**30 Sep 2011**

CDPOP

* woohoo! I found a ready-made distribution for python including numpy and scipy that I downloaded and installed and that worked!
* Ran the example file, no problem
* **Input files**
  + Xyfilename
    - Coordinates
  + Agefilename
    - Changed age distribution to agedistribution.csv, didn’t work
    - **Lesson: save as Windows-delimited csv, not just csv**
    - Tried to compare results from multiple runs, but Excel won’t let you open multiple files with the same name.
      * **Lesson: use NeoOffice or something else besides Excel to look at multiple files with the same name**
    - Results have negative Fit and Fis and F values. I didn’t think you could do that…
      * F = 1-O/E, so if O is greater than E, then it would be negative. So if you get negative F values, then the observed heterozygosity is higher than you would expect under HW. I think it could mean that there is selection for heterozygs or against homozygs. **Relearn your F statistics!**
    - Why are there subpopulations? **How is a subpopulation defined?**
  + Matecdmat (mating cost distance matrix)
    - ED-cdmatrix16 file: nxn cost distance matrix (ED is euclidean dist)
    - Mate CD: mating movement
  + Dispcdmat (dispersal cost distance matrix)
    - Dispersal CD: dispersal movement
      * I assume this means the greatest distance an individual can go in its lifetime, and that it couldn’t be less than Mate CD.
      * **More detail on Dispersal CD needed!**
      * Maybe it could be a way to distinguish road mortality and avoidance…
    - Mate and Dispersal CD can be the same file.
* **Model parameters:**
  + Mcruns: Number of Monte Carlo simulations
  + Looptime: Simulation run time (n generations)
    - Grid.csv files are output, which ones it writes depend on nthfile choice
    - What are infection, L0A0, L1A1, etc?
  + Nthfile\_choice: choice of which generations to write matrix
    - “list” entry means it’s not a file, but the values in nthfile\_list
    - “sequence” entry means it’s the number of runs in nthfile\_seq
      * I tested the change to sequence, and it did give me 0,1,4,7,10,13,16,19 (why 1 is first still a minor mystery)
  + nthfile\_list: separated by pipe character ( | )
  + nthfile\_seq: integer
  + oldmortprec: percent mortality in the adult population
    - 100 specifies non-overlapping generations
    - **it would be cool if I could specify mortality according to location rather than age… this would be the best way to go about simulation road-induced mortality**
* **Mating parameters:**
  + Matemoveno: mating movement number, the movement function answer for mating
    - 1 = linear
    - 2 = inverse square
    - 3 = nearest neighbor
    - 4 = random mixing
    - 5 = negative exponential
  + matemoveparA: neg exponential only
  + matemoveparB: neg exponential only
  + matemovethresh: cost distance units an individual can search for a mate
    - ‘max’ for all individuals
  + Freplace: can females mate more than once?
  + Mreplace: can males mate more than once?
  + Selfans: allow selfing?
  + Sexans: Y for separate sexes, N for hermaphroditic
  + Reproage: age at reproduction
    - Set to 0 for non-overlapping generations

CDPOP now allows you to specify fitness of genotypes so you can do selection studies. But we know perfectly well that fitness depends on environment. If CDPOP will let you designate differential fitness based on landscape characteristics, maybe it can do road mortality. Basically all genotypes would be less fit in that region. Aha! “mortality fitness surface” is part of CDEVOLVE. Could use it to trick CDPOP into causing general mortality by a road.

11 Jan 2012

Vary the parameters for mating dispersal cost. What if cost goes down for mating movement?

24 Jan 2012

Erin Landguth email on making this work:

Hi Karl.  
Yes, but it will be a bit tricky and you will have to have a few assumptions, the first is you can use as many loci as you want by you can only have 2 alleles.  
1. Use the cdevolve feature and input the mortality numbers based on location in a ascii grid. Put 1 in cdevolve answer spot. Make sure you have 2 as the allele input.  
However, this is a spatial selection (offspring viability) surface associated with genotype of the first 2 loci. But we can work around that.  
2. You will need to input an allele frequency file to initialize your genotype. The first 2 alleles you will make monomorphic.  
3. And then we set mutation to be 0, this is the next assumption.  
  
Then all individuals mating will always have the loci the same. As the offspring disperse on the landscape they will be applied a mortality percentage that you specify and make and convert to ascii and tell input parameters the name of it for the first loci location in the inputvariable file.  
  
This is my quick and dirty answer. It will work, but have to assume a few things.  
Feel free to ask more questions and see user manual for more input variable guidance.  
Cheers,  
Erin.

25 jan 2012

umm, I think CDPOP can’t do what I want it to do. The problem is that dispersal is not a path. Dispersal is disappearing from your original spot and reappearing in another spot that’s a particular distance away. So even if there’s a road in the middle with 100% mortality, it’s unlikely to have any effect because unless you land on the road, you won’t die. You just appear on the other side. You could go back and forth every generation, theoretically. Since the chance that you’d land on the road is small, the population size won’t go down most of the time.

