

Breeding ecology parameterization based on FdN data

Vic Quennessen, viquennessen@gmail.com

2025-11-19

Model parameters to estimate:

1. distribution of the number of males that females mate with
2. the way that multiple males fertilize offspring in a clutch
3. the remigration interval for resident males

Overview of reproductive sub-model:

For each sex, age class, year, and simulation, I am tracking the number of individuals as well as the mean, median, and variance of genotypes (which are tracked at the sex and age class level). For each year and simulation, I am also tracking the operational sex ratio of the population.

For each breeding season, after survival and maturity are stepped forward in the population dynamics sub-model, the number of adults that become available to breed are the sum of successful draws from a binomial distribution with probability equal to one over the remigration interval (Q3). For example, if there are 5 females with a remigration interval of 5 years, I would take 5 draws from a binomial distribution with probability equal to $1/5 = 0.2$.

Once I have the number of available males and females, I can calculate the operational sex ratio (number of males / number of all available adults), and use the mating function to translate that to reproductive success.

The number of actually breeding females is then the sum of draws from a binomial distribution with the probability of successfully finding a mate equal to the reproductive success. For each female who breeds, she will mate with a certain number of males (Q1), to be drawn from a 'breeding pool', where each male is represented X times based on the minimum OSR for each mating function. For example, for the mating function with a minimum OSR of 0.05, each male can theoretically mate with 19 females, so there will be 19 copies of each male in the breeding pool.

Then, she will lay a number of clutches and a number of eggs for each clutch drawn from the distributions of clutches per females and eggs per clutch from our FdN field data.

For each clutch, the maternal genotype is based on the breeding female, and eggs are assigned paternal genotypes based on the pattern of multiple male fertilization (Q2). These are used to calculate the offspring genotypes, and then the offspring phenotypes. Evolvable traits are the midpoints of the temperature-mediated emergence success and the thermal reaction norm (i.e. the pivotal temperature). Then, offspring survive based on temperature-mediated emergence success and are assigned sex based on the clutch temperature and the thermal reaction norm.

I'm only using data from clutches with a high nest sample size, but I am retaining data from Season 1 since the partial season should not affect how many males a female mates with (assuming all males are represented in all clutches of females they mate with), the way that ranked males fertilize offspring in individual clutches, or skew the remigration interval analysis in the way that we'd expect (i.e. to have migrating males that we'd miss observing).

1. how many males do females mate with in one season?

In order to more accurately model evolution with maternal and paternal genotypes, I will build a ‘breeding pool’ of males, where each male will be represented X times, where X is the average number of females that males can mate with, rounded to the nearest integer, dependent upon the mating function. Then, for each female, I need to be able to determine how many males she will mate with to draw that many unique males from the breeding pool.

Number of males	1	2	3	4	5	6	7
Count	13	44	18	52	0	12	7
Proportion	0.09	0.30	0.12	0.36	0	0.08	0.05

Table 1. Number of males that females mated with in one season

Min	1st Q	Median	Mean	3rd Q	Max	Variance	SD	Skew	Kurtosis
1.000	2.000	3.000	3.315	4.000	7.000	2.438	1.561	0.607	-0.199

