether a particular patch i is self-persistent:

$$\begin{aligned} SP_i &= \text{LEP} \times \frac{\text{recruits}}{\text{egg}} \times \frac{p_{i,i} \times \#\text{ recruits from site}}{\frac{\text{recruits}}{\text{egg}} \times \#\text{ eggs produced by patch i}} \\ SP_i &= \text{LEP} \times S_e \times p_{i,i}. \end{aligned} \tag{10}$$

A patch is self-persistent if SP ≥ 1. If at least one patch is self-persistent, the metapopulation as a whole is persistent as well (CTITATION).

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

Not sure where to put this, or if it should exist, but seems like it might be helpful:

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)

θ	shape parameter in dispersal kernel	0.5	NA	estimated using methods in Bode et al. (2018) in Catalano et al. (in prep)
$size_{recruit}$	size (cm) of recruited offspring	mean of size of offspring in parentage analysis = 4.4cm	3.5 - 6.0cm	
S_e	egg-recruit survival			
E_c	eggs per clutch	depends on female size (eqn. 4)		relationship from Yawdoszyn et al. (in prep)
b_e	coefficient for eyed eggs	-0.608		Yawdoszyn et al. (in prep)
b_l	slope 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	

Table 1:

Results

Figure 2: abundance trends through time with some sort of time series analysis

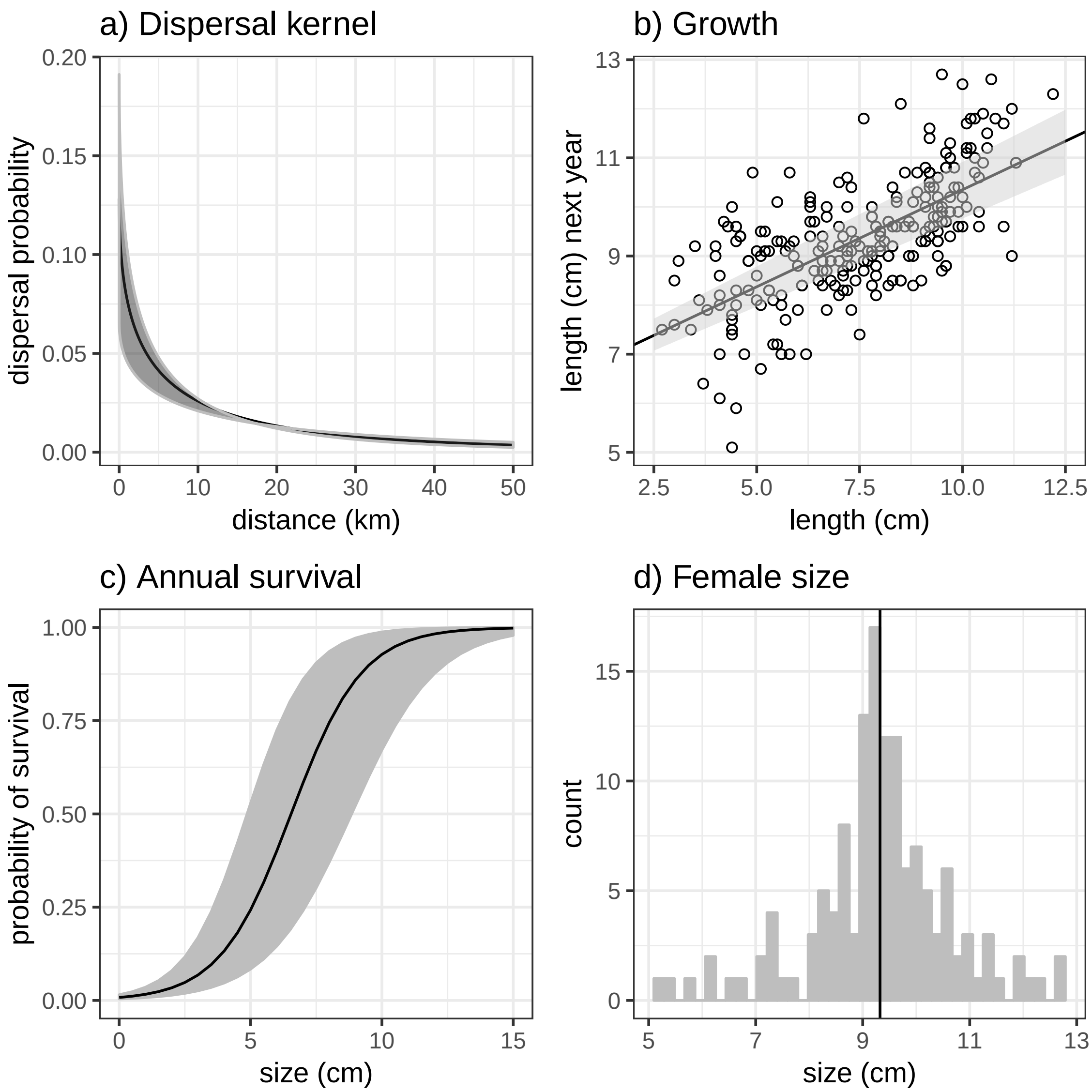

Figure 2: WRITE A CAPTION!

