Guidelines to Select Grid Resolution

Josh Cullen

17 March 2020

Background

During all previous analyses of snail kite space-use, these results were entirely based upon a discrete spatial domain where the grid resolution was subjectively determined. Ideally, we wanted a grid resolution that was relatively high, but also resulted in many long time segments of consistent grid cell use. These two goals are at relative odds with one another, which will likely require the user to find a balance of both of these properties. Additionally, the establishment of guidelines to determine a range of grid resolutions that balance number of occupied cells and duration of time spent in grid cells would increase reproducibility.

There are a few recent studies that evaluate ecological questions over a gridded spatial domain, each of which have different sets of limitations. Hanks et al. (2015) developed a continuous-time discrete-space mechanistic movement model and defined the grid resolution according to the rasters of available environmental covariates. While useful for the principles of their model, our method does not necessarily rely on evironmental covariate rasters to segment the time series of grid cell use. Bastille-Rousseau et al. (2018) treated the grid resolution of their connectivity study as an emergent property of step length, where they defined it specifically as the median of all step lengths. While including emergent properties of animal movement allow greater generalization for the analysis of multiple species with different movement patterns, the median step length does not necessarily help strike a balance between grid cell number and time period duration. Another study by Pasquaretta et al. (2020) evaluated network connectivity in many different organisms and estimated grid resolution by species using a multi-step process. This involved characterizing the sequence of grid cells occupied into sets of network motifs and then comparing motif sequences from different grid resolutions using a Dynamic Time Warping algorithm. Last, the authors calculate the Shannon diveristy index to determine the grid resolution that produces the most unique grid cells used (species abundance) that are used with comparable frequency (species evenness). This method also relies on step length as an emergent property of animal movement, but requires more extensive analyses than should be necessary to inform the initial step of discretizing space.

Instead of using any of these three aforementioned methods, I developed a method that directly evaluates the trade-off between grid cell number and the duration spent before moving to a new grid cell. Below I will show the steps used to estimate a suitable range of grid resolutions using the snail kite data as an example.

Snail Kite Case Study

First, the user must generate a large sequence of quantiles (~100) from animal displacement, which is the distance from each relocation to the initial tagging location. Therefore, this method treats the displacement, instead of step length, as the emergent property of animal movement. Displacement was used in place of step length since territorial species (e.g. armadillo, snow leopard) appear to exhibit step lengths whose values would not be large enough to provide a suitable trade-off between the number of grid cells and the time spent in a given grid cell. For snail kites, a sequence of 120 values from the 20th quantile (0.25 km) to

the 50th quantile (33 km) were generated on which to measure properties of their total number of occupied grid cells and the duration spent within a grid cell before moving to another.

```
#generate quantiles for displacement
quant<- seq(0.2, 0.5, length.out = 120)

#run function that measures median number of grid cells used and median time spent per
#grid cell across all individuals
quant.range<- res.function(quant = quant, dat.spdf = dat.spdf, crs = proj4string(dat.spdf))
```