Table 2. Summary statistics of number of males females mated with

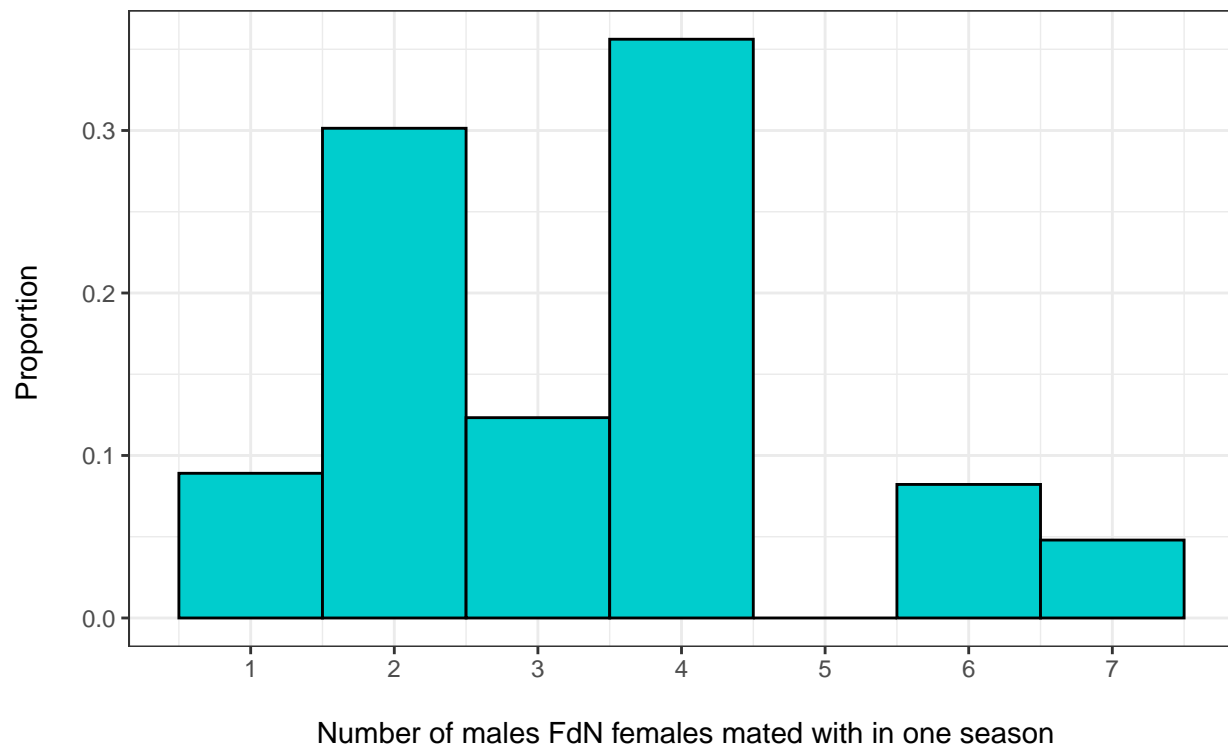


Figure 1. Density histogram of number of males females mated with per season

why fit a model?

we can't just use the proportions of females that mated with 1 - 7 males since sample sizes are small, and no recorded female mated with 5 males.

Distribution	Log likelihood	AIC	BIC
discrete gamma	-263.2182	530.4363	536.4035
poisson	-270.5726	543.1452	546.1288
negative binomial	-270.5726	545.1452	551.1124
geometric	-341.0678	684.1356	687.1193

Table 3. Model fits

The discrete gamma distribution had the best fit with the lowest log likelihood, AIC, and BIC values.

model output:

```
## [1] "coefficients:"
##      shape      rate
## 4.332093 1.306867
```