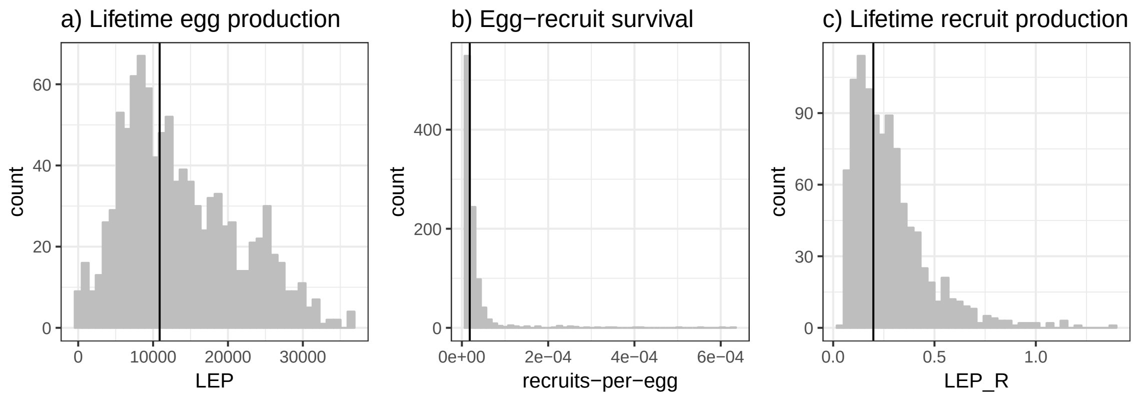

Figure 3: Metrics with best estimate (using recruit size of mean of offspring size) and uncertainty. WRITE A CAPTION!

