APPLICATIONS OF MARKOV CHAINS

ANIMAL TRACKING

BY GRACE CHANG

ANIMAL TRACKING & PAPER

INFERRING ANIMAL DENSITIES FROM TRACKING DATA USING MARKOV CHAINS

- Study of animal distributions and relative densities common in ecology
- Primarily electronic tags (most popular)
- Tagging used to detect movement, habitat use, foraging patterns, behavior, social structure, life history, response to changes

THE PROBLEM

- Track data is usually biased
- Locations of tracked animals are biased of release location ("start-bias')
- Currently no bias removal technique
- Learn less about animals and habitat
- Suffer of management and conservation efforts

THE PREMISE

- Envisage tracks as Markov chain (animal move between areas)
- Markov Chain method gives unbiased measures of relative density
  - done over geographical grid, water depth, or habitat type

THE MODEL: ASSUMPTIONS

1. Animals tracked = random subset of population in term of movement behavior

2. Movements form a time-homogenous Markov chain

    - Probability an animal moves from i to j at time p is independent of time

THE MODEL: THEORY

$$
\begin{aligned}
\mathbb{P}\left(\text{Animal in cell } i\right) &= \pi_i \\
&= \sum_j \mathbb{P}\left(\text{In cell } j\right) \mathbb{P}\left(\text{Move from } j \text{ to } i\right) \\
&= \sum_j \pi_j p_{ji}
\end{aligned}
$$

THE MODEL: THEORY

- Assumption 2: Movements form time homogenous Markov chain
    - $\pi = \{\pi_j\}$
    - $P = \{p_{ji}\}$
- Which gives us $\pi = \pi P$
- π = left eigenvector of P associated with eigenvalue 1

THE MODEL: THEORY

$$\pi = \pi P$$

- π = unbiased estimate of proportion of times animals spent in one area
  - Assumptions 1 and 2
  - π = proportion of population in different areas

$$n_{ji} = \text{number of steps starting in j and ending in i}$$

$$\hat{p_{ji}} = n_{ji} / \sum_j n_{ji}$$

THE MODEL: THEORY

$$\pi = \pi P$$

$$\hat{p_{ji}} = n_{ji} / \sum_j n_{ji}$$

- with the left eigenvector corresponding to the eigenvalue of 1
- estimate relative densities from corresponding elements
- normalize eigenvector so it sums to 1

METHODS OF THE PAPER

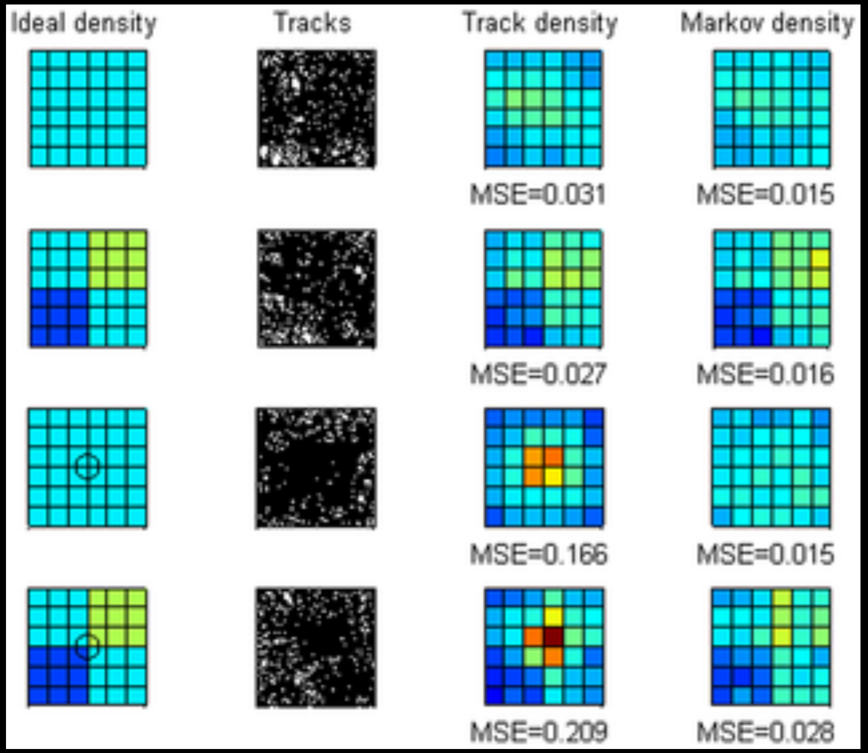
Study Areas:

1. Homogenous: Moves length 0.05
2. “Quadrants” moves different

0.05	$\frac{0.05}{\sqrt{1.5}}$
$\frac{0.05}{\sqrt{0.5}}$	0.05

Tracking Start

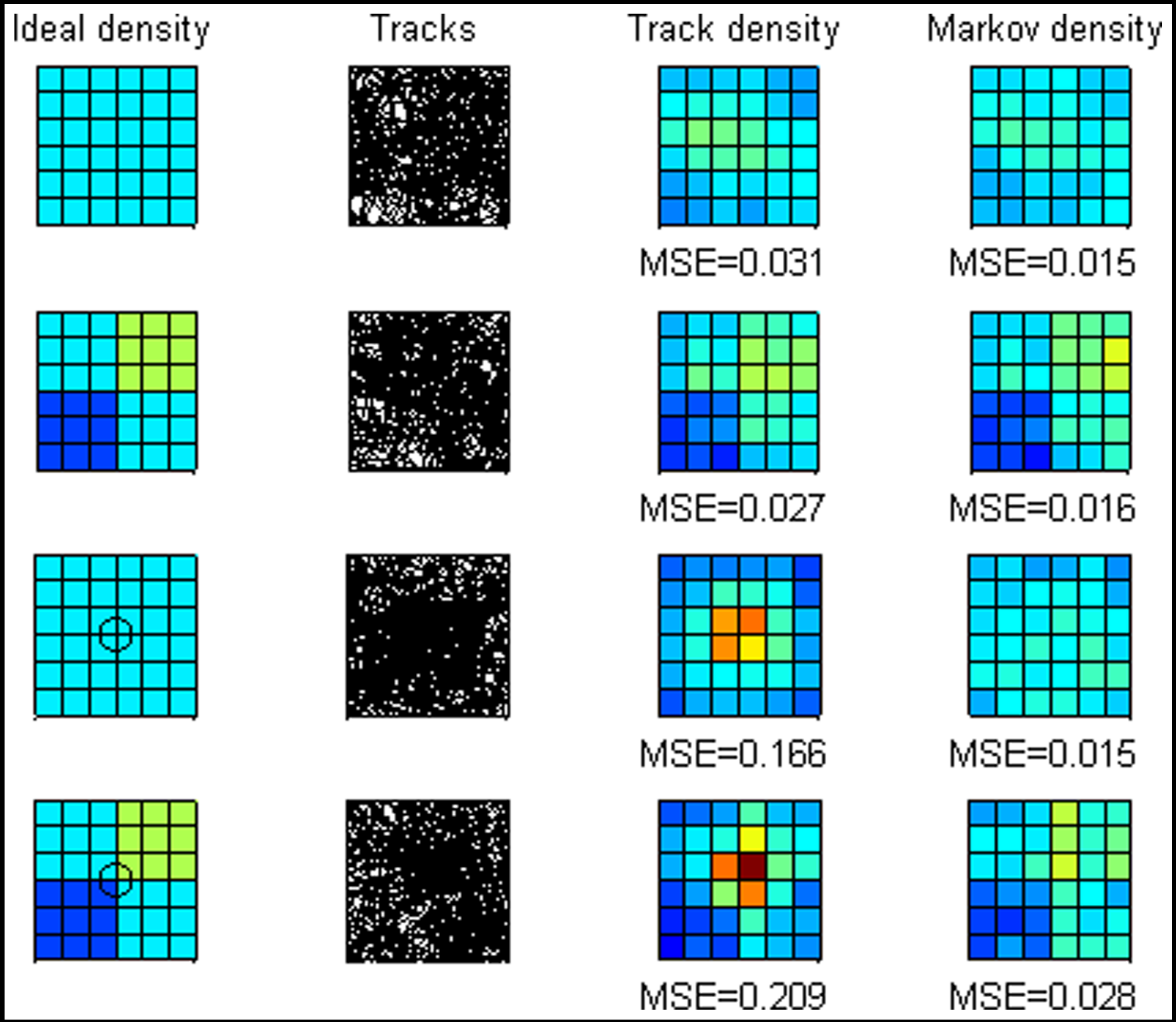
1. Randomly chosen position
2. Randomly chosen position within 0.1 units of the center [0.5,0.5]



DETAILS OF METHODS

- 36 Square cells, 1 Exterior cell for outside study area
- Track Density: Relative density from summing number of track positions in each cell
- Markov Density: Density using eigenvalue model
- 300 Agents, 1000 Moves

THE EXPERIMENT: RESULTS



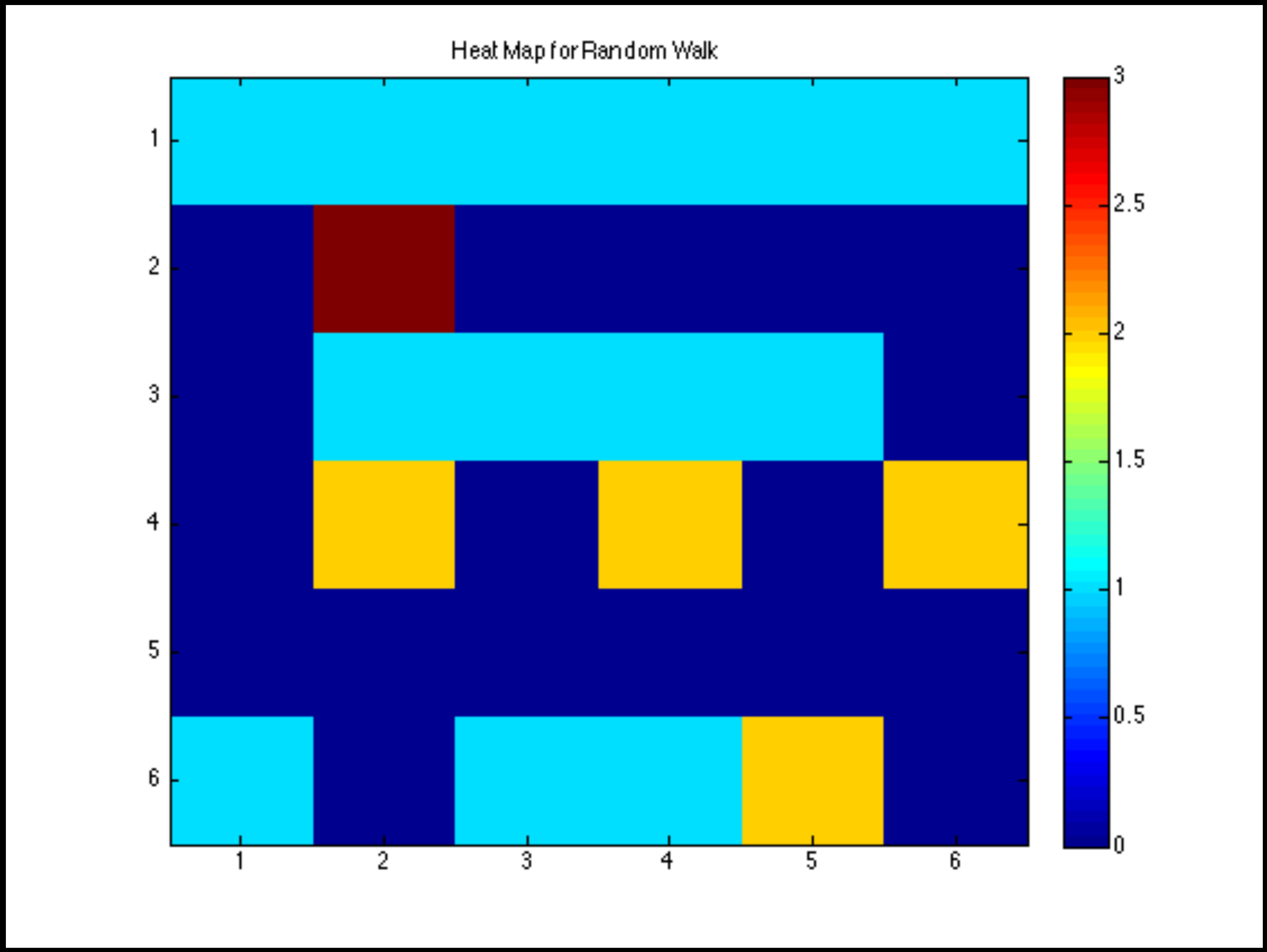
THE EXPERIMENT: MY TRACKS

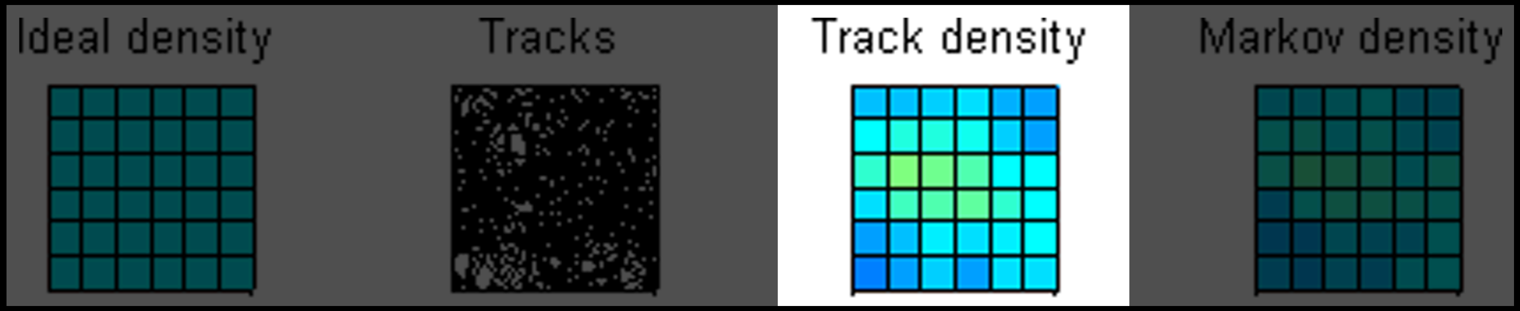
1. Create matrix: 2 columns, # agents row
2. Generate random starting point
3. Input final position in matrix
    - Roll to move x-direction or y-direction
    - Roll to move forward or backward
4. Reassign coordinates to grid number (1-36)
    - Created empty 6x6 matrix
    - Loop across agents, rows, and columns
    - If fulfilled both those conditions, then would +1 to 6x6 matrix (sum them)

$$\frac{1}{6}(\text{column number}) - \frac{1}{6} \leq \text{x-coordinate} < \frac{1}{6}(\text{column number})$$

$$\frac{1}{6}(\text{row number}) - \frac{1}{6} \leq \text{y-coordinate} < \frac{1}{6}(\text{row number})$$