Median values for each of these measures were estimated instead of the mean to remove the bias of extreme values; these were also estimated across all individuals of the dataset. Results from this analysis were visualized graphically to ease interpretation. Axes were standardized (0 to 1) to equally weight the impact of both variables. Additionally, a distance measure from each point was taken compared to the origin (0,0) and the grid resolutions that produced the 10 smallest distances were recommended as the range of possible grid resolutions to test. However, it is worth noting that the segmentation model may still work even with grid resolution outside of this range. This is meant to serve as a starting point for testing a grid resolution where it is also recommended to select the finest resolution that produces the best trade-off.

```
#standardize number of grid cells used and duration (length) of time per grid cell
quant.range[,2]<- quant.range[,2]/max(quant.range[,2])
quant.range[,3]<- quant.range[,3]/max(quant.range[,3])

#measure distance from point (n cells, n length) to origin (0,0)
quant.range<- quant.range %>%
  data.frame() %>%
  mutate(dist = sqrt((n.cells-0)^2 + (n.length-0)^2))

quant.range2<- quant.range %>% top_n(n=-10, wt=dist)   #12.7 - 28.3 km
quant.range2
```

```
##          res     n.cells  n.length      dist
## 1   12677.54 0.16393443 0.3333333 0.3714641
## 2   15733.56 0.16393443 0.3333333 0.3714641
## 3   15862.42 0.16393443 0.3333333 0.3714641
## 4   16026.17 0.15573770 0.3333333 0.3679203
## 5   16053.14 0.15573770 0.3333333 0.3679203
## 6   16097.56 0.16393443 0.3333333 0.3714641
## 7   16133.83 0.15573770 0.3333333 0.3679203
## 8   16150.95 0.16393443 0.3333333 0.3714641
## 9   16161.41 0.16393443 0.3333333 0.3714641
## 10  16185.78 0.16393443 0.3333333 0.3714641
## 11  16366.11 0.16393443 0.3333333 0.3714641
## 12  16394.35 0.16393443 0.3333333 0.3714641
## 13  16420.84 0.16393443 0.3333333 0.3714641
## 14  16462.44 0.16393443 0.3333333 0.3714641
## 15  17348.39 0.15573770 0.3333333 0.3679203
## 16  20185.35 0.13114754 0.3333333 0.3582050
## 17  28316.13 0.09836066 0.3333333 0.3475427
```
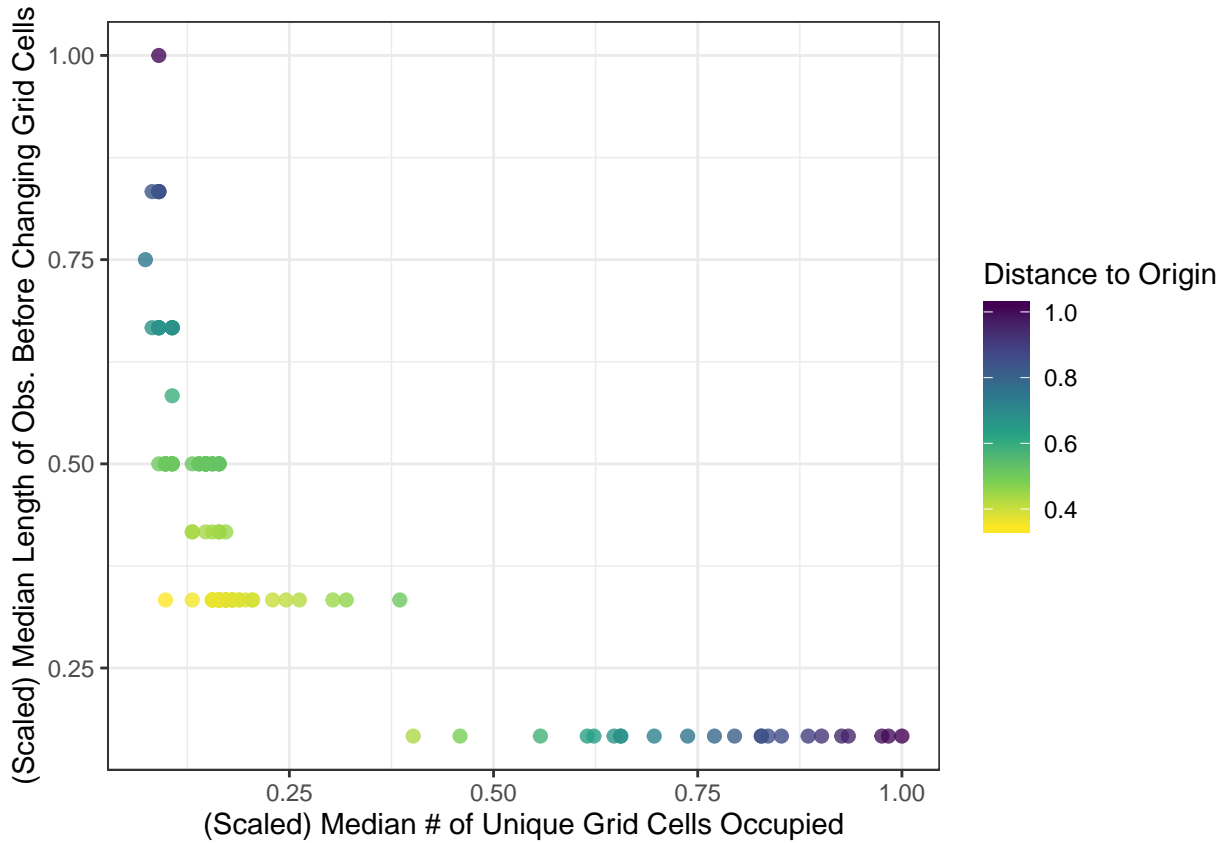
Figure 1: Points (defined by grid resolutions) that are closest to the origin and therefore have the smallest distance values represent the recommended grid resolutions with which to analyze the data.