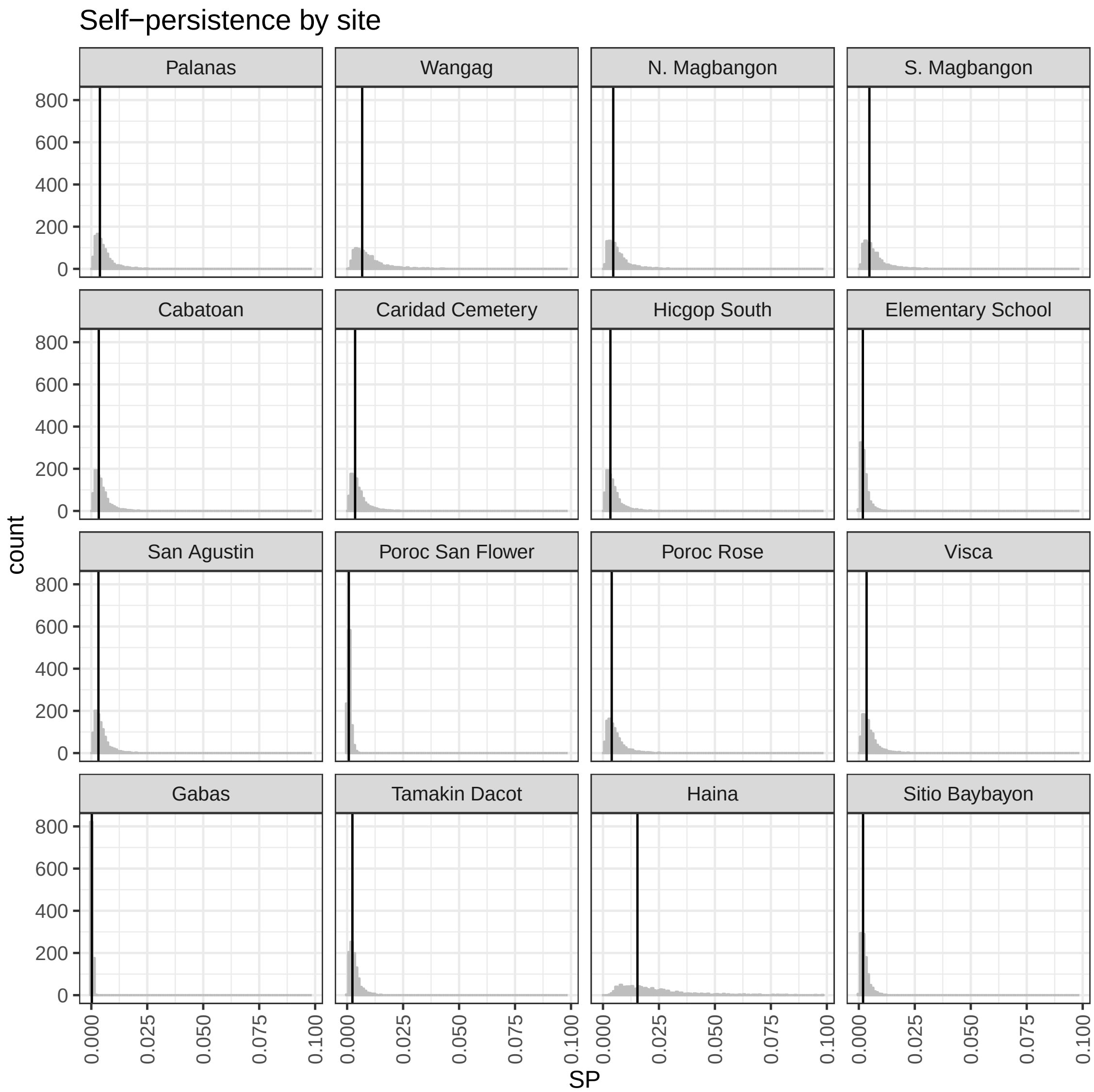

Figure 4: Metrics with best estimate (using recruit size of mean of offspring size) and uncertainty. WRITE A CAPTION!

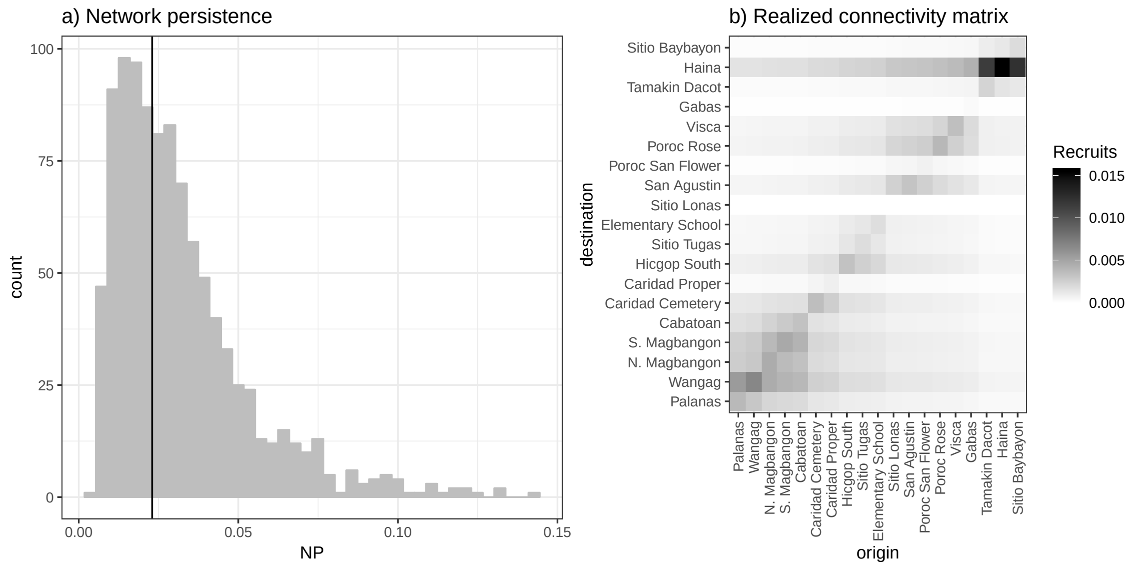

Figure 5: Metrics with best estimate (using recruit size of mean of offspring size) and uncertainty. WRITE A CAPTION!

