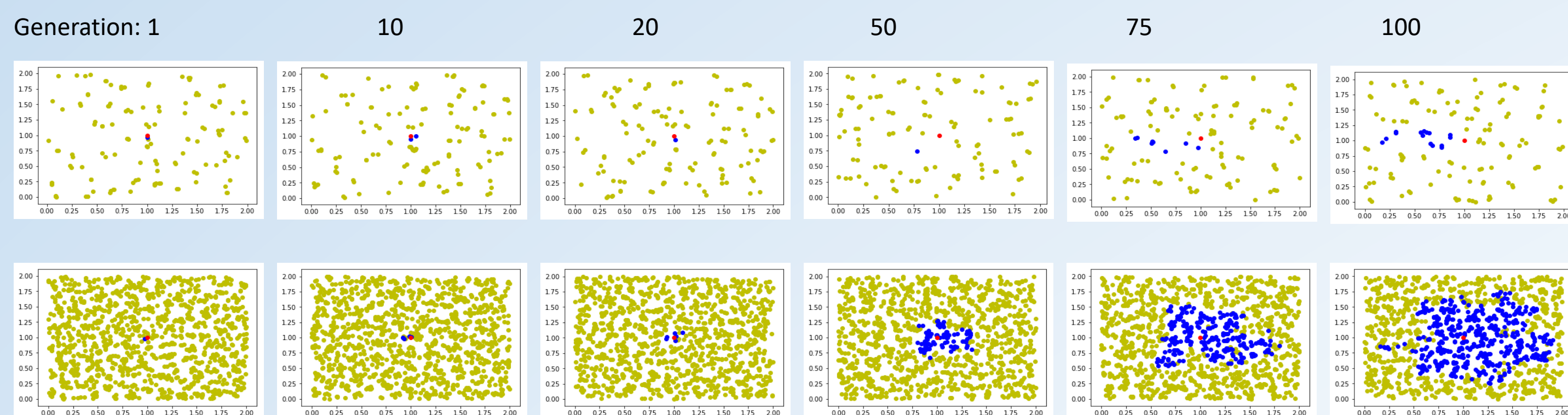




Mutations create genetic variation within the genetic code. This genetic code and thus the variation is passed on from parent to offspring each generation. When parents have multiple offspring a specific sequence of code can by chance, begin to spread through a population. Some genetic sequences, (genes) can increase the probability of an individual producing successful offspring, thus increasing the chance the sequence will spread.

We use simulations to model how quickly new mutations spread and compare that to several mathematical models.