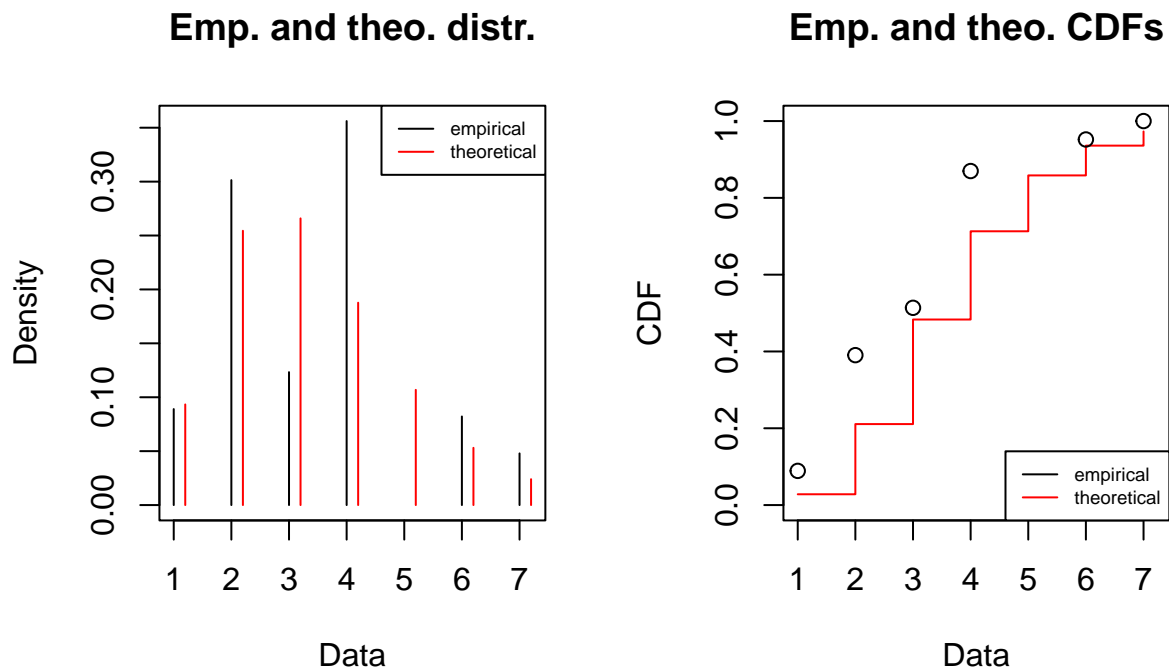


Figure 2. Discrete gamma model fit for the number of males females mate with in one season

Black lines indicate the density of males females mated with in our FdN dataset, while red lines indicate the predicted density based on the discrete gamma model fit to the data.

given this discrete model...

If we assume that females can mate with up to 10 males (i.e. we assume that we did not observe all contributing males in our dataset), then we can predict the proportion of females that would mate with 1 to 10 males. Unfortunately, the discrete gamma distribution has a tail that cannot be incorporated into a discrete number of potential male values, so the sum of the weights does not add up to 1. In order to use these probabilities in my model, I have to normalize the weights to add up to one.

Number of males	Model weight	Normalized weight
1	0.183	0.188
2	0.272	0.280
3	0.230	0.236
4	0.146	0.150
5	0.078	0.080
6	0.037	0.038
7	0.016	0.017
8	0.007	0.007
9	0.003	0.003
10	0.001	0.001
Sum	0.973	1.000

Table 4. Weighted probabilities for females to mate with 1 to 10 males in any given breeding season.

2. how do multiple males fertilize nests?

Within each clutch, do males fertilize offspring randomly, such that each male has the same probability of fertilizing any individual offspring? And do those probabilities change as the total number of contributing males changes?