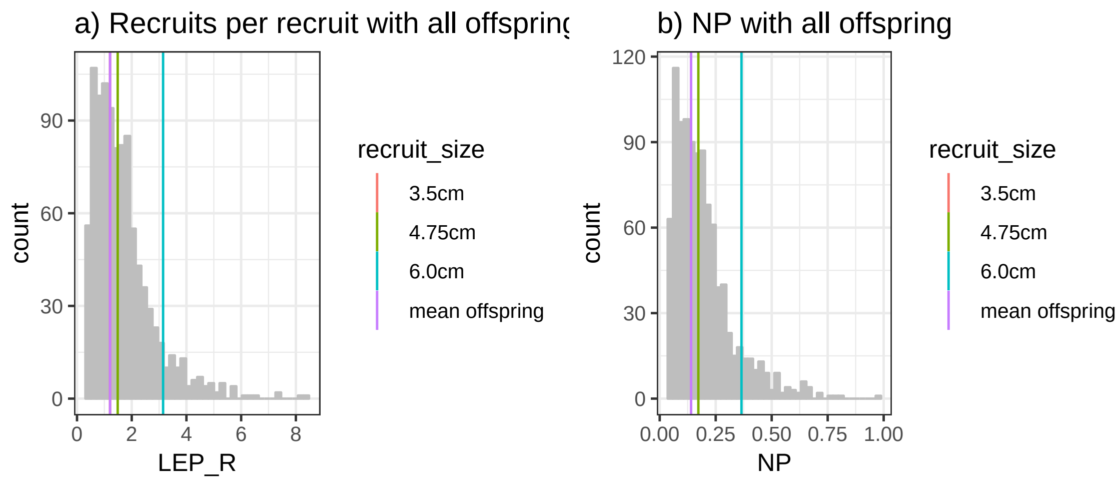

Figure 6: Range of parameter inputs for uncertainty runs with all uncertainty included. Census size is the size at fish are considered to have recruited, such that egg-recruit survival ends. Female transition is the size at which fish transition from male to female and their reproductive output is included in the estimate of lifetime egg production (LEP). FINISH LISTING PARAMS!

Discussion

previous research suggests 10km dispersal kernel spread for yellowtail clownfish (Pinsky et al., 2010)

Appendix

A Method details

Proportion of habitat sampled

Full set of MARK models

We consider the following set of models in MARK:

Model	**Model description**	**AICc**	**dAICc**
	survival size, recapture size+distance	3348.861	0
	survival size, recapture distance	3359.998	-11.1371
	survival constant, recapture distance	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 constant	3570.908	222.04690

Table A1:

B Uncertainty details

B.1 Sensitivity to parameters

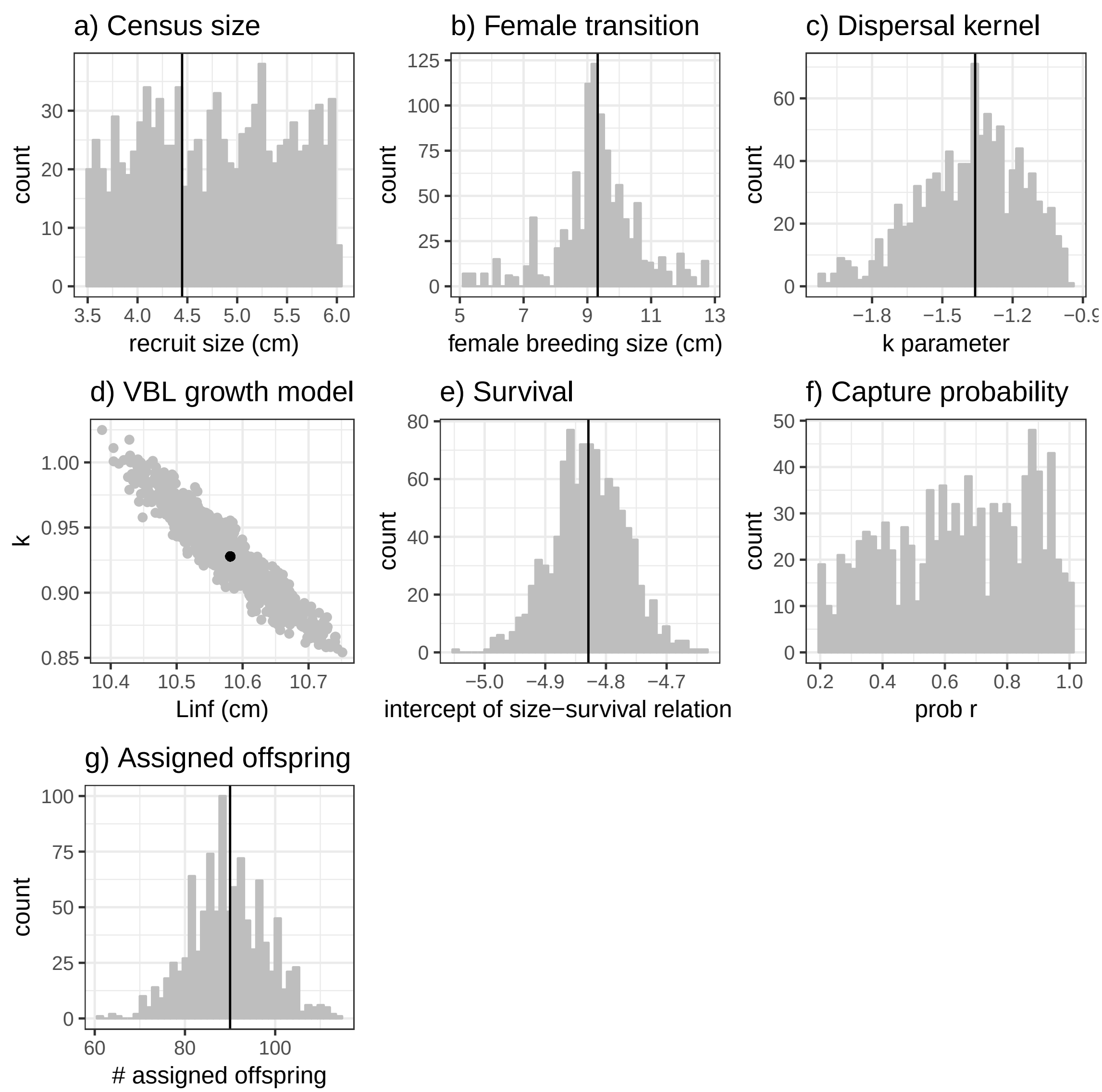

Figure B.1: Range of parameter inputs for uncertainty runs with all uncertainty included. Census size is the size at fish are considered to have recruited, such that egg-recruit survival ends. Female transition is the size at which fish transition from male to female and their reproductive output is included in the estimate of lifetime egg production (LEP). FINISH LISTING PARAMS!