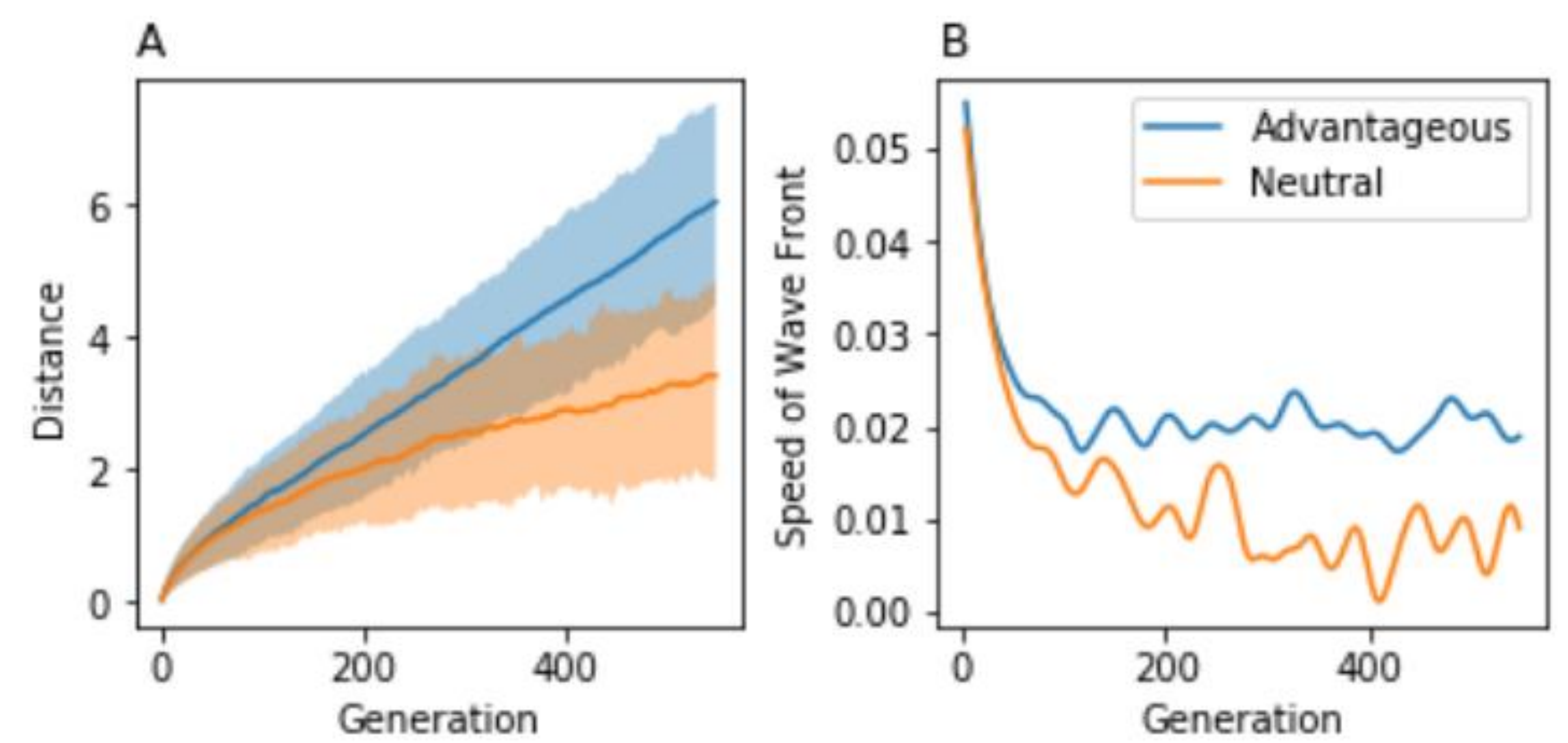
These graphs show a little slide show. From the left to the right, we see that as we travel forward in time the individuals who contain the advantageous allele (colored blue) spread out from the center. Top row: A sparse population. Bottom row: A dense population.

Simulations

Let N individuals be distributed uniformly in a $[0, B]^2$ space. At each time step (generation) individuals have k offspring, where k comes from a Poisson distribution. Offspring are placed into the space according to a Gaussian distribution centered about the parent with variance σ^2 . Density dependent selection then occurs. The more individuals there are in an area, the less likely each individual will survive. The probability of individual i surviving to adulthood is:

$$\min \left(1, \frac{2K\pi\sigma^2}{\sum_j e^{-\frac{d_{ij}^2}{2\sigma^2}}} \right)$$

Where K is the carrying capacity (Expected number of individuals per unit area) and d_{ij} is the distance between individuals i and j . From the simulations on the left it is clearly noticeable that advantageous alleles spread more quickly in dense populations.



Calculating the Spread

The individual closest to the center is given a new mutation with a predefined selective advantage s . Over time, the individuals who carry the mutated allele (mutants) either go extinct or begin to multiply and spread out. At each generation, after density dependent selection, the distance to the center is calculated for each mutant. The largest distance is saved as the front edge of the wave of advance. After running thousands of simulations, averages were used to calculate the numerical derivative.

Ronald Fisher (1937) first tried to estimate the speed of the wave of advance using a travelling wave solution to the PDE:

$$\frac{\partial p}{\partial t} = k \frac{\partial^2 p}{\partial x^2} + f(p).$$

Where p is the frequency of the new allele, and k is the diffusion constant. Using the appropriate initial conditions, the velocity of the travelling wave solution comes out to be $\sigma\sqrt{2f'(0)}$ which for the shown simulations comes out to be 0.0316. However, Fisher's equation doesn't account for population density, and essentially assumes an infinite population density.

The spread of Neutral mutations was first examined by Sewall Wright (1943, 46). He

estimated that ancestors could be found within $\sigma\sqrt{\frac{\pi t}{2}}$. At first glance it appears correct.