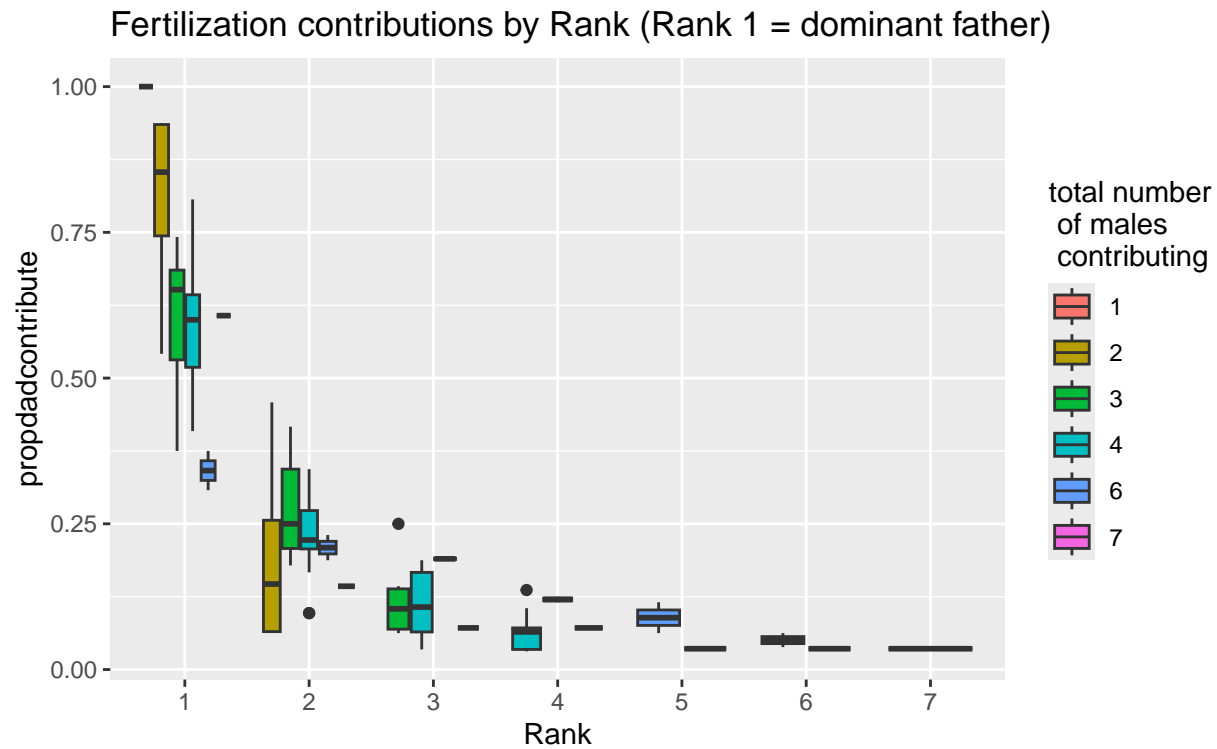


Figure 3. Fertilization contributions organized by father 'rank'

Where rank indicates proportion of offspring fertilized by the r^{th} most dominant father, i.e. Rank 1 indicates the father that fertilized the most offspring sampled from that clutch. Per this data, it looks like the total number of contributing fathers does matter, but samples sizes are too small to extrapolate the effect, so the model was fit to grouped data where only Rank mattered.

model fit

I chose to use a power law because mechanistically it makes sense that for the first male turtle, a certain proportion of eggs will be fertilized. For the second male, a certain proportion of what is left, and for the third male and so on it will always be a proportion of what eggs are left to fertilize, which can easily be quantified using a power rule with a negative exponent.

Power law: $y(x) = a \cdot p^{-x}$

```
## [1] "coefficients:"
```

```
##           Estimate Std. Error   t value    Pr(>|t|)
## (Intercept) -0.3764756 0.05857507  -6.427233 1.784865e-09
## log(Rank)    -1.7101305 0.07090767 -24.117708 1.988142e-52
```

```
## [1] "adjusted R^2:"
```

```
## [1] 0.800183
```

Where the intercept $a = \exp(-0.376) = 0.687$ and the exponent $x = -1.710$:

$$\text{Proportion father contributes}(\text{Rank}) = 0.687 \cdot \text{Rank}^{-1.710}$$