B.2 Effects of different types of uncertainty on metrics

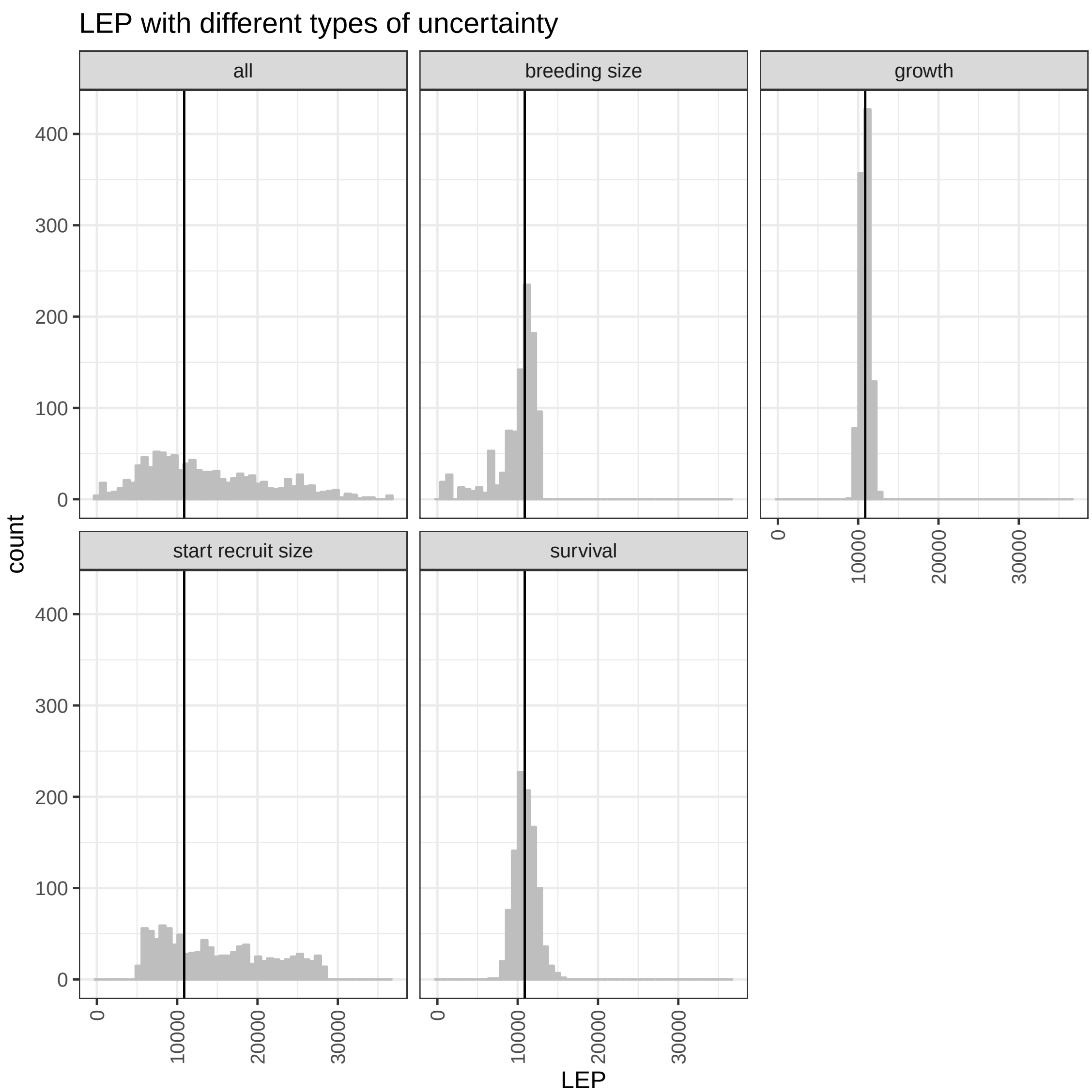

Figure B.2: WRITE A CAPTION!