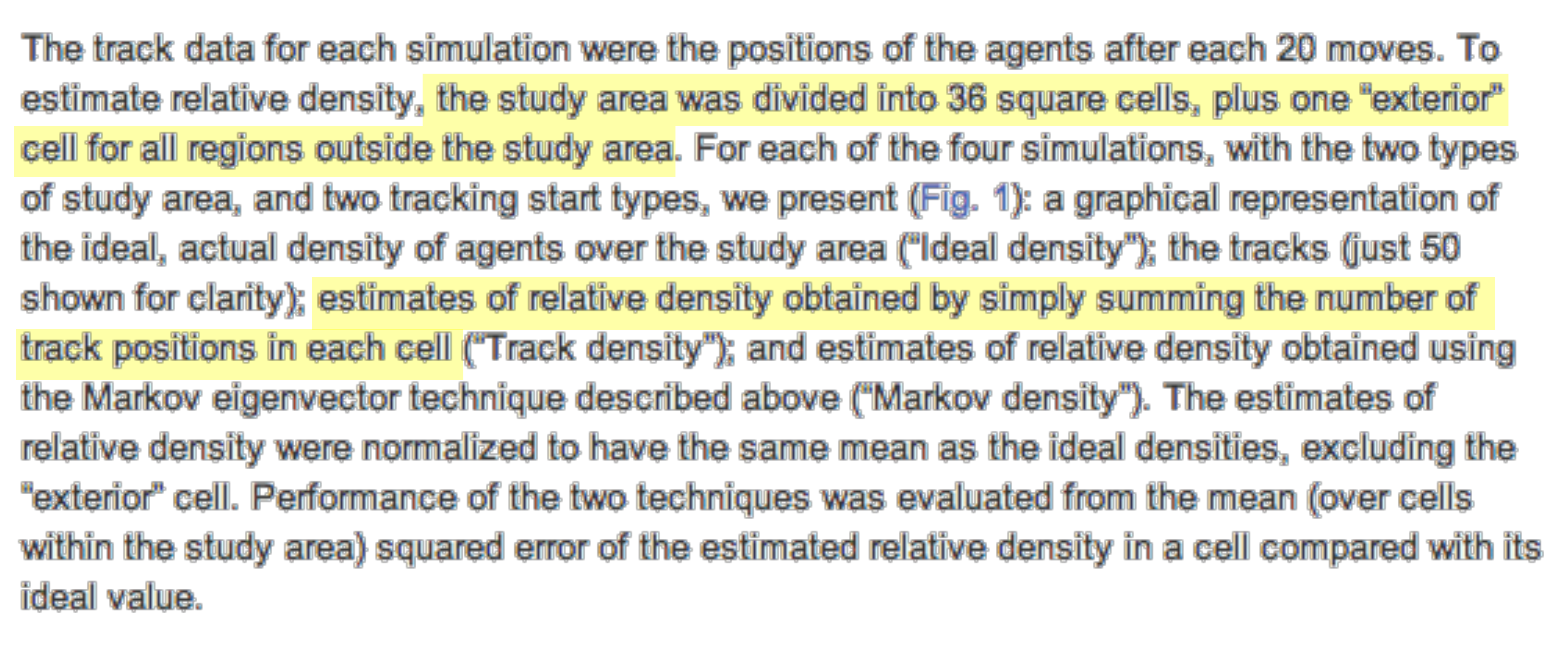
THE EXPERIMENT: MY TRACKS



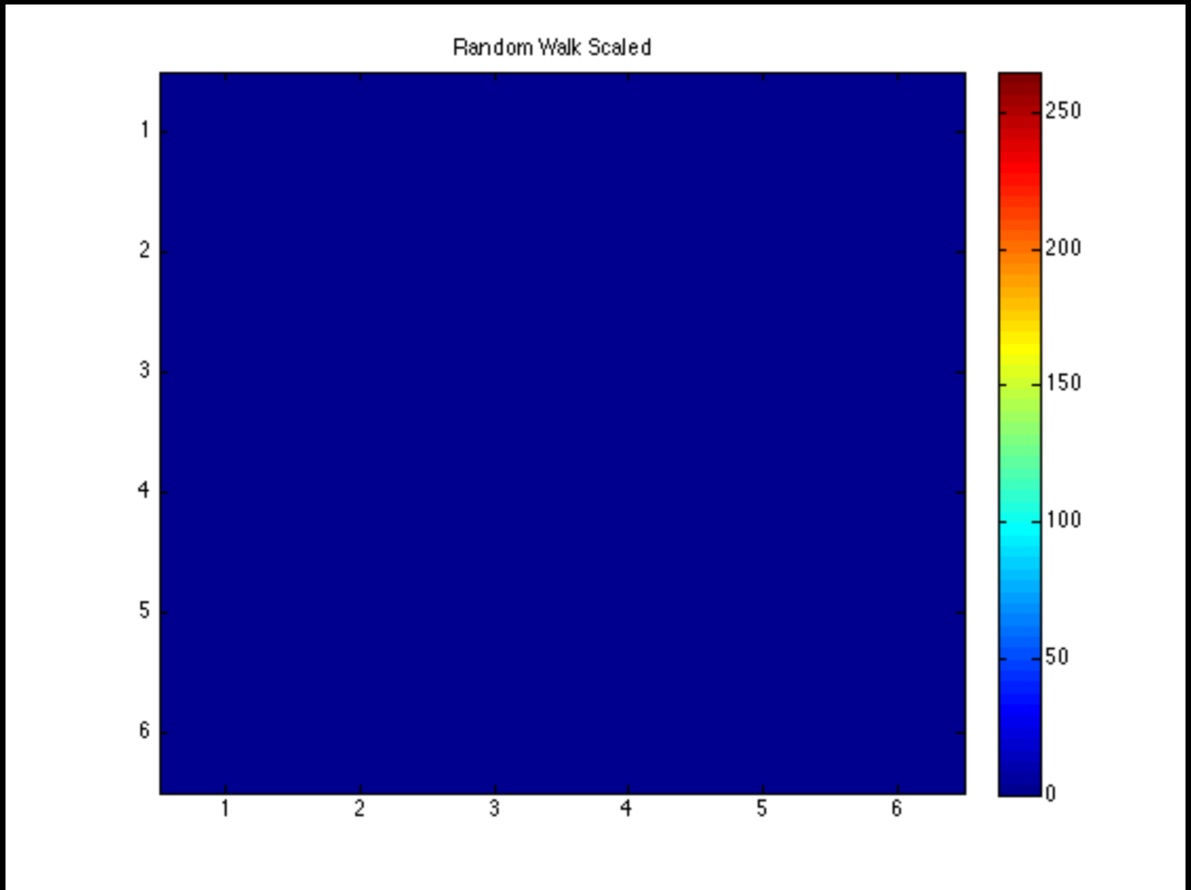


- Small range
- Generally "even'
- Integers make heat map look jumpy

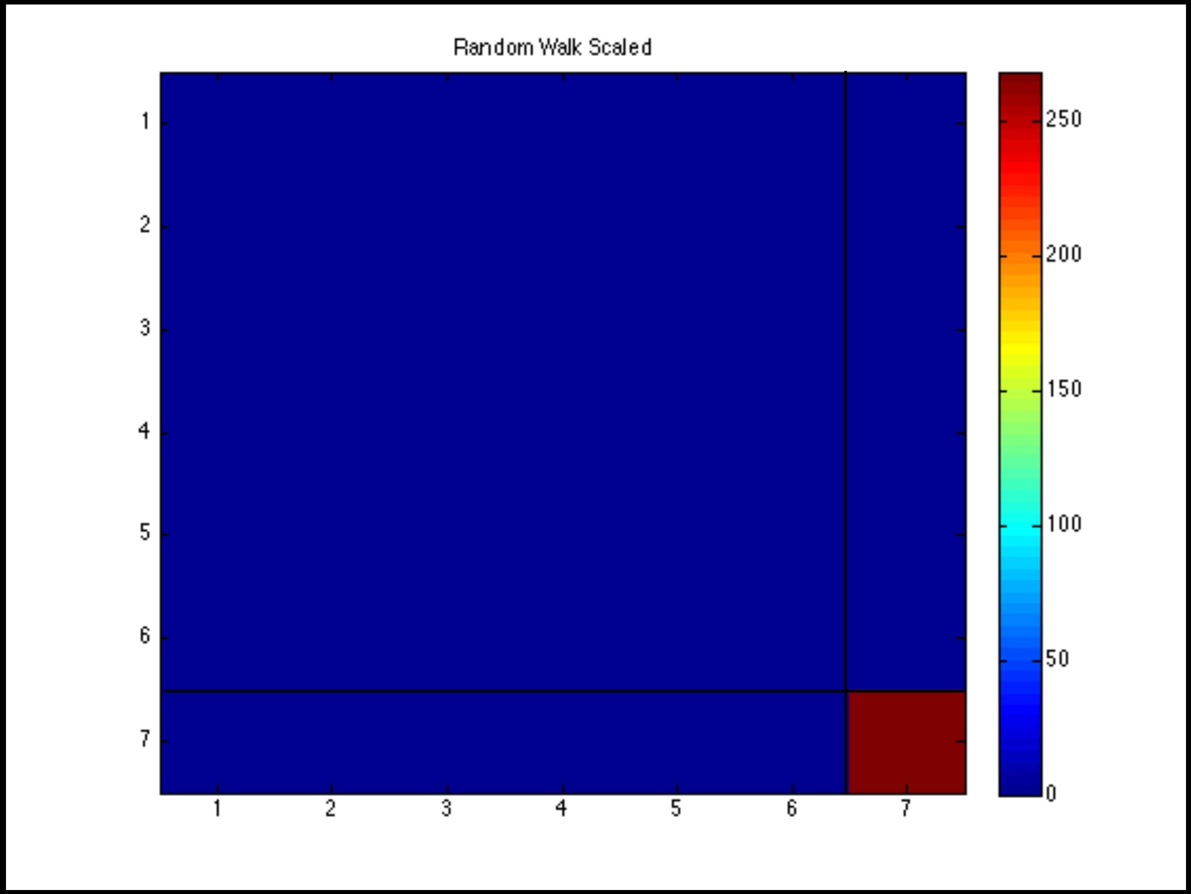
PAPER CLARIFICATIONS


The track data for each simulation were the positions of the agents after each 20 moves. To estimate relative density, the study area was divided into 36 square cells, plus one "exterior" cell for all regions outside the study area. For each of the four simulations, with the two types of study area, and two tracking start types, we present (Fig. 1): a graphical representation of the ideal, actual density of agents over the study area ("Ideal density"); the tracks (just 50 shown for clarity); estimates of relative density obtained by simply summing the number of track positions in each cell ("Track density"); and estimates of relative density obtained using the Markov eigenvector technique described above ("Markov density"). The estimates of relative density were normalized to have the same mean as the ideal densities, excluding the "exterior" cell. Performance of the two techniques was evaluated from the mean (over cells within the study area) squared error of the estimated relative density in a cell compared with its ideal value.