In the printed table above, the `res` column denotes the grid resolution, which is shown in meters. As determined from this analysis on the snail kite data, it appears that 12.7 - 28.3 km is the range of recommended grid resolutions. Since we would preferably like to analyze the data with the finest resolution possible, I will evaluate the snail kite data further using a grid resolution of 15 km since it is a round number on the low end of this range. Below is an example of what this grid looks like plotted over geographic space with the raw observations of all 27 snail kites.
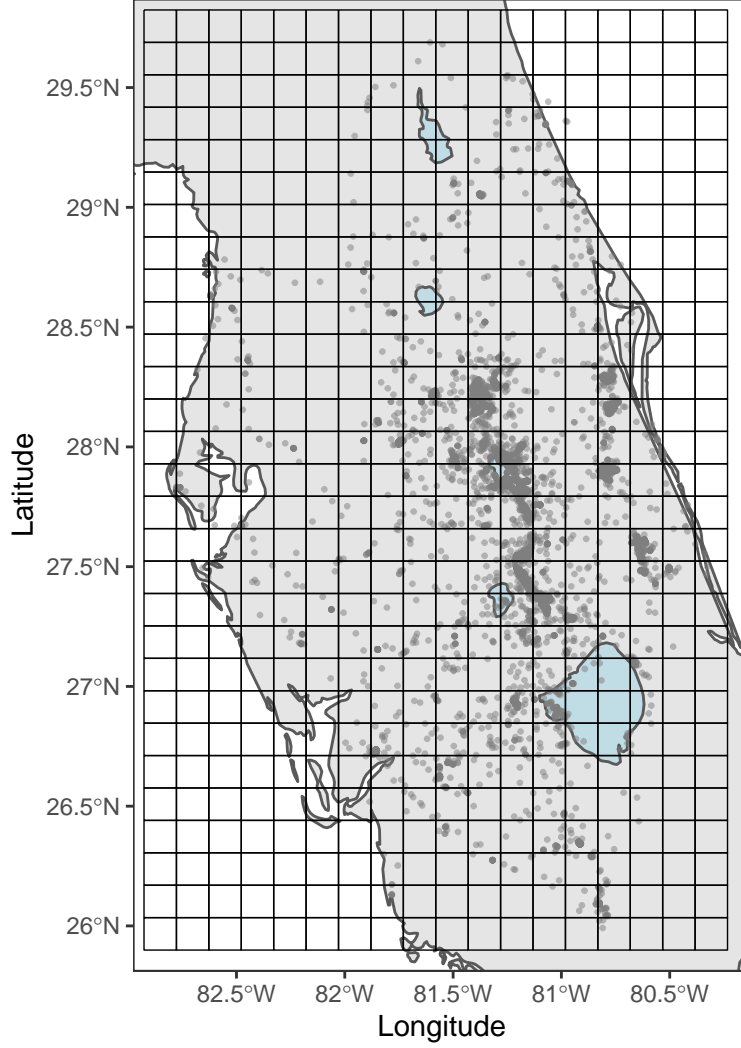
Figure 2: Snail kite observations plotted over a grid with 15 km resolution.