model predictions

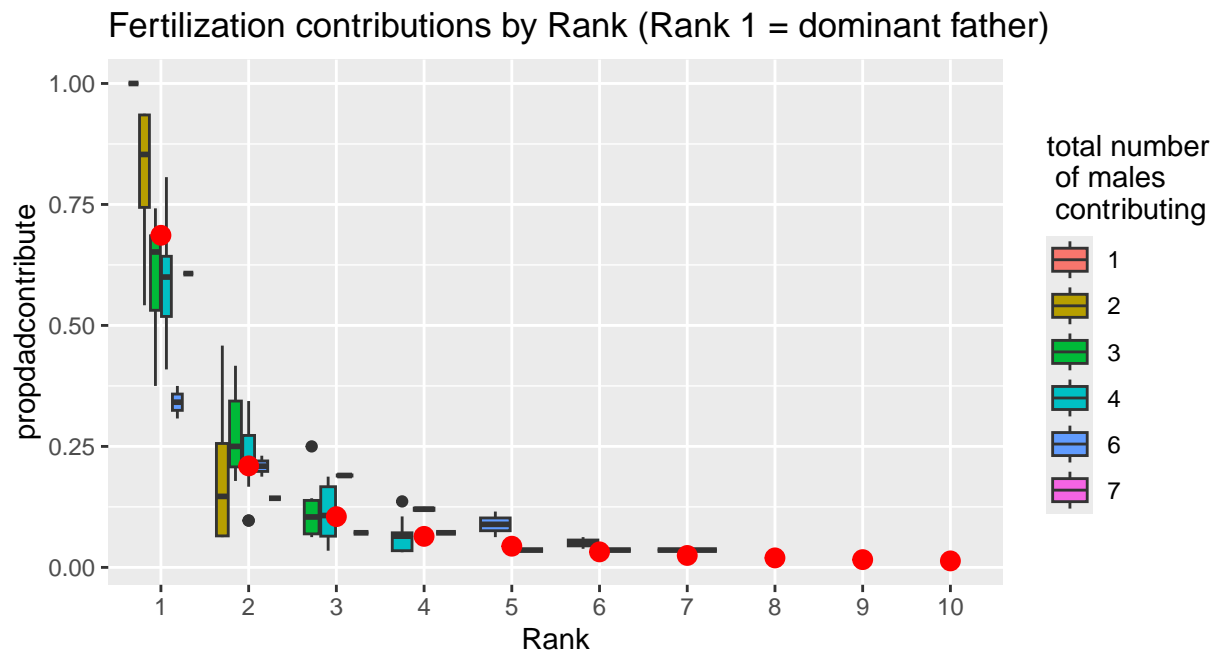


Figure 4. Power law model fit to paternal fertilization contributions organized by father 'rank'

Turtles work like gravitational waves! - Jerry

3. how often do males remigrate

male remigration interval used in Chapter 2: 1.47 (years)

for male green turtles at Atol das Rocas (Grossman et al., 2019)