RANDOM WALK: RANDOMNESS

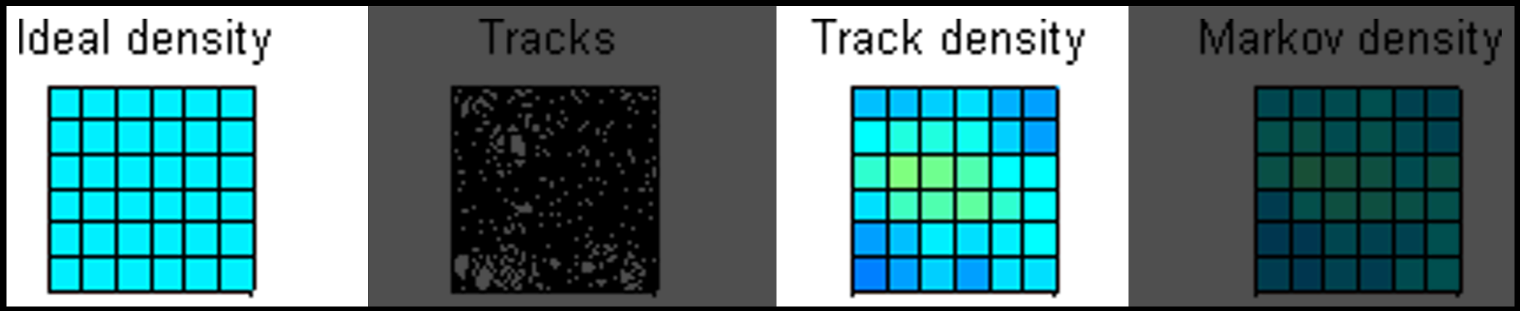


OUTSIDE CELL NOT SHOWN

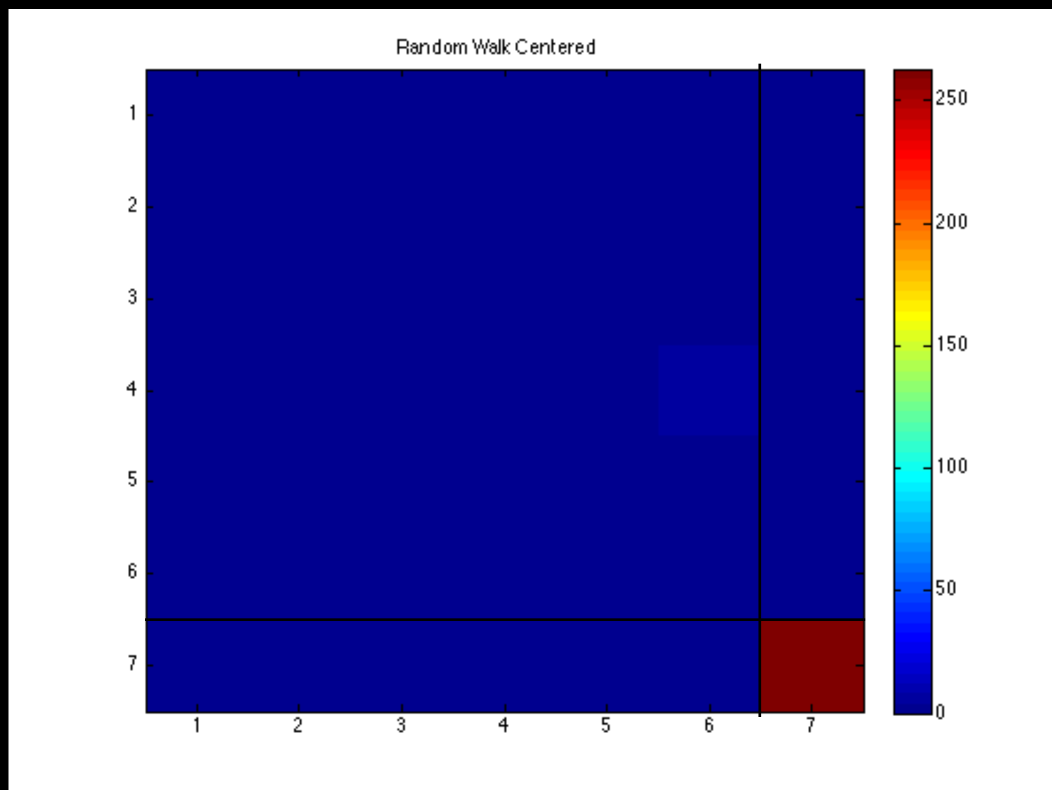


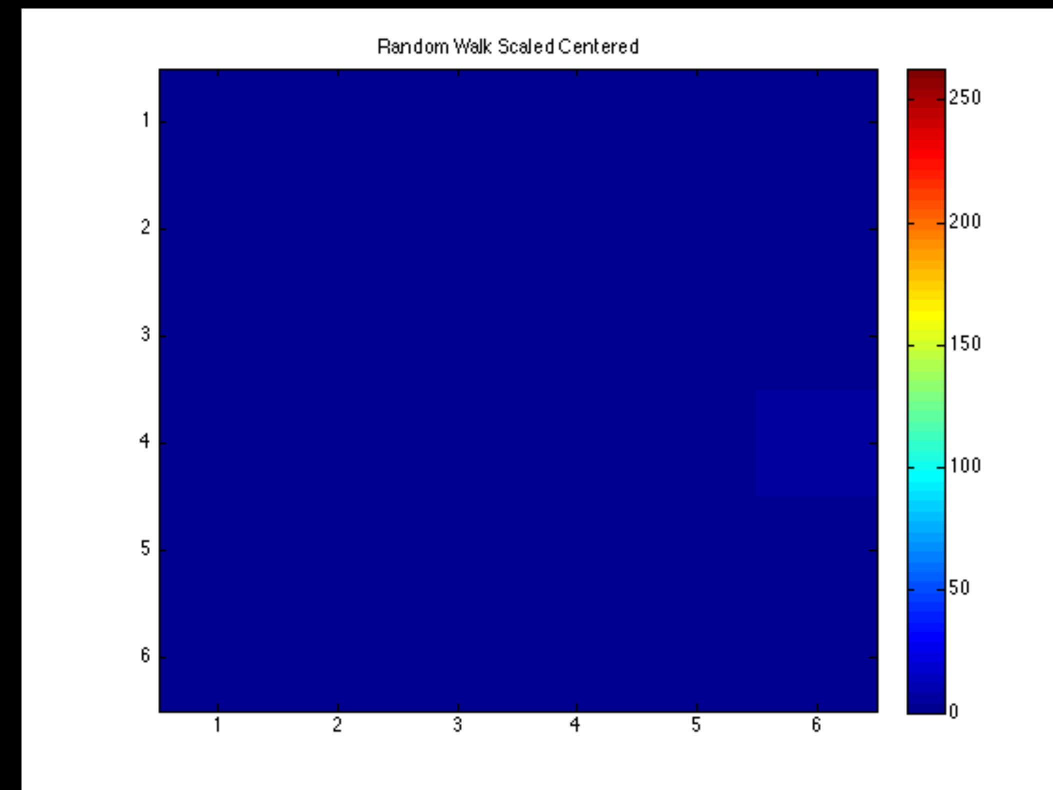
OUTSIDE CELL SHOWN

CELL WEIGHT ISSUE?
NO COLORBAR PROVIDED.



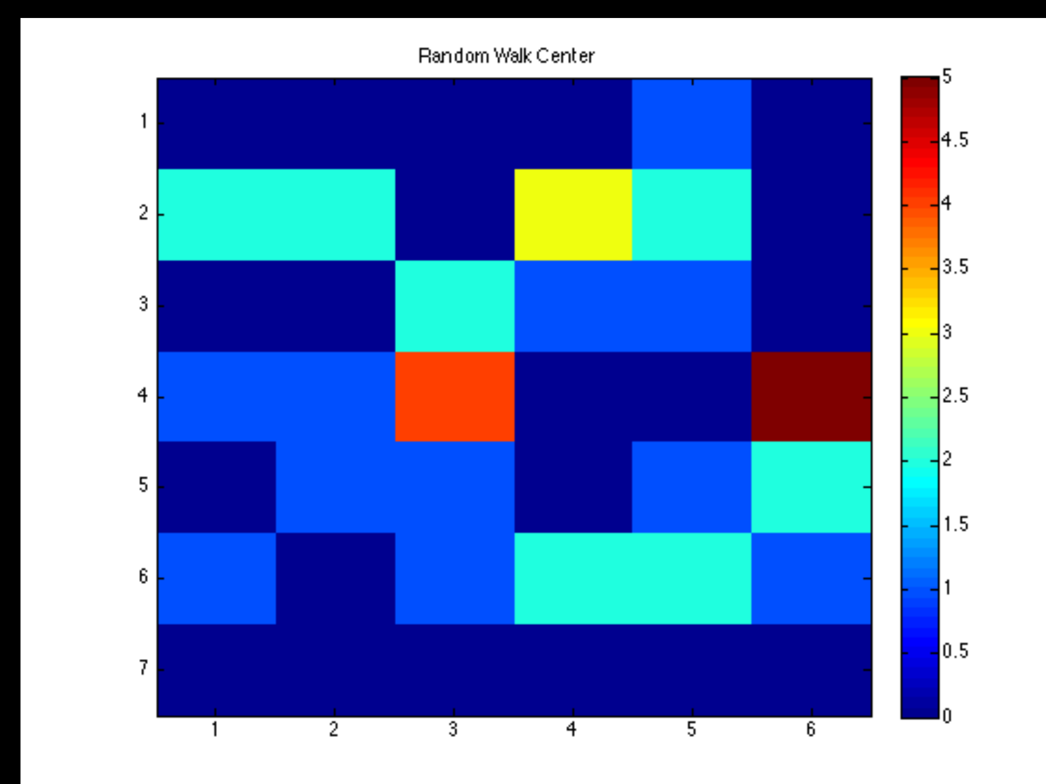
RANDOM WALK: CENTERED

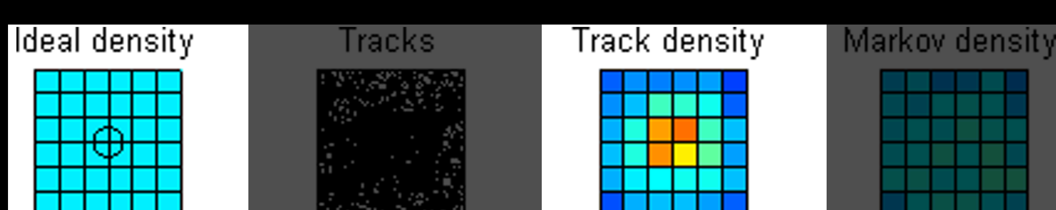




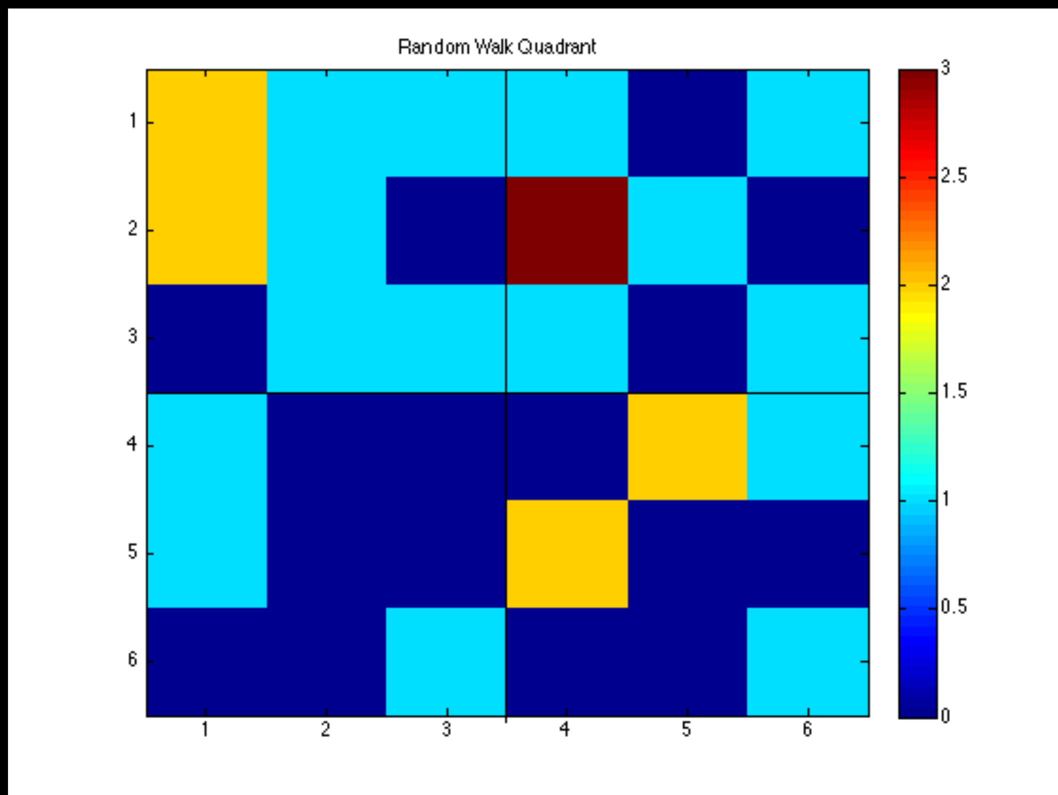
grid =

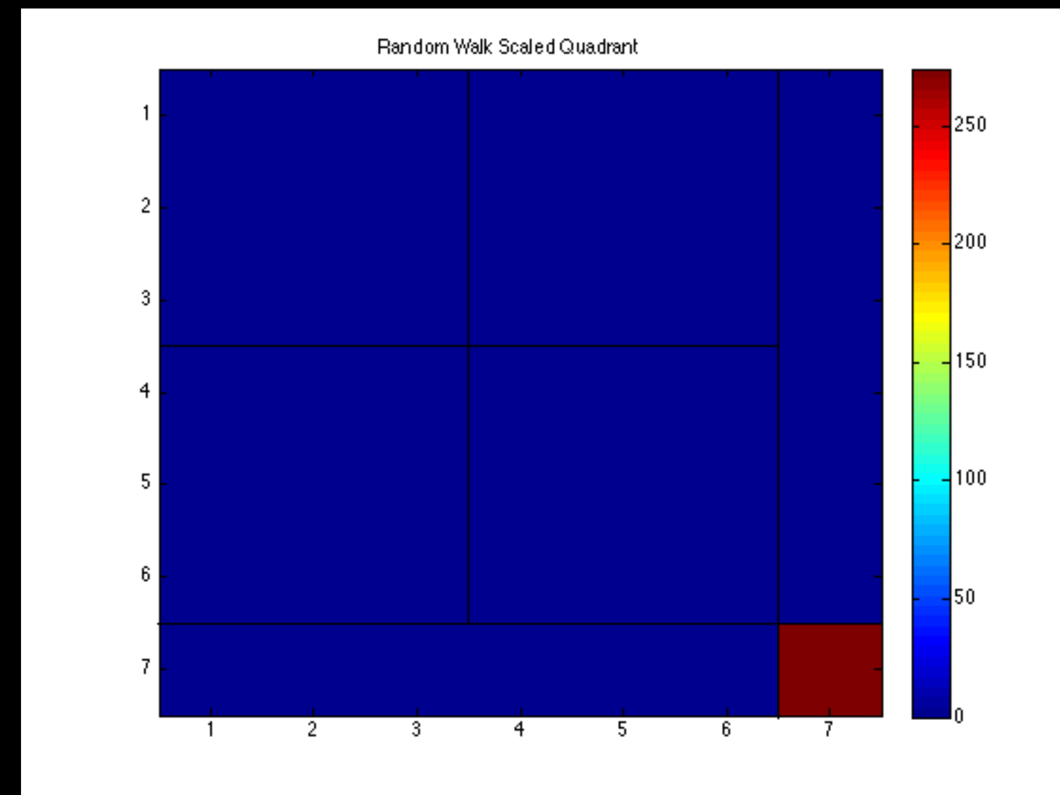
0	0	0	0	1	0	0
2	2	0	3	2	0	0
0	0	2	1	1	0	0
1	1	4	0	0	5	0
0	1	1	0	1	2	0
1	0	1	2	2	1	0
0	0	0	0	0	0	263