Since there are 27 snail kites, I will focus on the segmentation results of a single individual that is representative of most other tagged individuals (**SNIK 12**). Shown below is the heatmap showing grid cell use over time for **SNIK 12** with green vertical lines denoting breakpoints. *While it needs to be evaluated further, I have also analyzed this dataset setting $\alpha = 1$ for the prior where before it was 0.01. This results in fewer breakpoints, which benefited the interpretation during the analysis of the snow leopard data and may want to be used here as well for consistency. Alternatively, different species, movement patterns, and/or sample sizes may require different values for α which can be explored in subsequent analyses.*
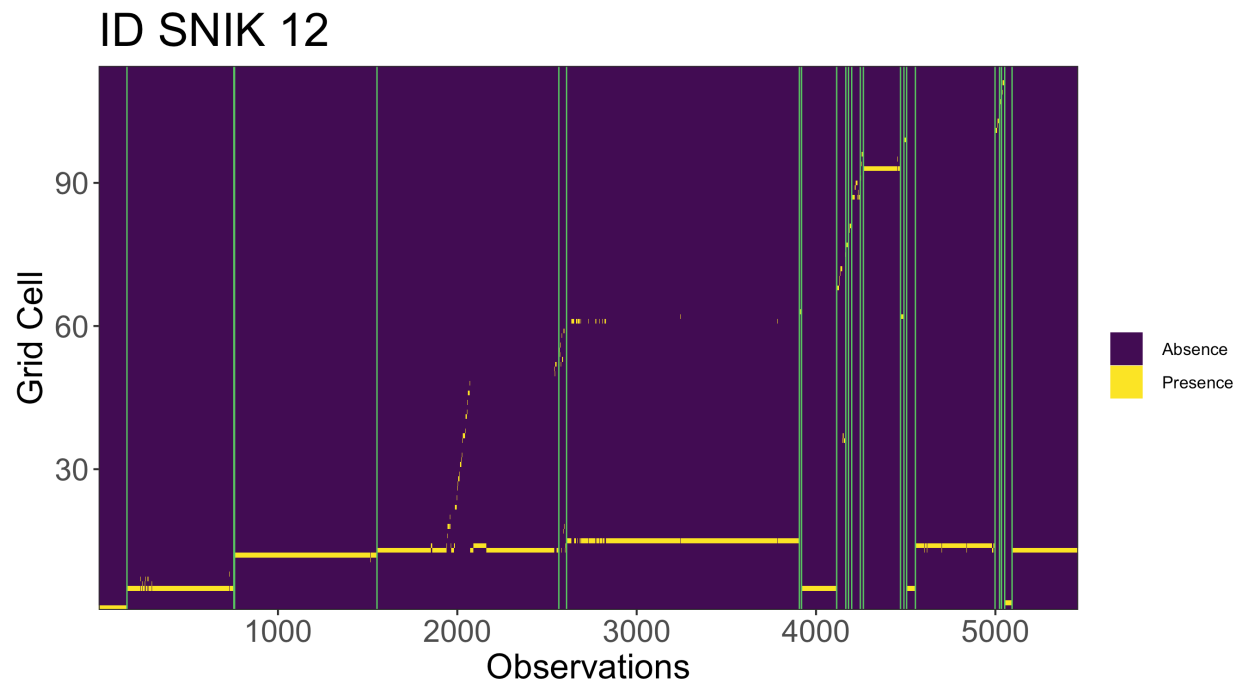
Figure 3: Time segmentation results using a resolution of 15 km.