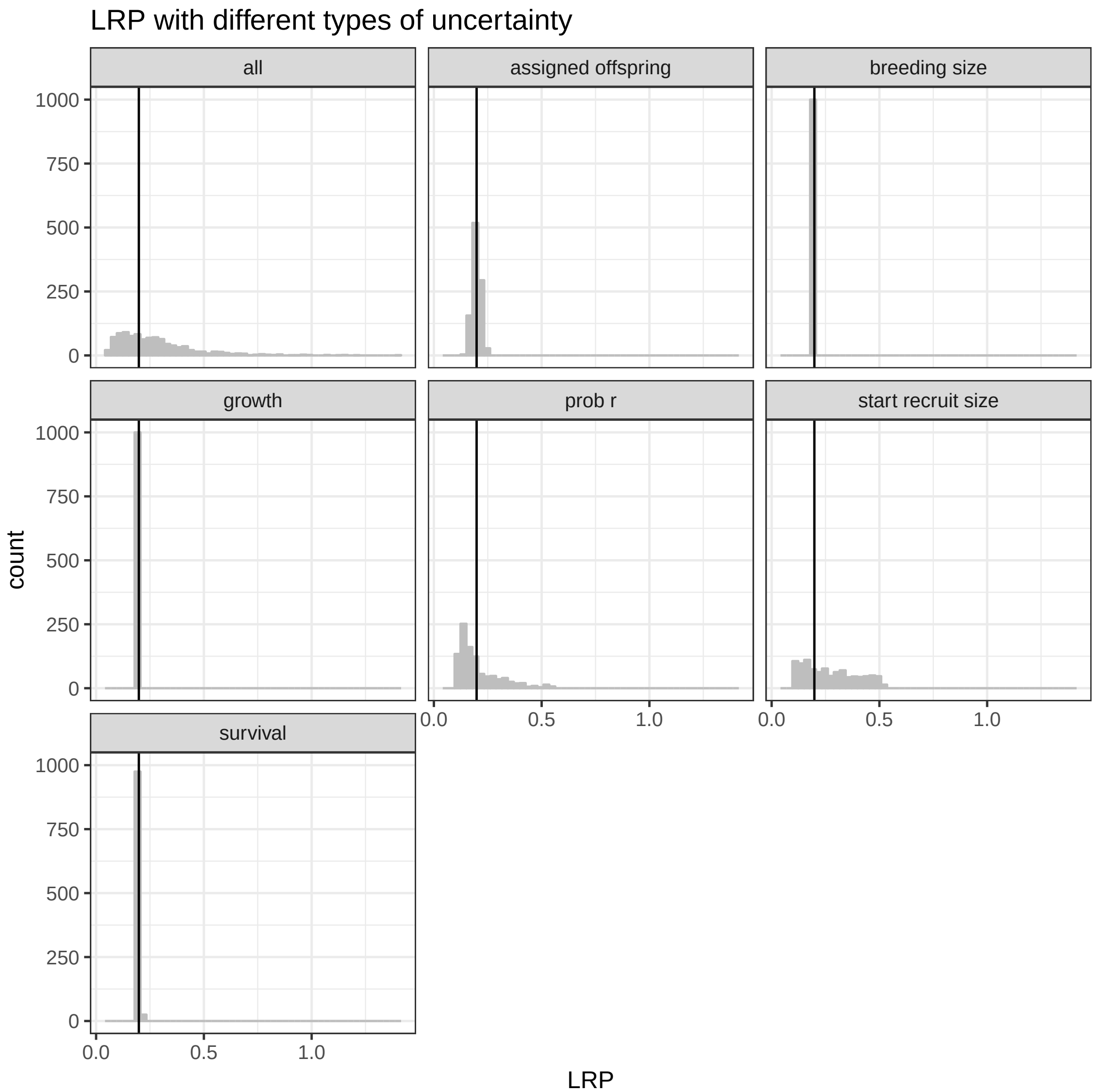

Figure B.3: WRITE A CAPTION!

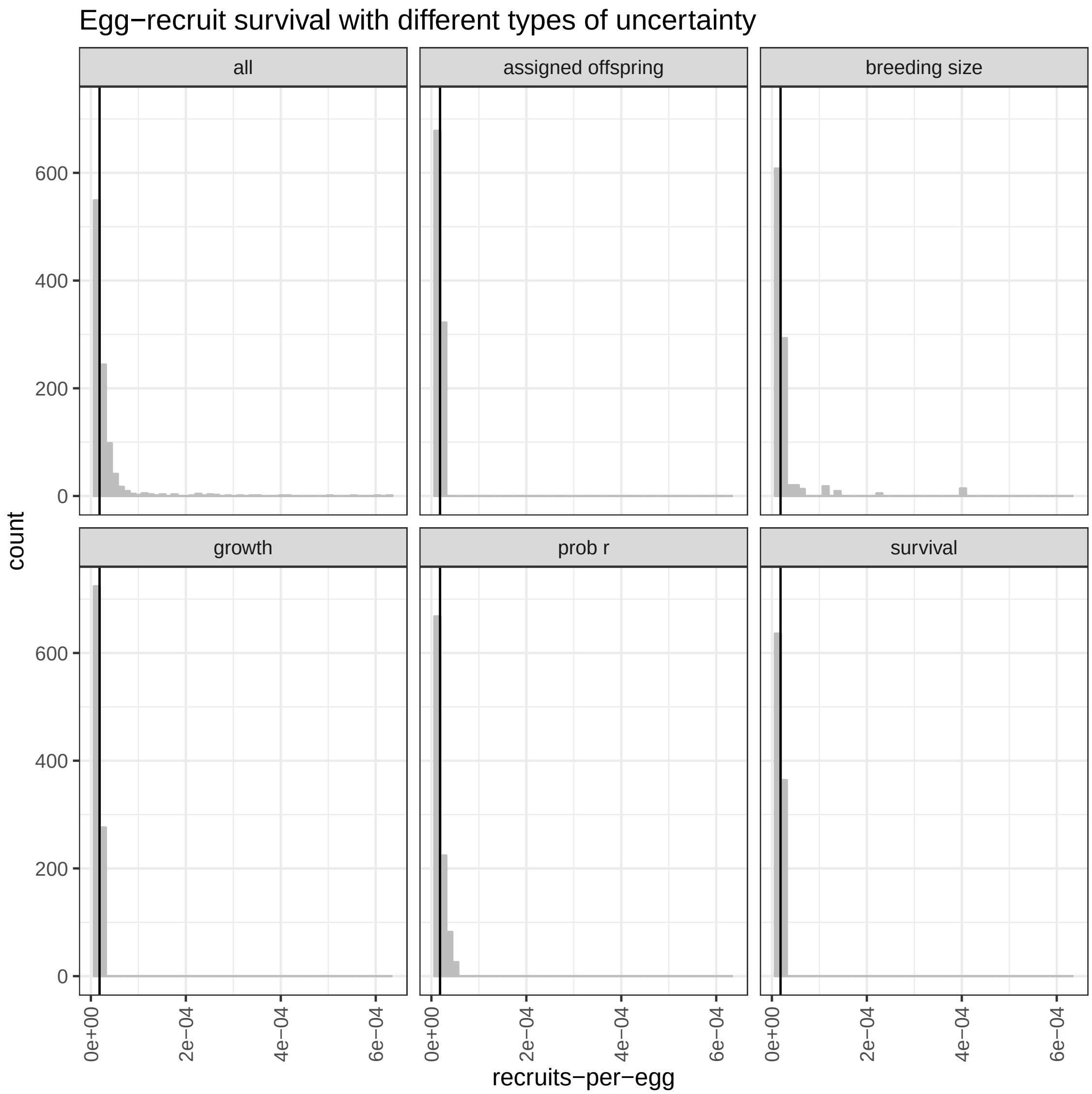

Figure B.4: WRITE A CAPTION!

