**Model design**

We developed a non-spatial agent-based simulation model of the evolution of emergence behavior in order to develop an understanding for when organisms should evolve to depend on different environmental cues, and if (and how) different climatic regimes should select for the use of different environmental cues.

In our simulation, a fixed number of organisms choose to emerge based on environmental conditions, gain fitness depending on the environmental conditions, and then produce offspring based on the fitness gained. Offspring largely resemble their parents, but with some chance of mutation in the way they decide to emerge.

**Traits**

We have a set of ten traits from which some subset of traits (usually 2-3) is used for any given scenario.

Those traits are as follows:

1. b.day – responsiveness to day of year
2. b.temp – responsiveness to temperature of current day
3. b.precip – responsiveness to precip of current day
4. b.cutemp – responsiveness to cumulative temperature for the year
5. b.cuprecip – responsiveness to cumulative precip for the year
6. b.daysq – responsiveness to the square of the current day
7. b.tempsq – responsiveness to the square of the current temperature
8. b.precipsq – responsiveness to the square of the current precip
9. b.cutempsq – responsiveness to the sum of the squares of all of the daily temperatures up to and including the current day
10. b.cuprecipsq – responsiveness to the sum of the squares of all of the daily precips up to and including the current day

At the start of a scenario, we set initial trait values for a population (for the subset of traits in use) as follows:

We calculate the minimum and maximum values of the associated cue for a subset of the year data to be used (currently the first 5 years). We then assign trait values to each individual drawn from a normal distribution with a minimum of the minimum found from the data, and a maximum of the maximum value found in the data times the total number of traits the population is using. These values were selected to provide a distribution of emergence times for the initial population that are distributed somewhat evenly between day 1 and day 365.

**Emergence Decisions**

The backbone of our model is how organisms decide when to emerge. They do so by following a simple rule. Each individual is characterized by some number of traits (see the traits section). As examples here, we use b.day, b.temp, and b.precip. Those traits are numbers from the set of real numbers (so positive or neg, can be fractions). Every day, organisms decide to emerge or not by calculating an emerge value E based on their traits and the environmental conditions:

E = 100/b.day \* day + 100/b.temp \* temp + 100/b.precip \* precip

where day is the day of the year (which organisms can determine using photoperiod), temp is the day’s max temperature, and precip is the day’s precipitation. The first time an individual has an E value greater than 100, they emerge. Note that if a trait value is zero, it is instead set to 10^-10.

Our original method for calculating E this called for a more simplified equation:

E = b.day \* day + b.temp \* temp + b.precip \* precip

The problem with the simplified version was that, since the decision to emerge is when E > 100, the same small mutation had a disproportionately large effect when the original trait was small. With the current version, the b.day trait represents the day the organism will emerge if all other traits are zero. Not only is this interpretation easier, but the effect (day of emergence) now scales linearly with the trait (instead of with 1/trait).

Because organisms emerge the first time their E value exceeds the threshold of 100, they are biased favor of earlier emergence times; this is biologically reasonable as we’re interested in emergence from overwintering.

Individuals emerge on the last day of the year if they haven’t before.

**Fitness**

When an individual emerges, they “collect” daily fitness for every day of their lifespan (the default is 10 days). Fitness for each day is calculated from precipitation and temperature, using the equation

F = f(precip)\*g(temp)

Where f and g are Gaussian functions centered on an optimal precipitation (currently 10 mm) and temperature (currently 30 degrees C), and precip and temp are the current day’s precipitation and temperature measures, respectively. This measure of fitness is the same for all individuals regardless of their traits – for any given year there is a single fitness curve through time, and organisms do well or poorly depending on where on that curve they emerge. The exact values were selected to give fitness surfaces that qualitatively matched known patterns of insect fitness through time.

We have been discussing using an asymmetric fitness surface for added realism, but have not yet decided on the distribution or implemented it.

Fitness is not gained for days after the last day of the year, so if an individual emerged on day 364, they would only obtain fitness for two days (364 and 365) instead of the full ten days most other individuals had.

**Reproduction and mutation**

Once all individuals emerge and their fitness is calculated, the reproduction step occurs. Parents are assigned offspring using a multinomial distribution where the probabilities are proportional to the parents’ fitnesses. We assume that all individuals can reproduce, and produce offspring that are identical to the parent unless there is mutation, so offspring are assigned trait values identical to their parents. The use of the multinomial distribution makes it trivial to keep the population sized fixed through all generations.

After reproduction, the offspring individuals are assigned mutations randomly. Each trait (b.temp, b.day, b.precip) of each individual has a small independent chance (currently .01) of mutating, and individuals can receive zero mutations or several (only one per trait, though). Any trait that was selected for mutation has a new number added to the current trait value, with the new number drawn from a normal distribution with mean zero, and standard deviation equal to `mutdist’ (currently .01) times the calculated maximum cue corresponding to the trait. As in generating the initial distribution of traits, the calculated maximum cue is the maximum value found in the first five years of climate used.

These new individuals were used as the generation for the following year. This means that the model is representing a univoltine organism.

**Climate Data**

The environmental values are taken from actual data. 101 years of daily weather at Davis. We used the daily precipitation and the daily max temperature as our precip and temp values, respectively. Missing days were imputed using the Amelia package, with negative precipitation values set to zero.