<https://www.int-res.com/abstracts/meps/v609/p197-207/>

FdN genetic data

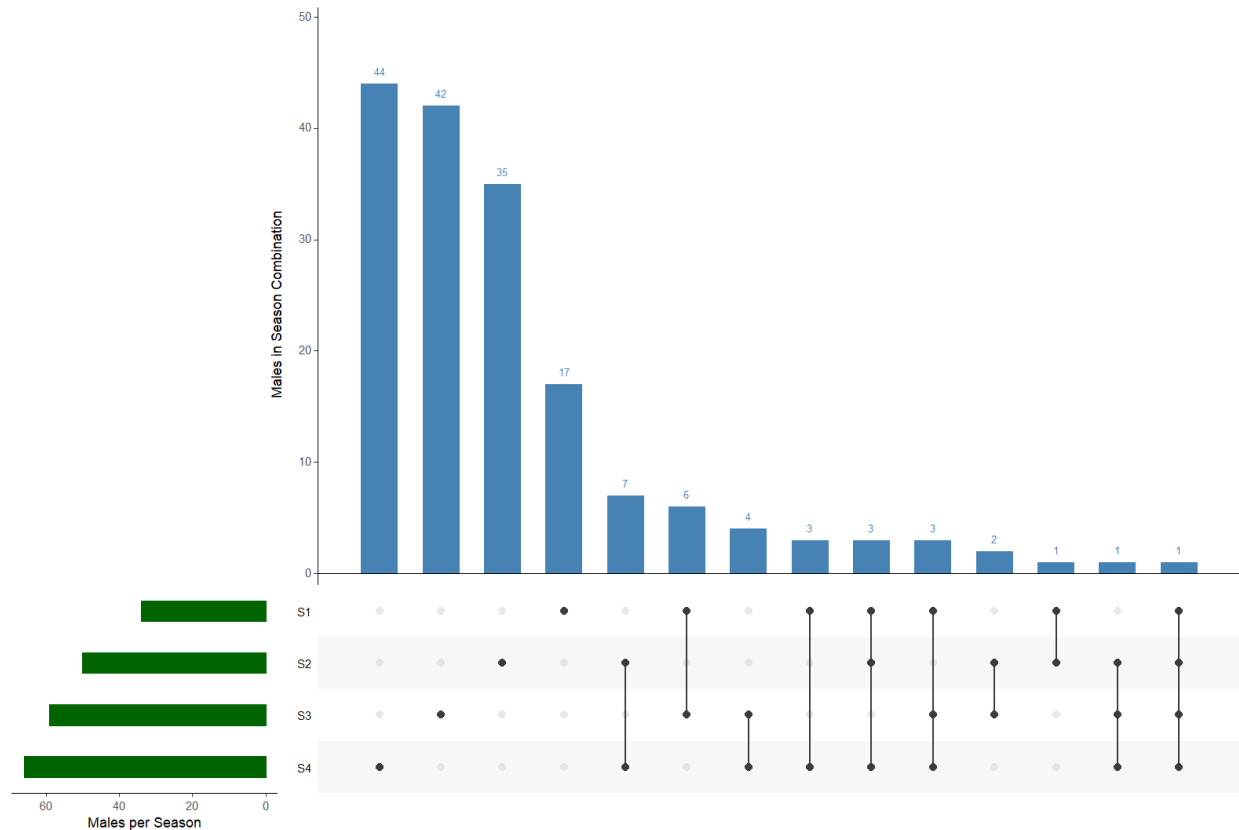


Figure 4. FdN male remigration data

Original data, obtained from Dr. Lisa Komoroske (lkomoroske@umass.edu) on Thursday October 9, 2025. Note: "...Estefany made this upsetR graph for Mariana for an upcoming talk, helps visualize the observed remigration interval a bit for the males (again, keep in mind the difference in sampling effort in S3 and S4, these will need to be normalized to the number of samples run eventually).

fit a binomial model

marginalized likelihoods over number of 0 and 1 successes

Given a binomial model, or any resident male green turtle, the probability of seeing it $k = 1, 2, 3, \text{ or } 4$ times across $n = 4$ years is equal to:

$$L(k) = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k}$$

We know that $n = 4$, since our data were collected over 4 years, giving 4 chances to observe each male. 66 males were observed in season 4, meaning they could have been observed for all previous seasons. We can assume that those males that were observed more than once were residents. For the remaining 44 males, some unknown number of them were residents and others were transient males. For RM1 resident males observed once, there were RM1 + 22 total resident males observed in our dataset.

Given probability p , the likelihoods of observing RM1 males once, 14 males twice, 7 males 3 times, and 1 male 4 times is the probability of seeing a resident male $k = 1, 2, 3, 4$ times in any given trial of size $n = 4$ years to the power of the count $c = RM1, 14, 7, 1$ times it was observed:

$$L(k = 1 \mid p) = \left(\binom{4}{1} \cdot p^1 \cdot (1 - p)^3 \right)^{RM1}$$

...

$$L(k = 4 \mid p) = \left(\binom{4}{4} \cdot p^4 \cdot (1 - p)^0 \right)^1$$

And the total likelihood of seeing all of these is the product of each probability times the number of potential combinations of the order in which they were observed:

$$L(data \mid RM1, p) = \frac{(RM1 + 14 + 7 + 1)!}{RM1! 14! 7! 1!} \cdot \prod_{k=1}^{k=4} P(k \mid RM1, p)$$

The conditional likelihood is then the probability of observing RM1 males once out of all males observed (RM1 + 22) given probability p :

$$P(RM1) = \binom{RM1 + 22}{RM1} \cdot (p(1 - p)^3)^{RM1} \cdot (1 - p(1 - p)^3)^{22}$$
$$P(RM1 = 1) = \binom{23}{1} \cdot (p(1 - p)^3)^1 \cdot (1 - p(1 - p)^3)^{22}$$

The marginalized likelihood for each value of RM1 is then the likelihood times the conditional likelihood, and the total likelihood for each probability P is equal to the sum of all the marginalized likelihoods across all potential values of RM1.

should we use all males or only males seen in Season 4?