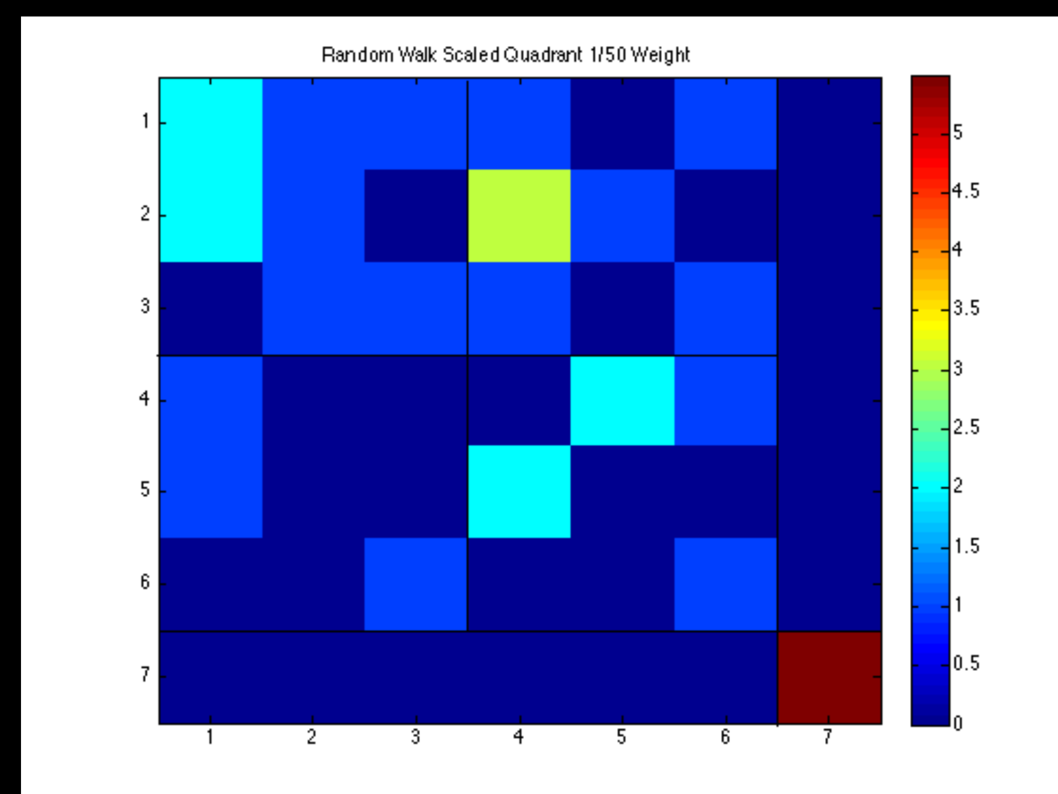
RANDOM WALK: QUADRANT

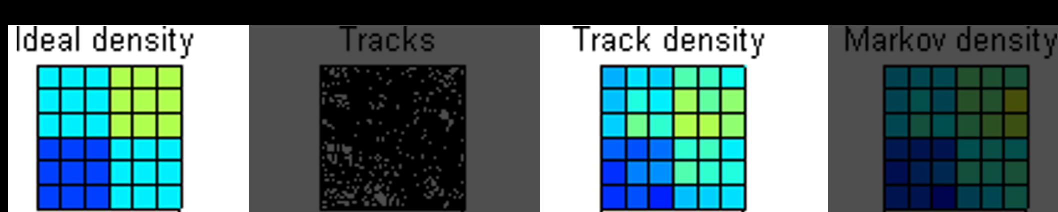




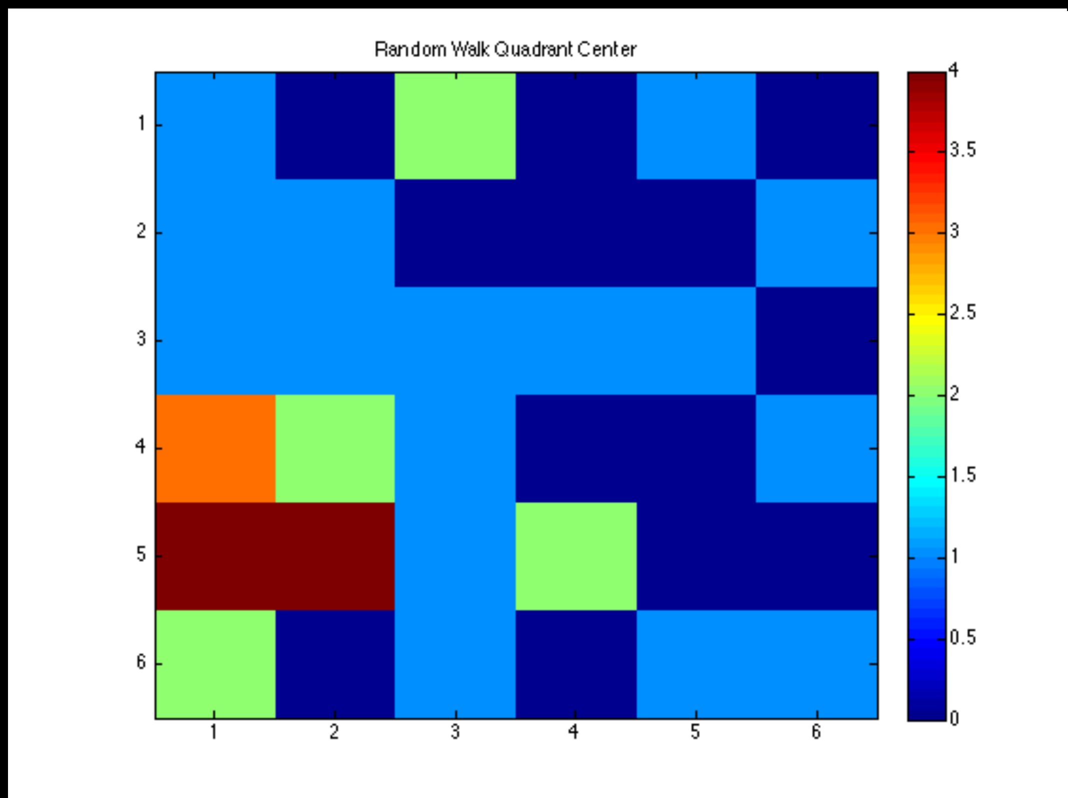
grid =

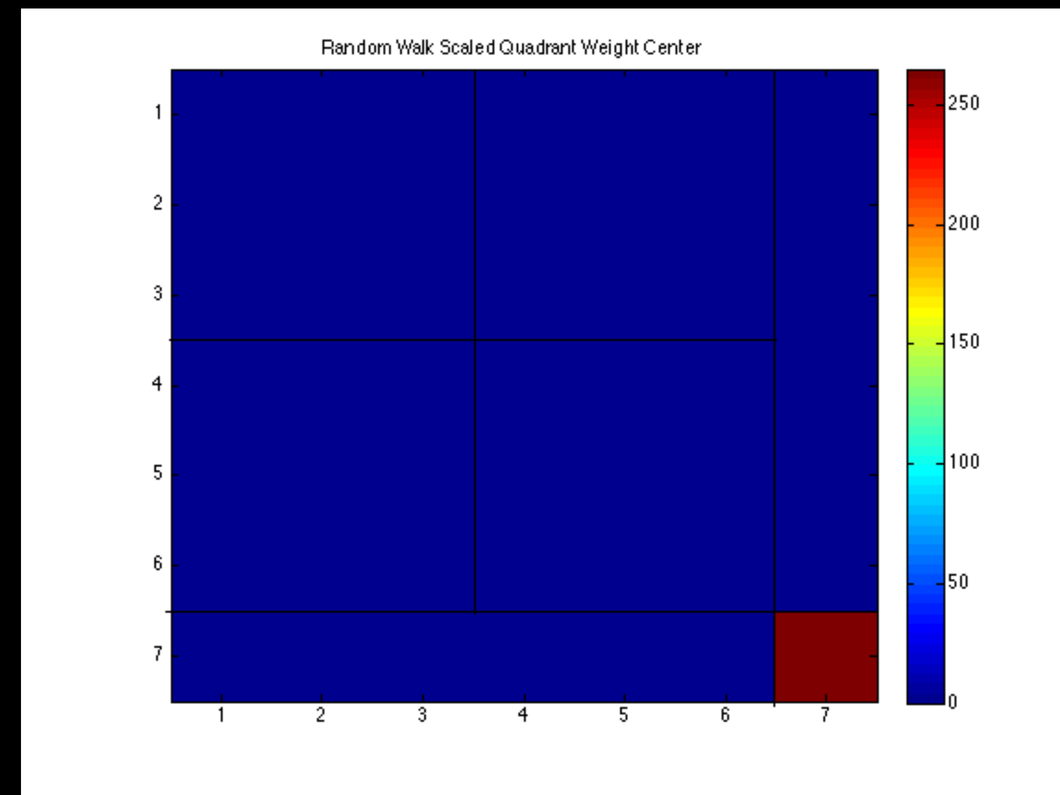
2	1	1	1	0	1	0
2	1	0	3	1	0	0
0	1	1	1	0	1	0
1	0	0	0	2	1	0
1	0	0	2	0	0	0
0	0	1	0	0	1	0
0	0	0	0	0	0	274





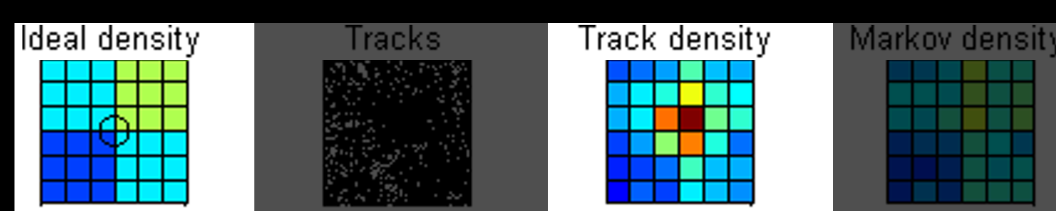
RANDOM WALK: QUADRANT CENTER






```
grid =

     1     0     2     0     1     0     0
     1     1     0     0     0     1     0
     1     1     1     1     1     0     0
     3     2     1     0     0     1     0
     4     4     1     2     0     0     0
     2     0     1     0     1     1     0
     0     0     0     0     0     0   265
```