Improving the Calculations

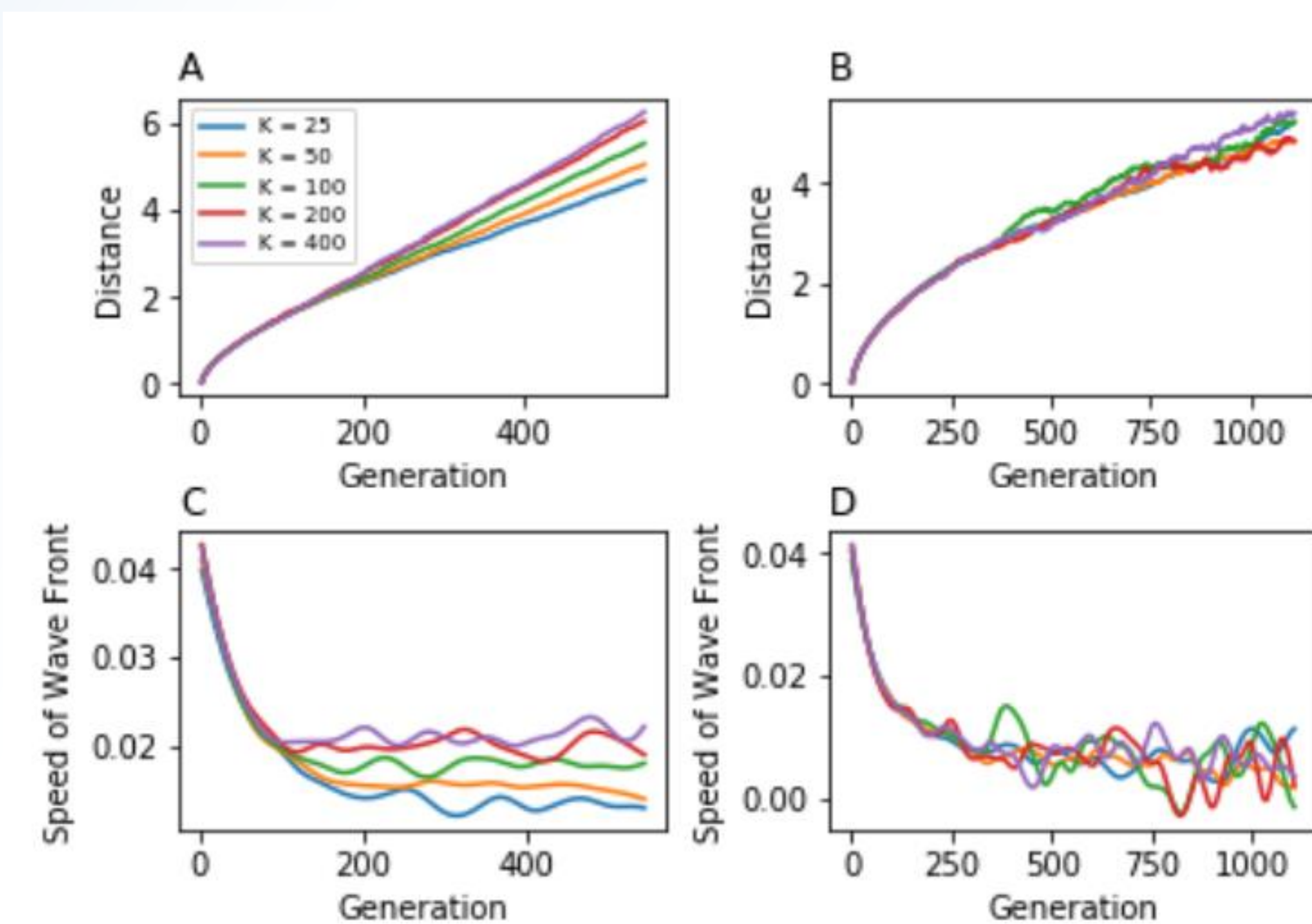
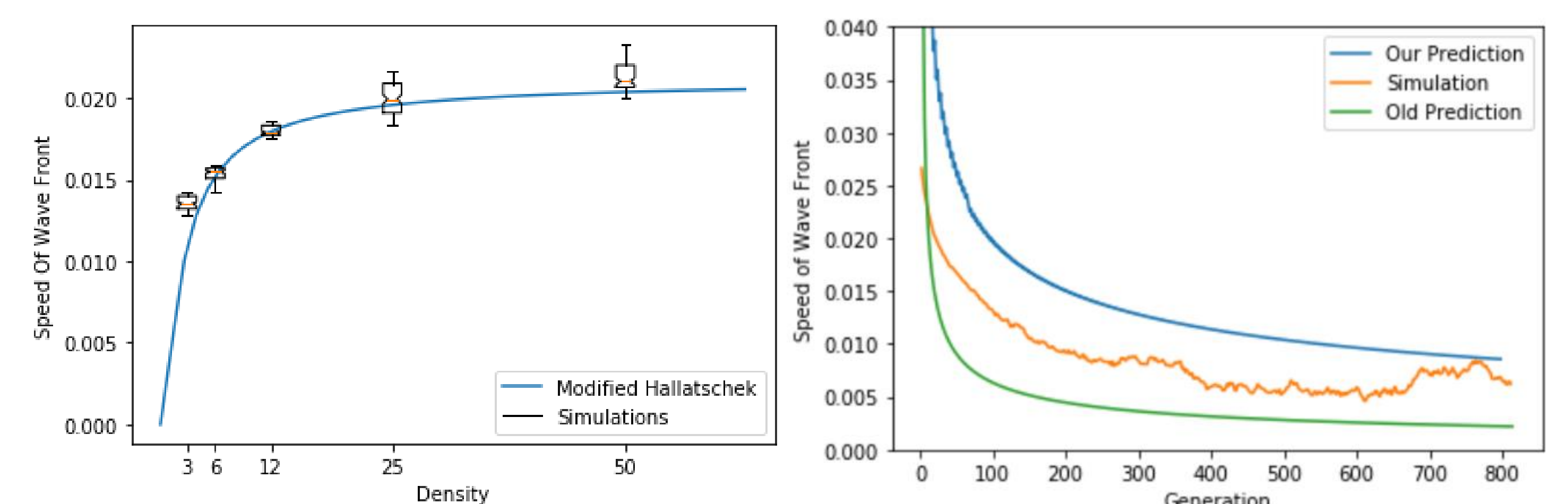
Many have notice that finite population sizes, (Especially sparse populations) decreases the speed of the wave of advance of advantageous alleles. Using an idea from Hallatschek (2010) and extending it to 2d we obtain the equation $\sigma\sqrt{\frac{4cs}{\pi}}$ where c represents the effect of a finite population. A small c corresponds to a slow speed while a large c leads to a large speed. We know that there is an upper limit to the speed as the population density tends to infinity.