If we compare the results from analysing only those resident males observed in Season 4 (as opposed to all resident males), the results do not substantially change. The maximum likelihood estimator for P drops from 0.55 to 0.52, but the overall likelihoods are much smaller.

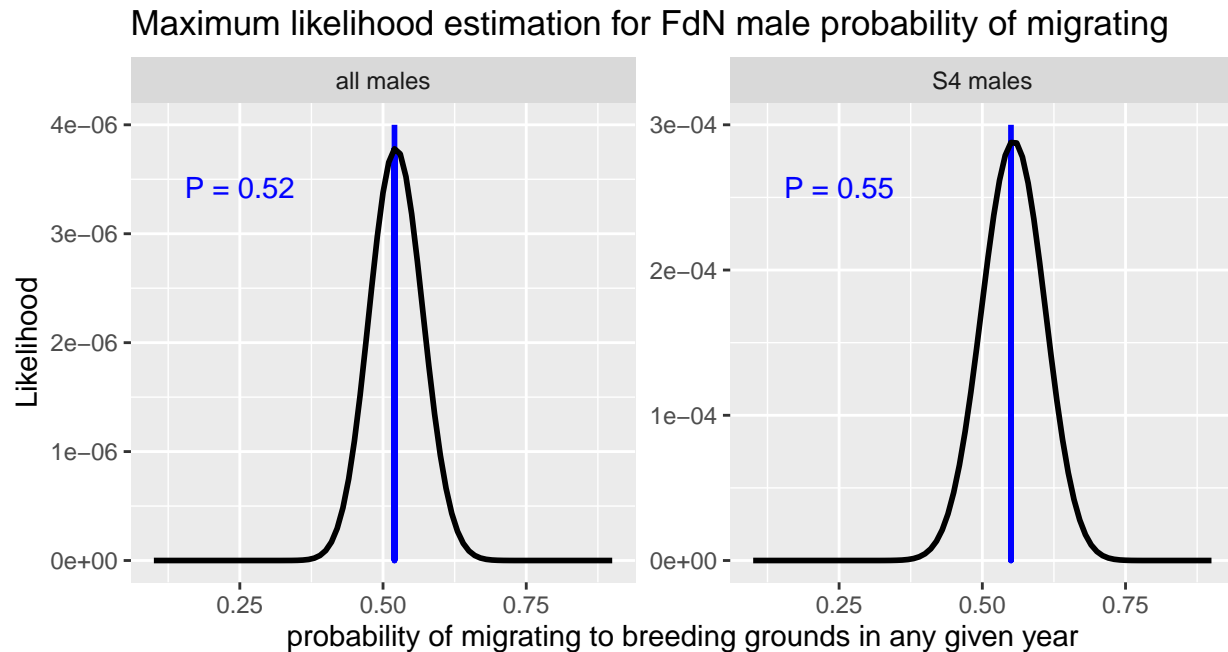


Figure 5. Likelihood of different probabilities P of FdN resident males mating in any given breeding year

Analysis	Log likelihood	Probability	Remigration Interval
All males	0.00000378	0.52	1.92
Season 4 males only	0.000288	0.55	1.82

Table 5. Maximum likelihood estimators for probability P of any one resident FdN male migrating to breeding grounds in any breeding season.

The final maximum likelihood estimator for the binomial probability of males mating is therefore $p = 0.55$, with an associated average remigration interval of 1.82 years.

Acknowledgments

Research leads:

- Lisa Komoroske
- Mariana Fuentes
- Will White

Field work:

- Armando J. B. Santos
- Cintia Miranda
- Blair Bentley
- Estefany Argueta Herrera

Lab work:

- Estefany Argueta Herrera