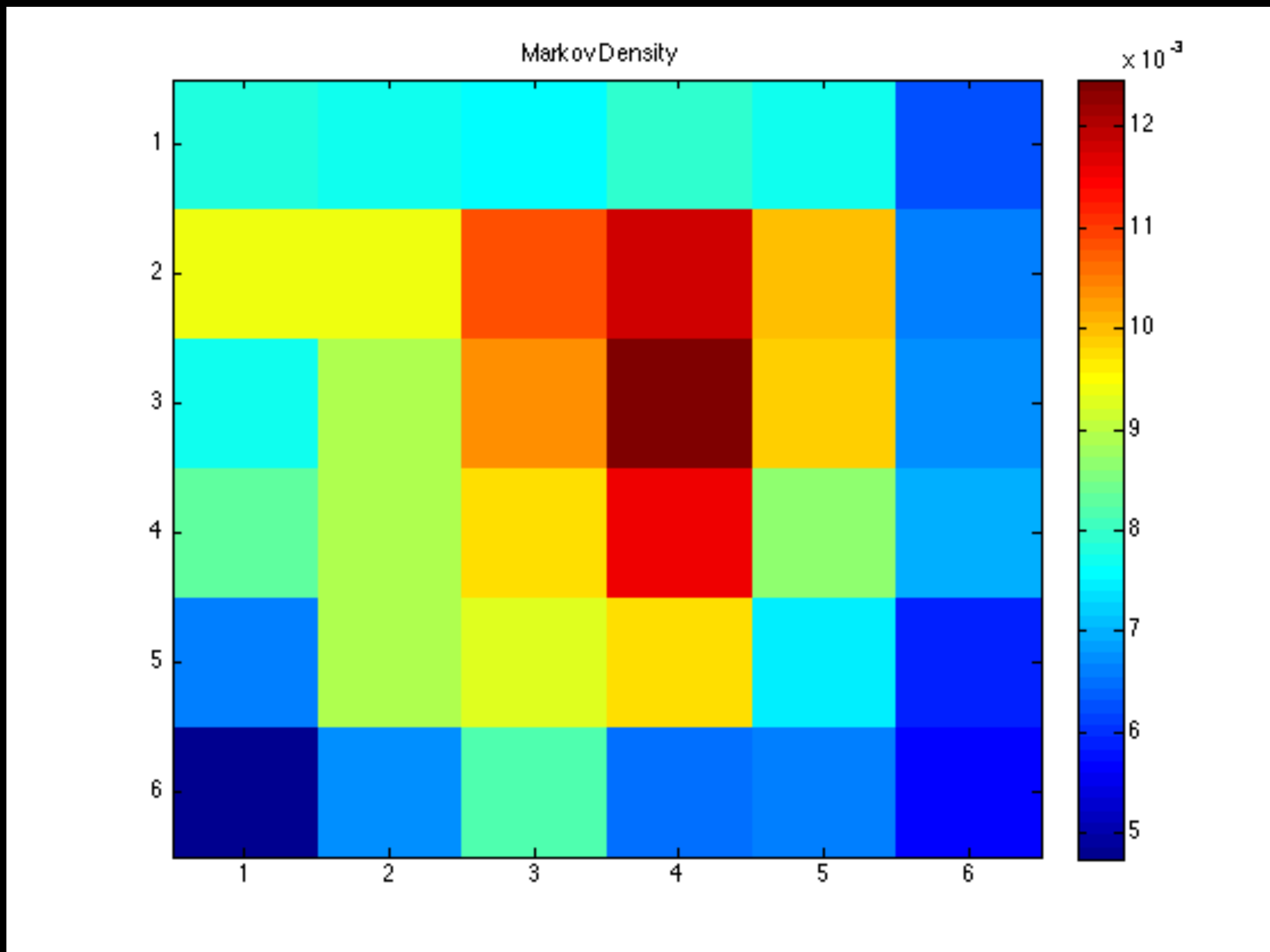


MARKOV STEPS

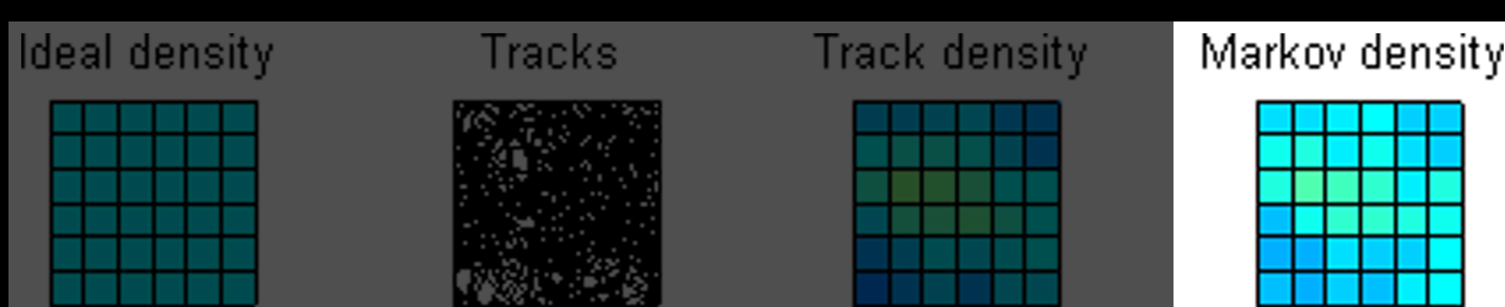
1. Create matrix: 2 columns, # agents row
2. Generate random starting point
3. Input final position in matrix
    - Roll to move x-direction or y-direction
    - Roll to move forward or backward
    - Created count matrix each time move into cell
4. Calculate p_ij
    - Summed count matrix for denominator
    - Divided count matrix by totals
5. Calculate Eigenvalue for Final Density
    - Calculate eigenvalue for p_{ji}
        - script to find the eigenvector corresponding to eigenvalue 1
    - Normalized eigenvector
6. Split into grid

$$\hat{p_{ji}} = n_{ji} / \sum_j n_{ji}$$

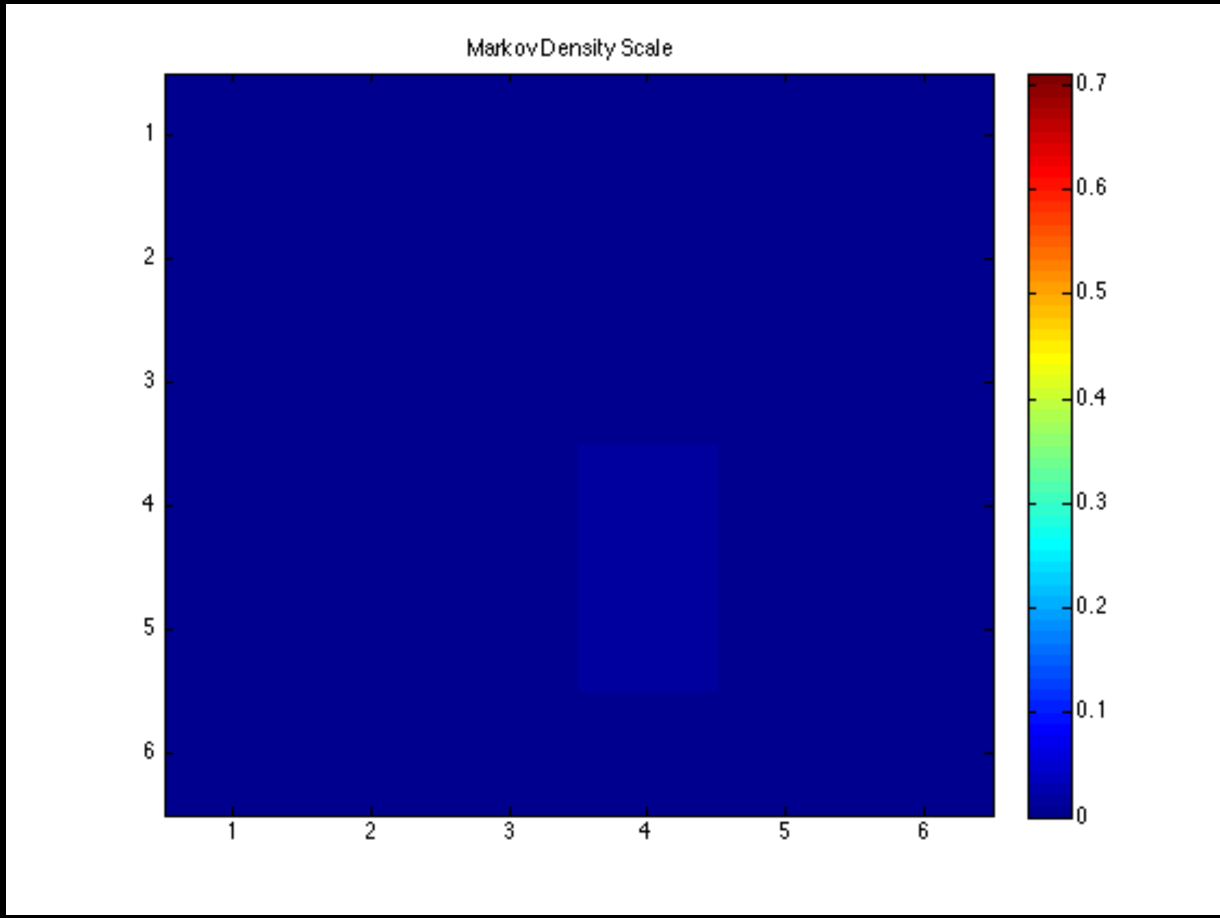
MARKOV: RANDOM



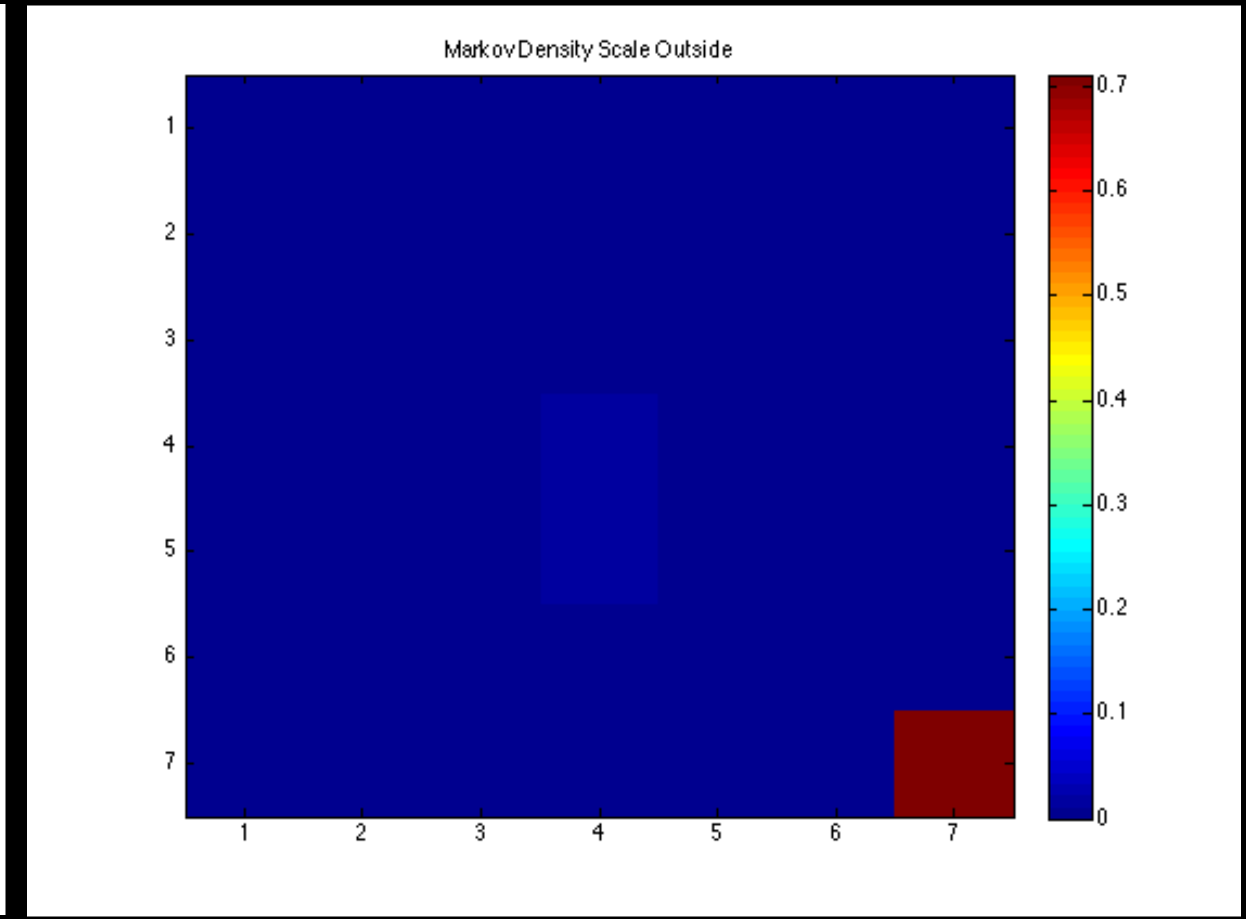
- Small range
- Continuous numbers make it more smooth
- scale difference