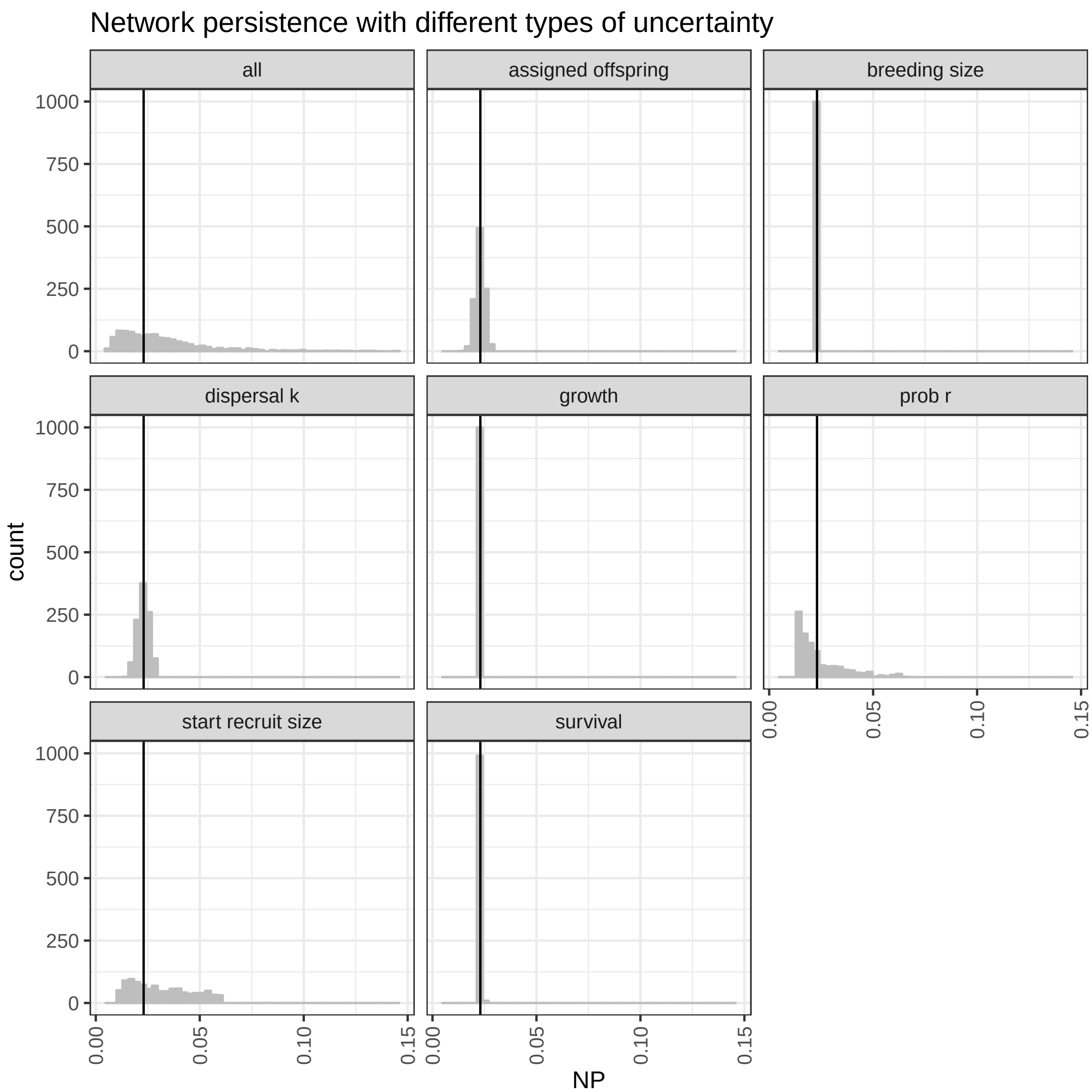

Figure B.5: WRITE A CAPTION!

References

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.

Scott C Burgess, Kerry J Nickols, Chris D Griesemer, Lewis AK Barnett, Allison 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. *Ecological Applications*, 24(2):257–270, 2014.

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

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

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.

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 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 returning home. *Proceedings of the National Academy of Sciences*, 103:6067–6072, 2006.

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, 2017.

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.

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/Publications/ProcRpt/PR2013-01.pdf`.

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.

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.

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.

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/`.