Maybe I’ll have to do HexSim.

26 jan 2012

spent some quality time with CDPOP, and I think I can trick it to do what I want it to do. Since offspring dispersal is really just a point determined by a probability that depends on the location of the parents, there is not spatial path that it takes to get to its destination and a region of high mortality, such as a road, has no impact on it. However, if dispersal is limited and the road is pretty wide in relation to the dispersal function – that is, if it’s set up so that it’s unlikely but possible for an individual to pop up on the other side, then we’re onto something. Not ideal, but it’s something.

What I’ve done so far is to create a scenario with 64 individuals uniformly spread across a landscape that’s 256x256 pixels. I have a selection surface that is uniform, one that has a 10-pixel wide band where survival probability is 0, and one that is about half 0-prob. So far it’s just panmictic, so location means nothing in terms of gene flow. There is less area for individuals to inhabit in the latter 2 scenarios, but they are otherwise the same as the former. Since population size is constrained to remain constant, they just fill in wherever they are.

What I need to do next is create a cost distance matrix from an IBD resistance surface. With that I can actually end up with a central band acting as a road – causing mortality.

14 Feb 2012

Great chat with Erin Languth on the phone just now.

She sees that what I want to do is vary mort and barrier effects independently, and that the program as it is now set up can’t really do that. So she’s developing another element to CDPOP – the road mortality percentage. Migrants from each population would be tracked, and if they end up in their own population they are fine, but if they end up on the other side of the road, they are subject to a mortality rate. Then their offspring arises from the other population and is tracked and the same rules apply to it. What was wrong with the old version is that even if you assign a genotype to all the individuals on one side and have a cost surface that determines the location of mortality, when an individual crosses the road, it would not only be subject to a probability of survival itself, its offspring would most likely be too. So the process would break down after one generation. The new approach would not use genotype or cost surfaces. It would use an extra column on the xy matrix file to determine the locations of the two populations, and then use mortality percentages for what it would cost to enter a new population. She would go even farther in order to make it apply to other applications, where there could be an additional cost for moving even farther away from the source population.

She is also sending a suite of cost surfaces that she used for looking at partial barrier effects on differentiation. Which would save me tons of time. Of course, I need to learn how to do this and would want to play with them and change them, but that’s not an issue for me. I’ll end up changing them quite a bit for sure.

She said she’d work on it right then. Super nice. I owe her for sure. She didn’t accept my offer to end up as coauthor on any work I do but I’ll keep that door open if she wants it. She claimed it would not take very long, so that would be awesome.

9/19/2012

haven’t kept this log up but I’ve done quite a bit.

Sam generated the surfaces for it, which I’m using.

I’m generating CD mats in R package gdistance.

I just ran into an issue where my population would drop below 1000 to a number slightly lower and stay there in all simulations. I think it’s because of how I’m generating the points. Perhaps if they are too close to the edges they aren’t recognized because of the resolution.

9/28/2012

When I allow points to be simulated on top of the road in PAN and IBD, they cause issues – they have an intermediate CD between those on the same vs different side of the road and the pop size drops a certain number, like to 996, and stays there, doesn’t support any individuals. I don’t understand why not, but it doesn’t make sense to have them there, anyway, so I’ve created a buffer just big enough to keep them from being on the road but not enough to make it a ‘gap’ between the two sides.

When the dispersal values are too low on the IBR, I end up with missing individuals some of the time in the corners with bigger high-resistance patches.

I tested r5 vs r10 vs r20, with 10max, 25max, 40max dispersal. They all do fine with 40max, but pop size goes down as dispersal decreases and as patch resistance increases. I think they just don’t get any migrants because it’s too hard to get there.

23 Oct 2012

sGD issue: cannot deal with NAs! Ask Shirk about how to deal with that.

29 Mar 2013

I need to know how plain landscapes behave without a barrier, and how the diversity looks on them.

1. PAN, just to make sure it looks random
2. IBD. What’s up with the circle in the middle?
3. Try out rectangular landscapes
4. Vary dispersal

1 apr 2013

Before I get ahead of myself, I need to:

1. make sure I know what the various raster commands are doing
2. make sure my rasters are what I think they are
3. set rasters to an extent that makes sense

2 apr 2013

1. checked on all that. Good. Confident that my rasters and cd matrices are in good shape
2. just realized that I have way more decimal places than I realized. No need for 9 places! If that becomes an issue I can round them all.
3. Working on CDPOP.
   1. Need to settle on parameter values!
      1. If I’m modeling k-rats, initial dispersal is small, but mating dispersal is large.
      2. I did some comparisons before. Take some of those results
   2. Need to settle on n generations!
      1. Last time I was finding that 200-250 was enough to get heterozygosity to equilibrium.
      2. Het takes longer to get to equilibrium than mantel corr.