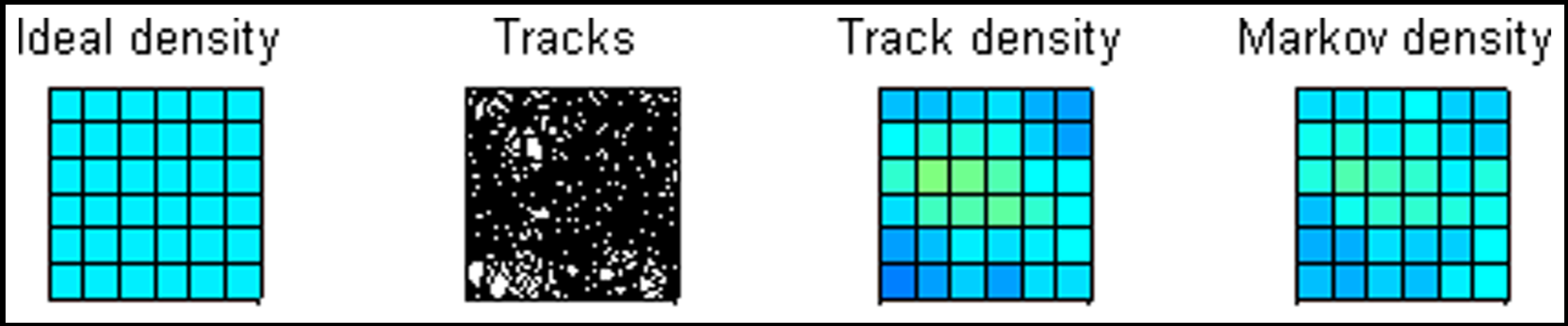
THE EXPERIMENT: MY MARKOV



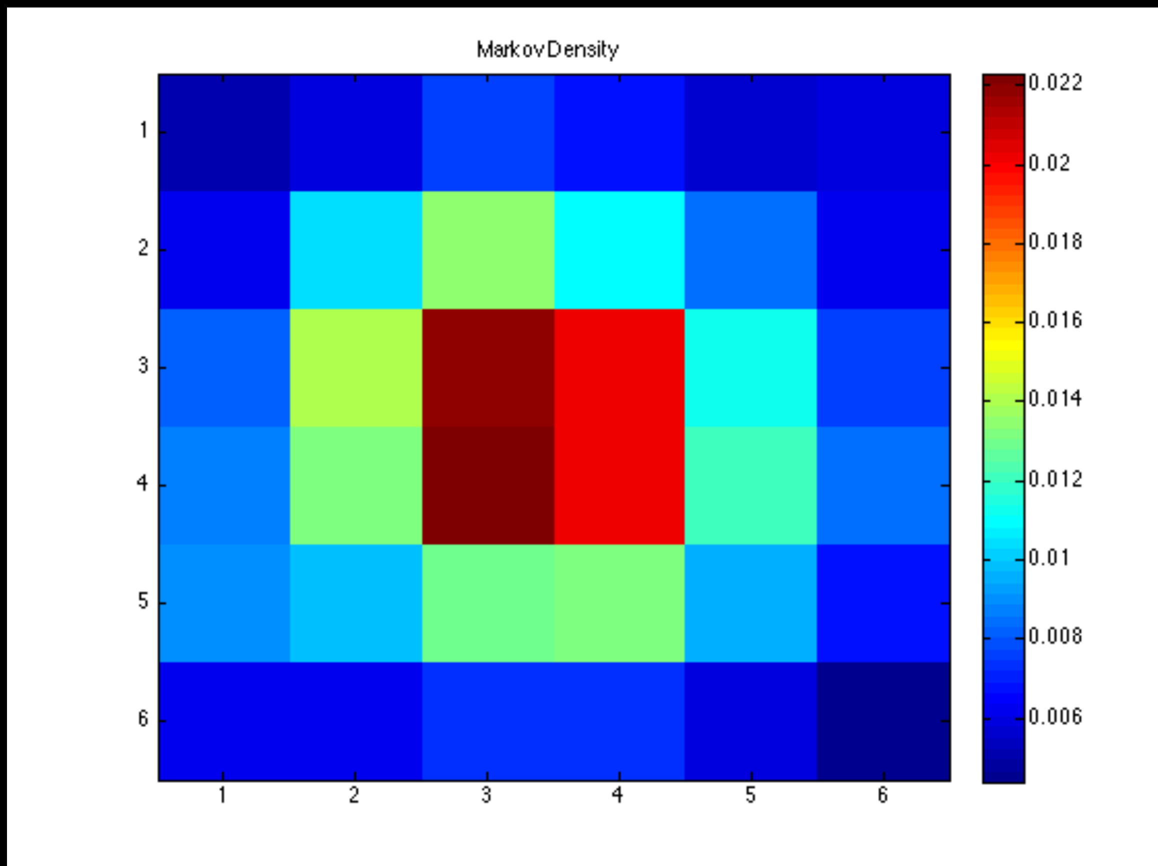
OUTSIDE CELL NOT SHOWN

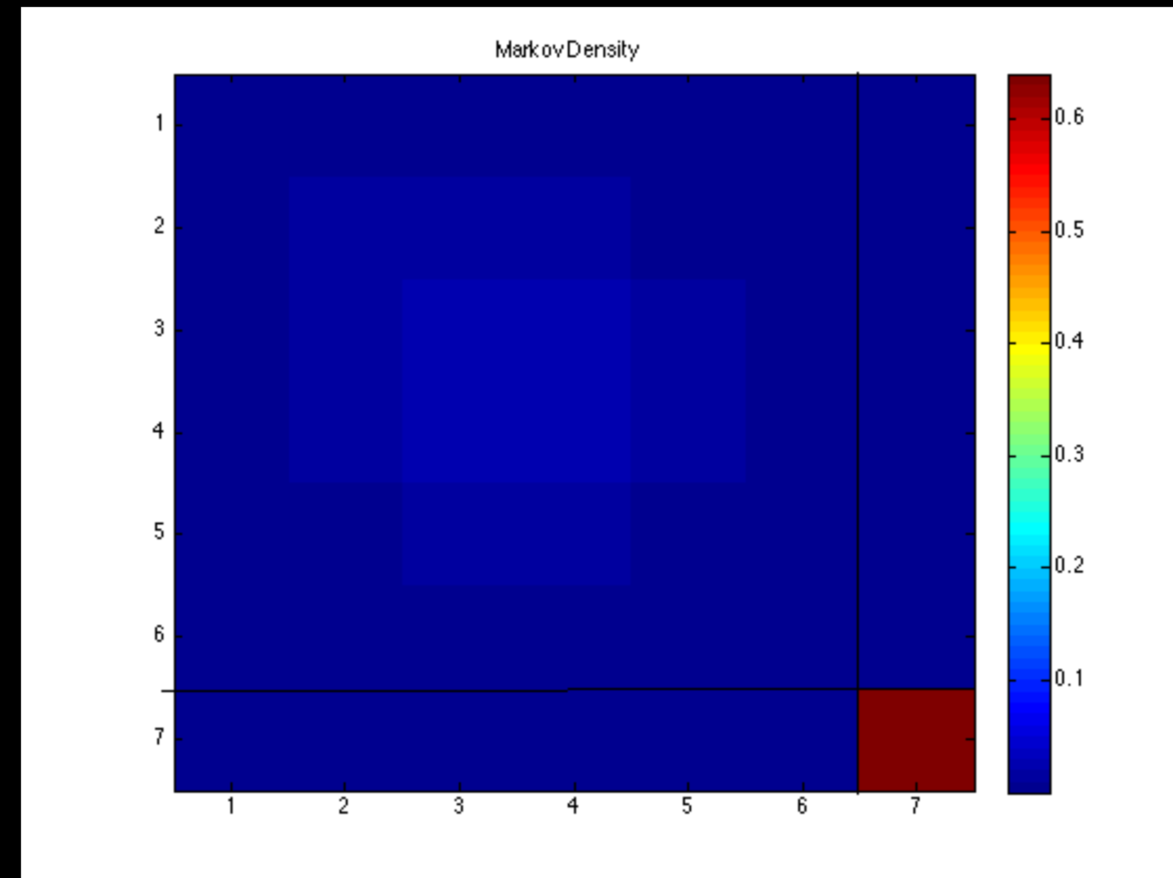


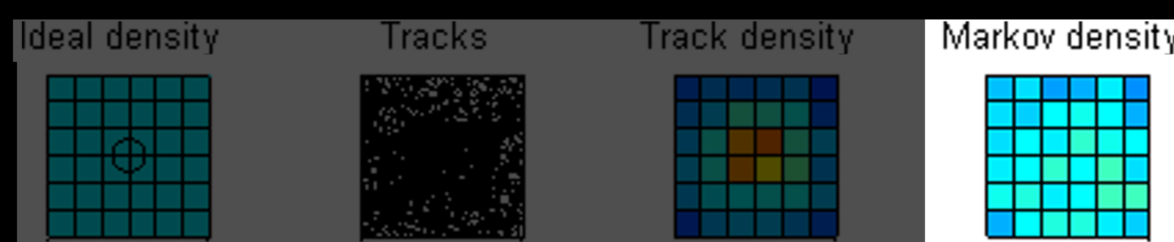
OUTSIDE CELL SHOWN



MARKOV: CENTERED



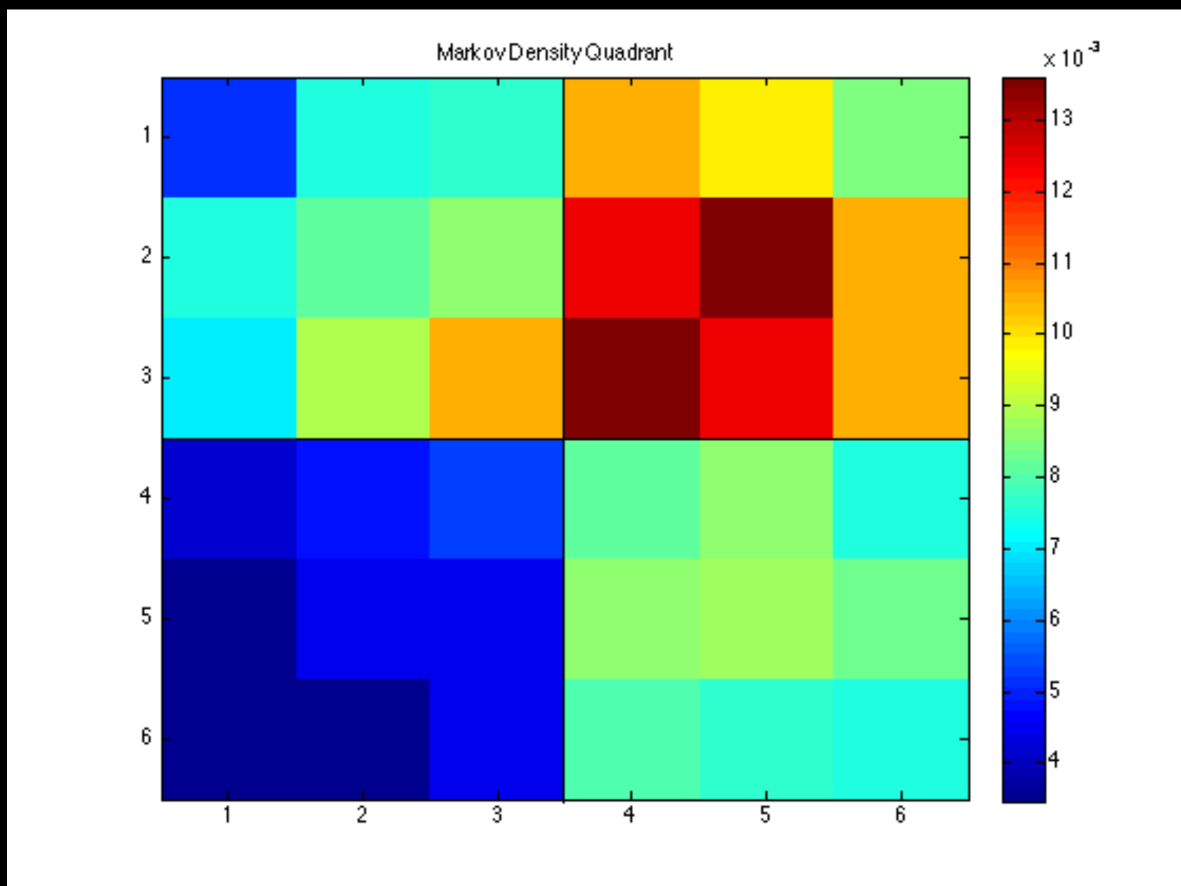


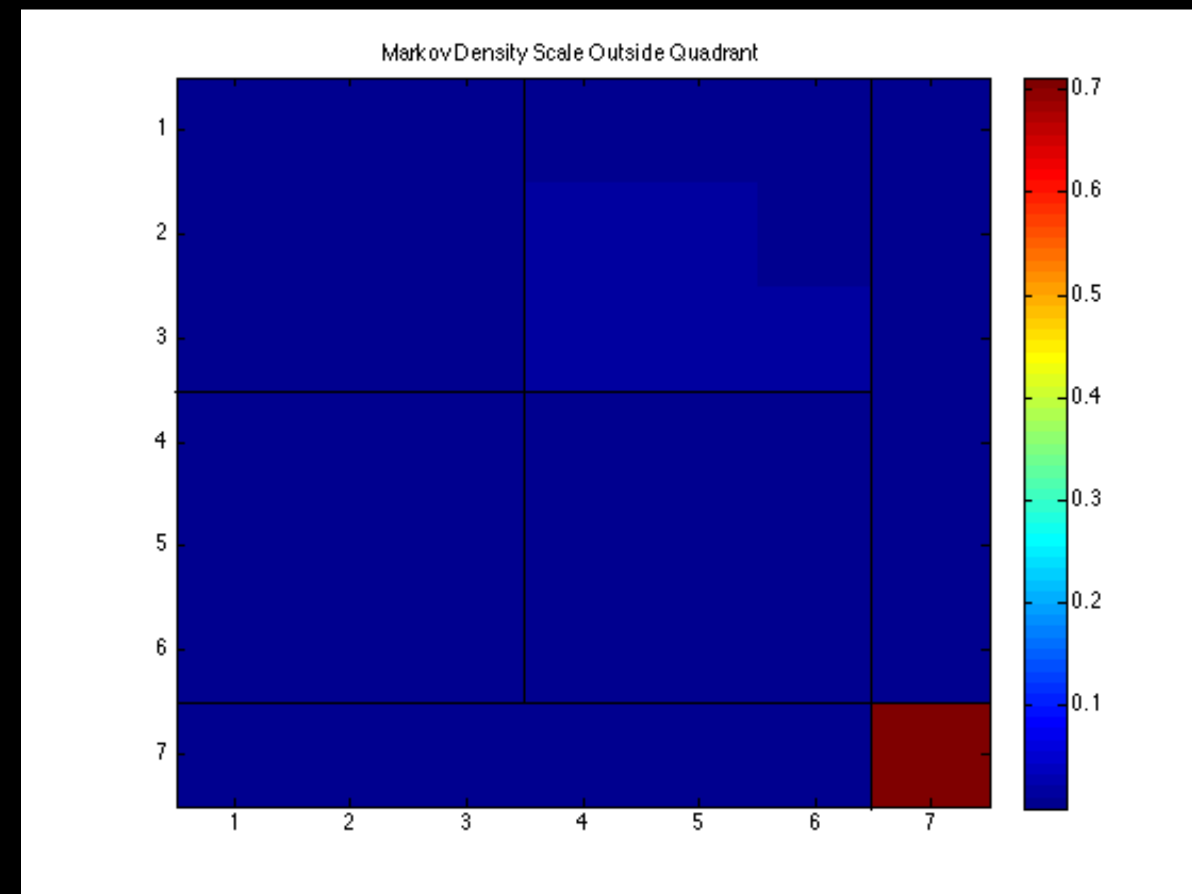


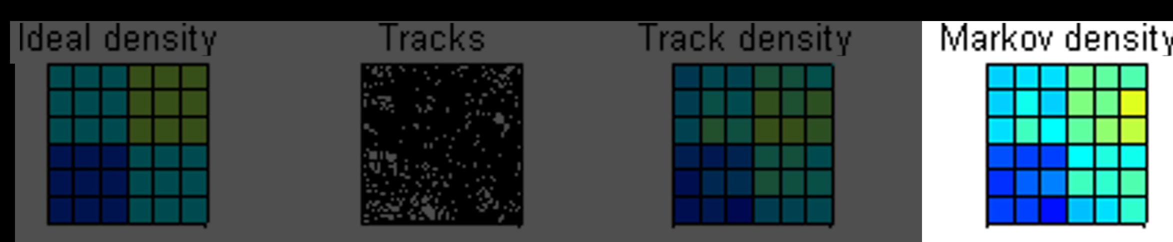

```
markov_heat =

    0.0052    0.0060    0.0076    0.0066    0.0056    0.0059         0
    0.0061    0.0105    0.0134    0.0111    0.0085    0.0061         0
    0.0081    0.0141    0.0219    0.0202    0.0114    0.0077         0
    0.0088    0.0131    0.0223    0.0201    0.0120    0.0084         0
    0.0091    0.0098    0.0129    0.0132    0.0097    0.0069         0
    0.0061    0.0063    0.0072    0.0075    0.0059    0.0044         0
         0         0         0         0         0         0    0.6406
```