This segmentation model detected 23 breakpoints for **SNIK 12**, which was much fewer than the 81 breakpoints previously estimated for this individual when using a grid resolution of 5 km and setting $\alpha = 0.01$. However, the time segments here are much longer and homogeneous than in the prior analysis. Next, I will show the locations of 20 modeled activity centers (ACs) from all 27 snail kites using this grid resolution.
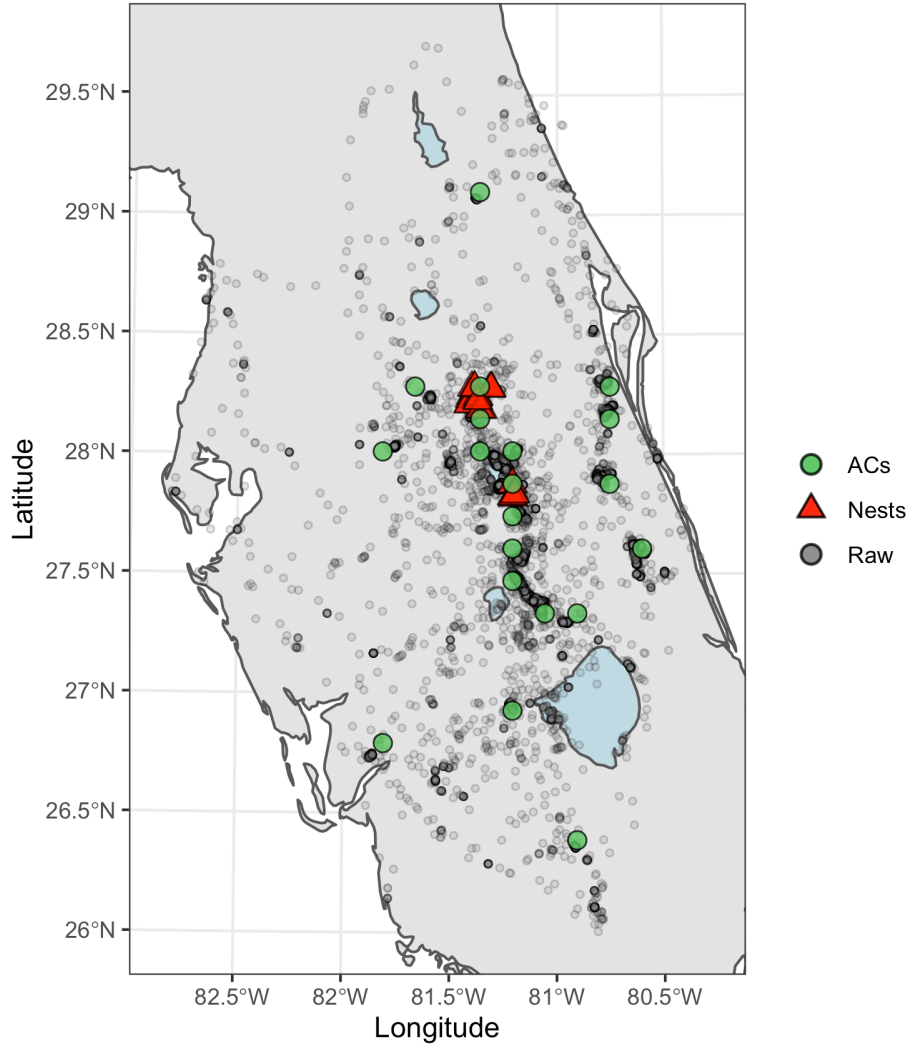
Figure 4: Raw observations from all 27 snail kites are shown in relation to the nest sites and the AC estimated by the model.

The ACs mapped in this figure appear to not overlap locations of high point density as well as the model using a 5 km grid resolution. Additionally, the coarser resolution has resulted in fewer locations along the Kissimmmee River watershed since these are covered by fewer grid cells at a 15 km resolution. A similar result is seen for the number of ACs estimated to overlap with nest locations, as there are three fewer ACs that overlap these sites and they do not overlap as well. While the mixture model ran much quicker using this data compared to the 5 km resolution analysis (2 min vs 25 min), the resulting AC locations appear to be poor by comparison. This may require the use of the other model that estimates ACs over continuous space. Additionally, the user could test increasingly finer grid resolutions to find something more suitable.