Using our simulations to tune c , we define $c = \frac{e^{-\frac{1}{N_w}}}{\sqrt{2}}$, where N_w is population density scaled by dispersal distance.

We also noticed, that while the spread of neutral alleles follows the same trend as expected by Sewall Wright, they do seem to spread somewhat faster. We calculate the expected maximum distance of k independent random Gaussian walks as an upper bound for the distance a neutral allele

$$E(d) = \sigma k \sqrt{\frac{t\pi}{2}} \sum_i^{k-1} \binom{k-1}{i} \frac{-1^i}{(i+1)^{\frac{1}{2k}}}$$

Noticing that the spread of neutral alleles is independent of population density we calculate the expected number of offspring k in generation t using a Poisson distribution conditional upon non-extinction.



References

O. Hallatschek and D. R. Nelson. Life at the front of an expanding population. *Evolution: International Journal of Organic Evolution*, 64(1):193–206, 2010.

S. Wright. Isolation by distance. *Genetics*, 28(2):114–138, 1943. ISSN 0016-6731. URL <https://www.genetics.org/content/28/2/114>.

S. Wright. Isolation by distance under diverse systems of mating. *Genetics*, 31(1):39, 1946.

R. A. Fisher. The wave of advance of advantageous genes. *Annals of Eugenics*, 7(4):355–369, 1937. doi: 10.1111/j.1469-1809.1937.tb02153.x. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1469-1809.1937.tb02153.x>.

Results

Our Modified-Halletchek equation follows the trend of the data quite nicely, it is much more accurate than the original Fisher equation. And although c was tuned to fit the simulations of a single set of parameters, the curve fits just as nicely with a wide range of selective coefficients and dispersal distances.

Our upper bound for the neutral allele also follows the trend and seems to sandwich the actual speed right in the middle between our bound and the prediction of Wright.