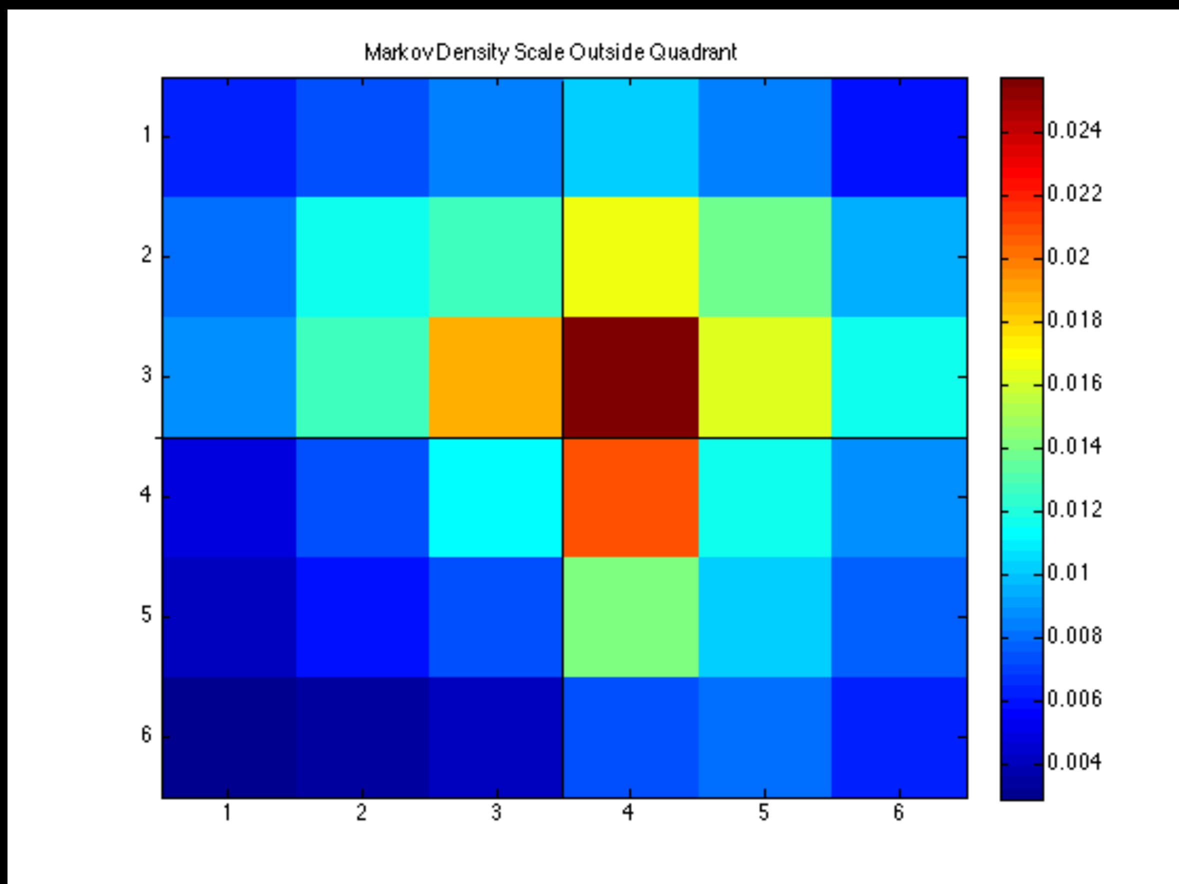
MARKOV: QUADRANT

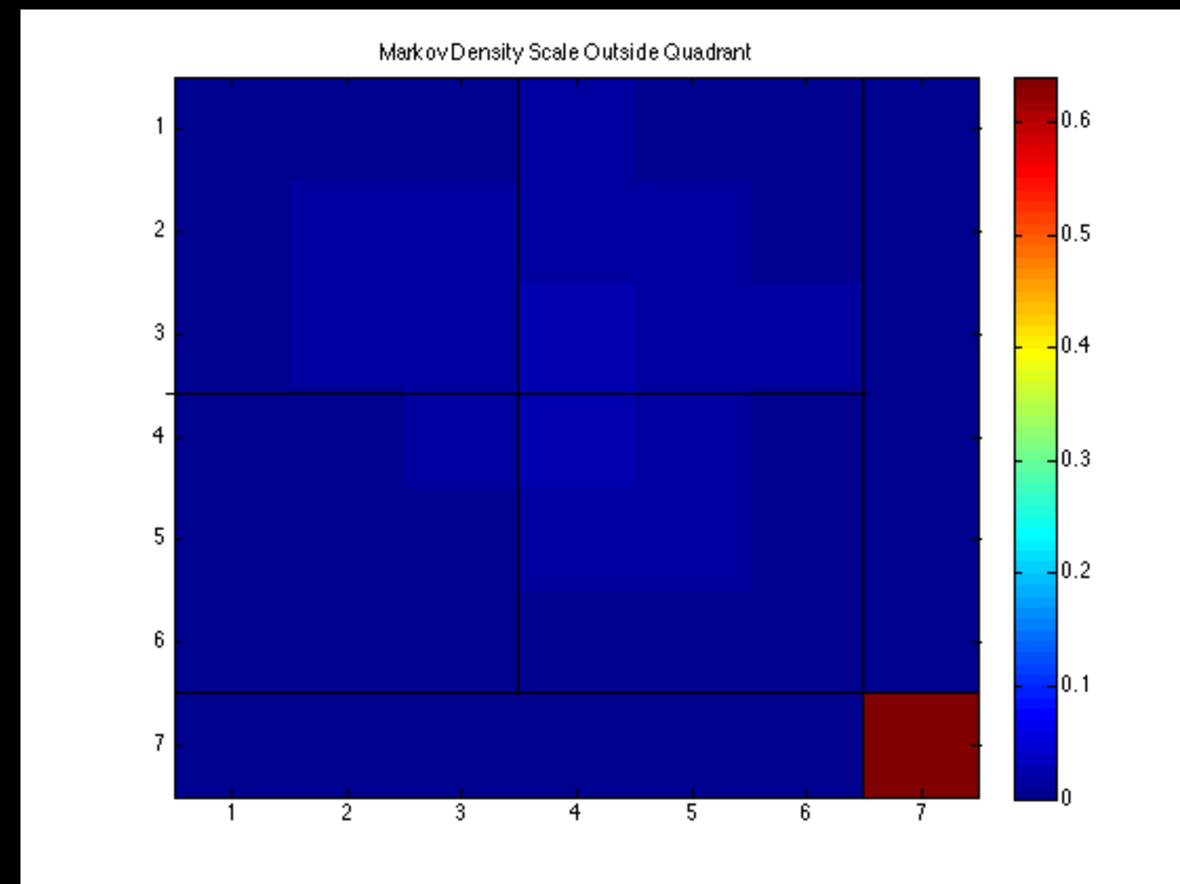


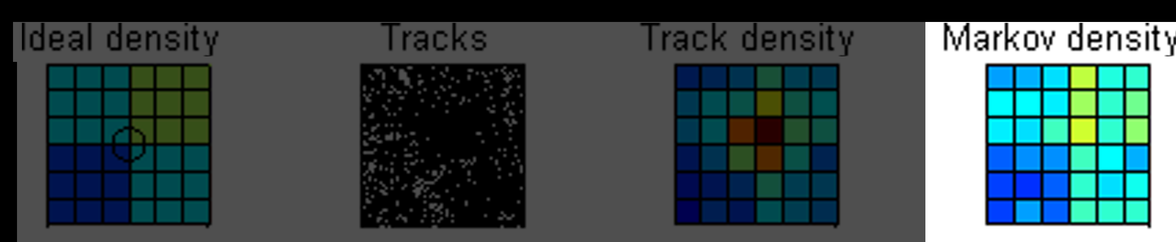


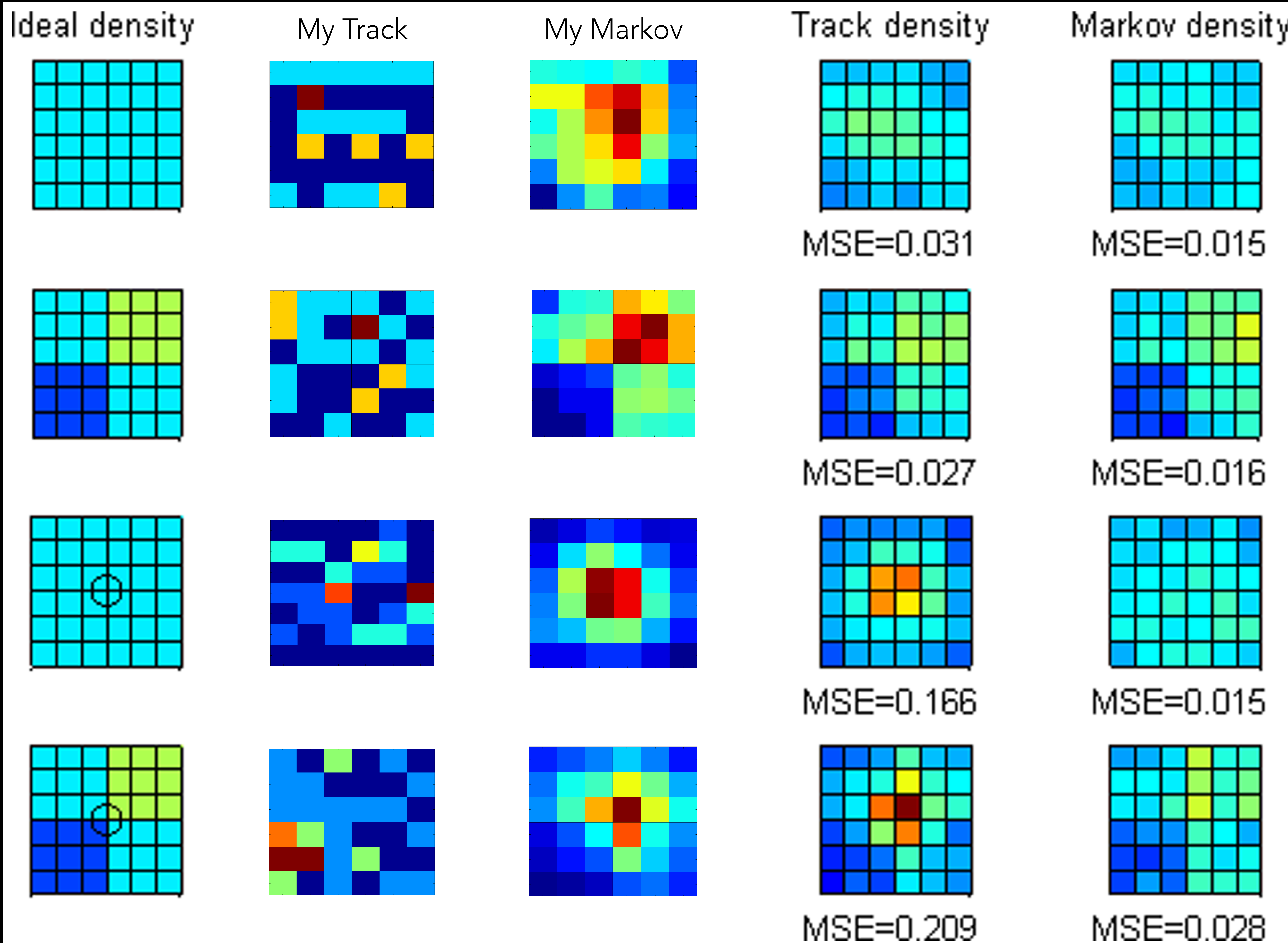


MARKOV: QUADRANT CENTERED








Ideal density
My Track
My Markov
Track density
Markov density
MSE=0.031
MSE=0.015
MSE=0.027
MSE=0.016
MSE=0.166
MSE=0.015
MSE=0.209
MSE=0.028

WEIRD. DOUBLE CHECK

The track data for each simulation were the positions of the agents after each 20 moves. To estimate relative density, the study area was divided into 36 square cells, plus one "exterior" cell for all regions outside the study area. For each of the four simulations, with the two types of study area, and two tracking start types, we present (Fig. 1): a graphical representation of the ideal, actual density of agents over the study area ("Ideal density"); the tracks (just 50 shown for clarity); estimates of relative density obtained by simply summing the number of track positions in each cell ("Track density"); and estimates of relative density obtained using the Markov eigenvector technique described above ("Markov density"). The estimates of relative density were normalized to have the same mean as the ideal densities, excluding the "exterior" cell. Performance of the two techniques was evaluated from the mean (over cells within the study area) squared error of the estimated relative density in a cell compared with its ideal value.

Did not do track data. Only heat map.

Done and accounted for

Done and accounted for

Done and accounted for

Paper never mentioned how they got ideal densities and never shared the averages. Everything else normalized.

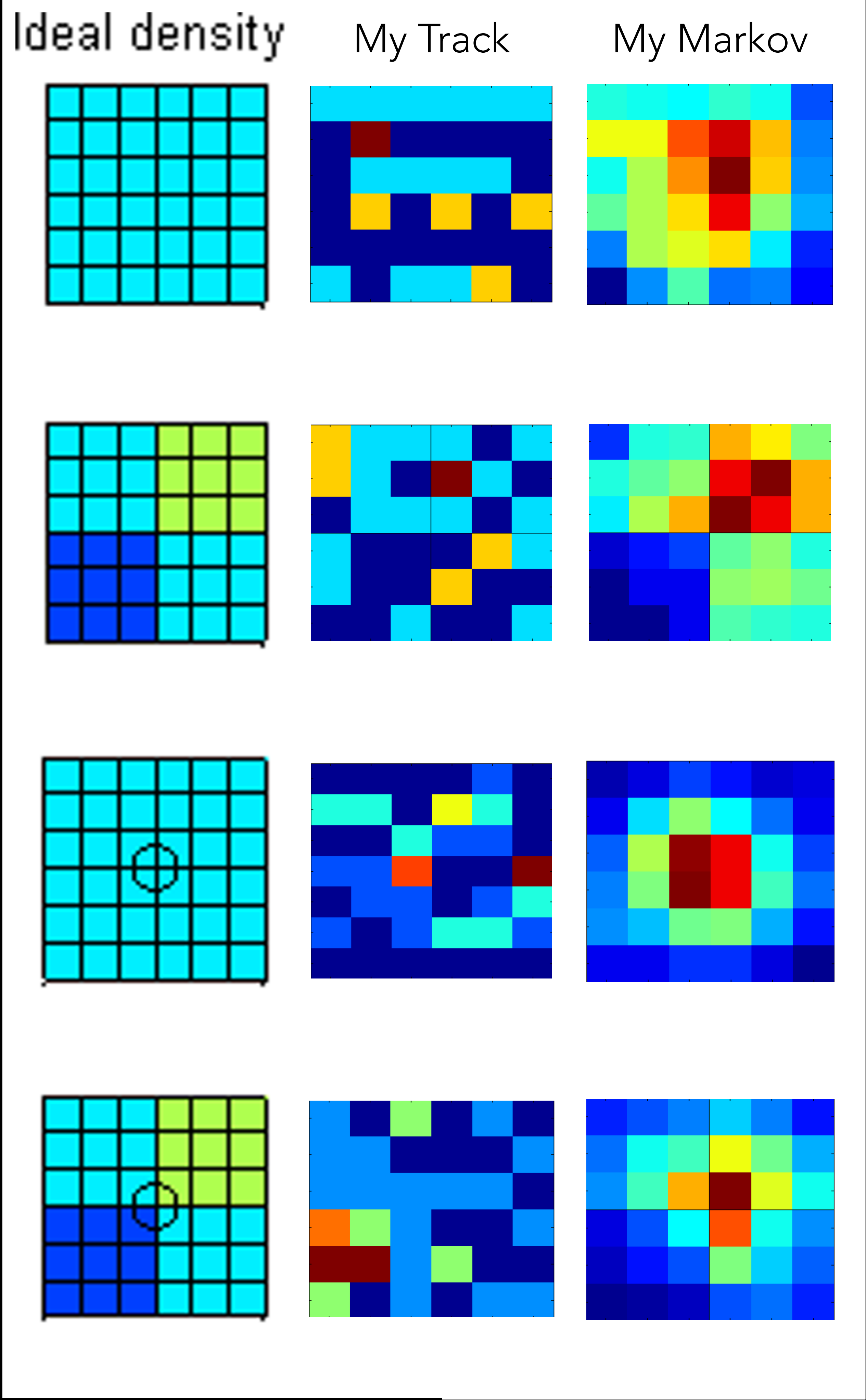
EXPERIMENT PROBLEMS

Methods: Steps are 0.05 -> 20 steps to go across whole thing, wanted 1,000 steps

- bad parameter setting
  - steps size so small, study area so small
- most animals will probably leave
- difficulty in real world animal tracking (only move 20 times until leave?)

DATA INFERENCE

- Paper says 1 and 3 should be the same and 2 and 4 should be the same
- My data: inconclusive, because of the lack of additional information



PAPER PROBLEMS

- Authors have different heat map weight in weird way that might lead to misinterpretation
  - Perhaps difference in steps calculation (some continuous method)
  - Didn't show the color bars to know scale

Illustration using simulated data

We simulated the method in a wide range of conditions, but to illustrate its general performance we present four of these simulations. In each of these simulations, 300 agents made uncorrelated random walks [4] over an *x*-*y* plane with 1,000 moves. The study area was a 1 unit by 1 unit square centered on [0.5, 0.5], although individuals could move outside the study area, or back into it. Two types of study area were simulated:

1. "Homogeneous", in which all moves are of length 0.05. Here the ideal, expected, distribution is homogeneous across the study area.
2. "Quadrants", in which move lengths were $0.05/\sqrt{1.5}$ in the quadrant with $x>0.5$ and $y>0.5$, $0.05/\sqrt{0.5}$ in the quadrant with $x\leq0.5$ and $y\leq0.5$, and 0.05 in the other two quadrants. These differences in move length, in other words the speed of the agents, give relative densities, ideal distributions, of 0.5, 1.0, 1.0, 1.5 in the four quadrants [4].

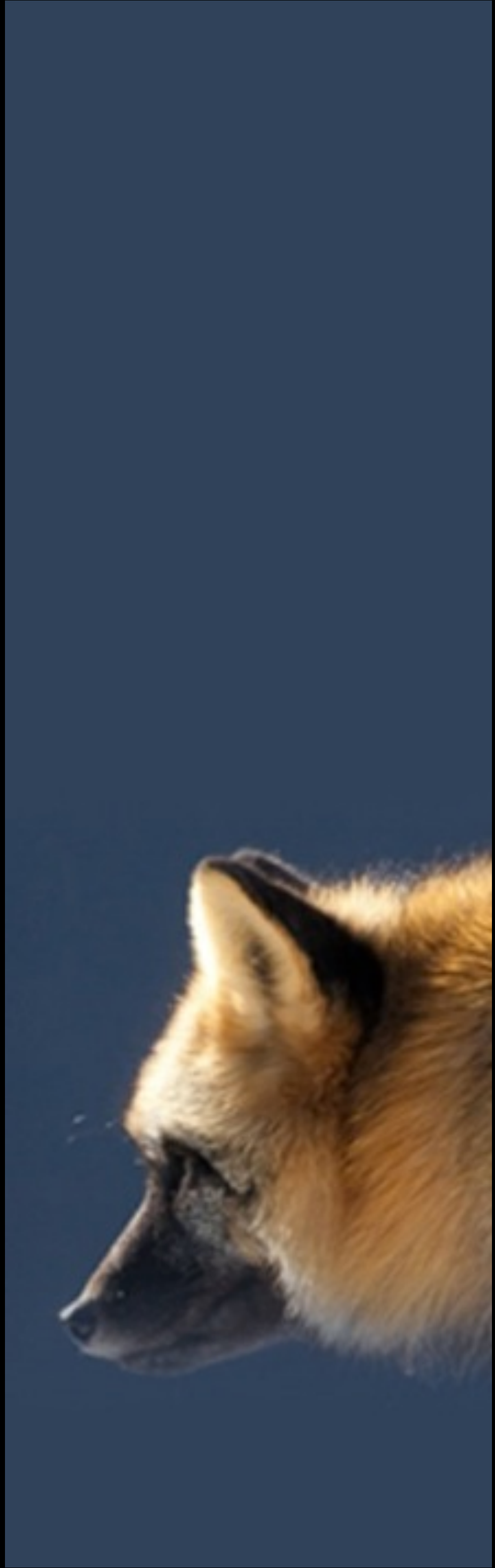
We considered two ways in which tracking was initiated:

1. tracks started at randomly-chosen positions within the study area;
2. tracks started at randomly-chosen positions within 0.1 units of the center of the study area ([0.5, 0.5]).

The track data for each simulation were the positions of the agents after each 20 moves. To estimate relative density, the study area was divided into 36 square cells, plus one "exterior" cell for all regions outside the study area. For each of the four simulations, with the two types of study area, and two tracking start types, we present (Fig. 1): a graphical representation of the ideal, actual density of agents over the study area ("Ideal density"); the tracks (just 50 shown for clarity); estimates of relative density obtained by simply summing the number of track positions in each cell ("Track density"); and estimates of relative density obtained using the Markov eigenvector technique described above ("Markov density"). The estimates of relative density were normalized to have the same mean as the ideal densities, excluding the "exterior" cell. Performance of the two techniques was evaluated from the mean (over cells within the study area) squared error of the estimated relative density in a cell compared with its ideal value.

PAPER PROBLEMS

1. MATLAB code included was incomplete and could not run on its own. Did not show heat map code.
    - from what sense I was able to make, Markov calculated the same way
2. Simulated data not provided
3. Lack of information in implementation
4. Talk of use of ideal densities, but no information
5. Unclear on role of “exterior” cell



REST OF THE PAPER

- Showing applications of sperm whale data
- Doesn't go into detail on how calculated
  - impossible to replicate

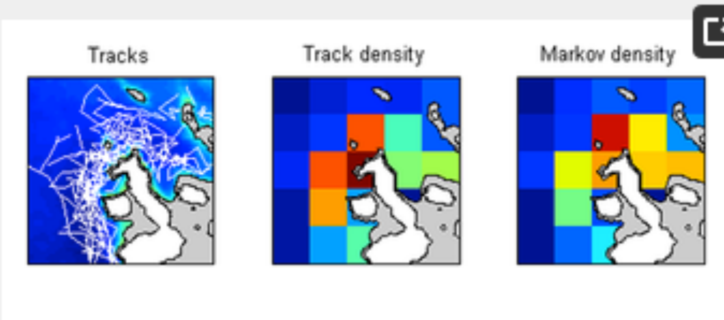
DATA

CALC

Illustration using real data

We also illustrate the method using real tracking data that come from tracks of groups of female and immature sperm whales (*Physeter macrocephalus*) off the Galapagos Islands, Ecuador, between 1985–1995 (study area: 1°S - 1°N; 90° 30'W - 92° 30'W) [5]. The groups of whales were tracked both visually (mainly daylight) and acoustically by listening for the sounds of the whales (mainly at night time) from 10–12 m auxiliary sailing vessels [6]. Positions were determined by SATNAV or GPS (Global Positioning System), and were interpolated every six hours. There were 57 tracks containing 460 locations (i.e. 115 24-hr days tracking).

The study area was divided into 25 square cells, plus one exterior cell. Waters less than 1,000 m deep were masked out. We plot, in Fig. 2, the estimated densities in the study area: the track density from the number of tracking locations divided by area within the cell greater than 1,000 meters deep, and the Markov density, also using the usable area in each cell as a divisor. We also analyzed the same data using depth-delineated cells. The depth ranges chosen were 0–750 m, 750–1,250 m, 1,250–1,750 m, and 3,250–3,750 m.



Download:

PPT PowerPoint slide

PNG larger image (737KB)

TIFF original image (1.06MB)

Figure 2. Estimating spatial sperm whale distributions.
Distribution of groups of sperm whales off the Galapagos Islands (1°S - 1°N; 90° 30'W - 92° 30'W, so the cells are 44.5 km square) from tracking data (shown at left, with water depths); estimated densities from summing tracking positions ("Track density") and the Markov technique ("Markov density"). Only waters greater than 1,000 m deep were considered. Islands are shown in white, and waters less than 1,000 m deep in grey. The density plots use the same normalized color scale ranging from dark blue (near zero) to turquoise (medium) to dark red (maximum).
doi:10.1371/journal.pone.0060901.g002

CONCLUSION

1. **DATA:** Markov looked to be promising
   - would need more data
   - can help environmental studies
2. **EXPERIMENT:** Have data that is not susceptible to loss of information
3. **PAPER**: Be clear and detailed in experiment

SOURCES

- Cover photo from http://www.sej.org/publications/sejournal-su13/tangled-tale-endangered-wolf
- Tag deer picture from http://cloudfront-media.reason.com/mc/_external/2013_07/tagged-animal.jpg?h=230&w=300
- Goose picture from http://www.project-e-track.eu/Portals/4/Images/news%20images/gans%20met%20nieuwe%20E-Track%20tag_small.jpg
- Fox in snow picture from: http://ohbythewayblog.blogspot.com/2010/12/animals-dont-mind.html
- All other pictures from: Fox pictures from dynamicecology.wordpress.com
- Bear Photo from: http://yourshot.nationalgeographic.com/profile/453622
- Paper: http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0060901#pone.0060901.s001