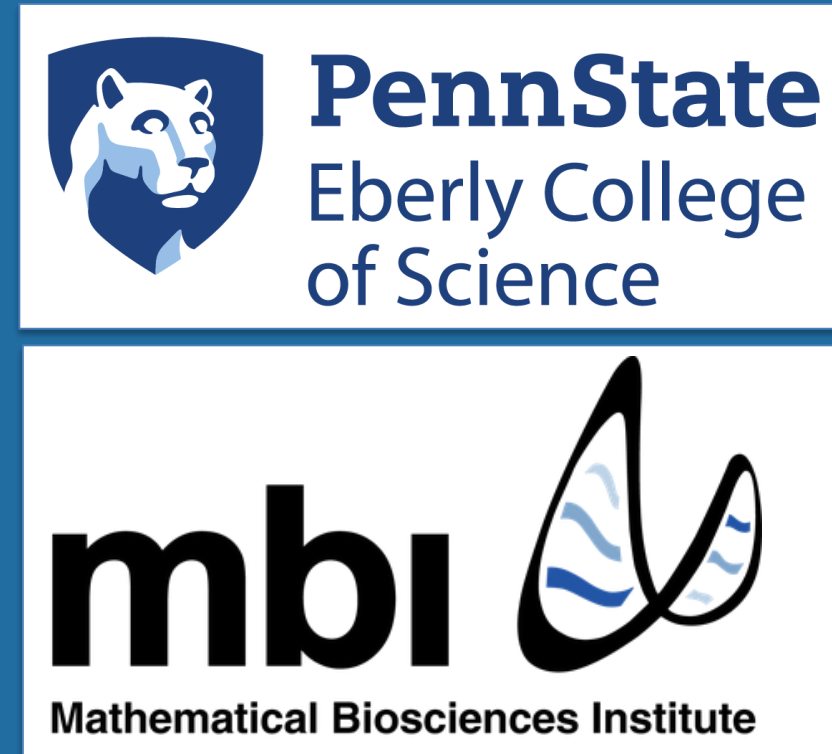


Distinguishing Resource Selection from Heavy-Tailed Dispersal in Spatial Epidemic Models

Maike Morrison ^{*1}, Emily Strong ^{*2}, Ephraim Hanks ³



1. Department of Mathematics, The University of Texas at Austin, Austin, TX 2. Department of Mathematics, St. John Fisher College, Rochester, NY 3. Department of Statistics, the Pennsylvania State University, University Park, PA

* These authors contributed equally to this work. Contact: maike.morrison@gmail.com or ers00083@sjfc.edu

Background

- Patchy terrain is often ignored in models of population movement and disease spread
- Recent work¹ suggests that this could lead to misleading conclusions about future spread

Clark Dispersal Kernel

- This formula from the literature² defines the probability of a movement based on its distance (x) from the start location

$$f(x) = \frac{c}{2\alpha\Gamma(1/c)} \exp\left(-\left|\frac{x}{\alpha}\right|^c\right)$$

α = distance parameter, c = shape parameter, x = distance

- Light-tailed dispersal kernel: infinitesimal probability of long distance movements; $c > 1$
- Heavy-tailed dispersal kernel: non-infinitesimal probability of long distance movements; $c < 1$

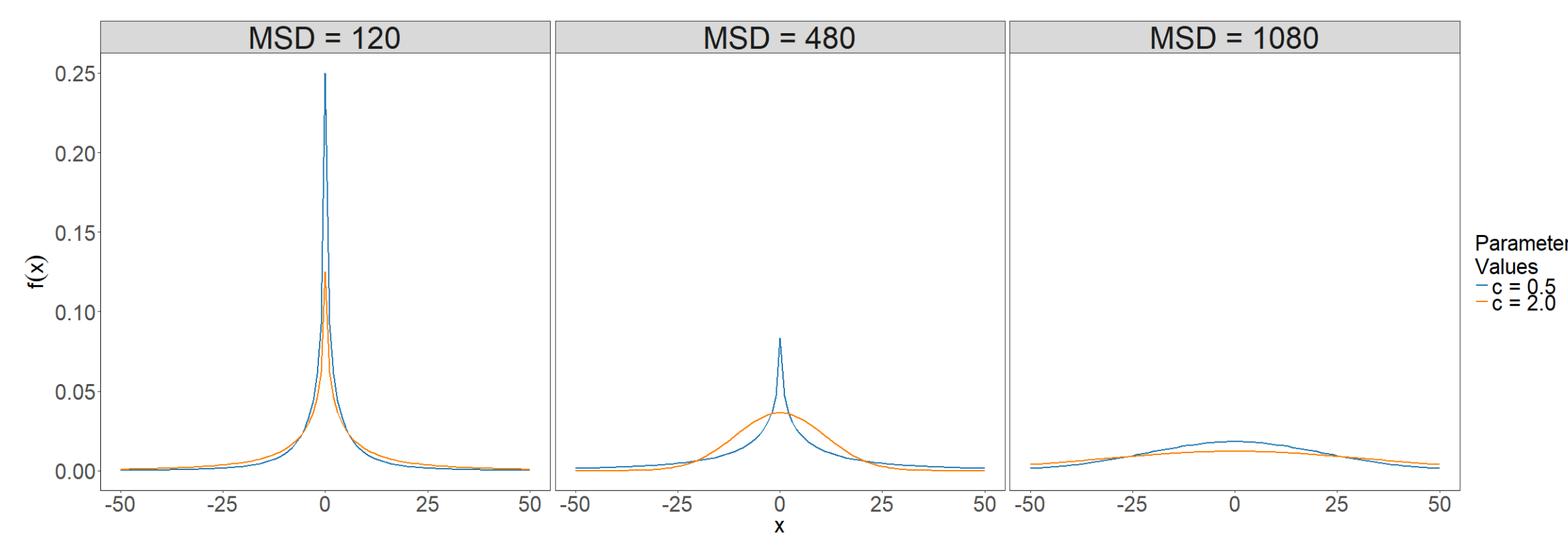


Figure 1. Heavy- and light-tailed Clark dispersal kernel distributions for values of α and c that yield three different mean-squared displacements (MSD).

Transition Matrix with Patchy Terrain

M is a probability transition matrix such that $M_{j,i}$ is the probability of a movement from node j to node i .

$$M_{j,i} = \frac{w(i|\beta)f(|j-i||\alpha,c)}{\sum_{i=1}^k w(i|\beta)f(|j-i||\alpha,c)}$$

- $w(i|\beta) = \exp(x(i)\beta)$ is a weighting function that models ending location's terrain features ($x(i)$)
- $f(|j-i||\alpha,c)$ is the probability of a movement of distance $|j-i|$ given by the Clark dispersal kernel
- β determines the preference for good terrain over bad terrain

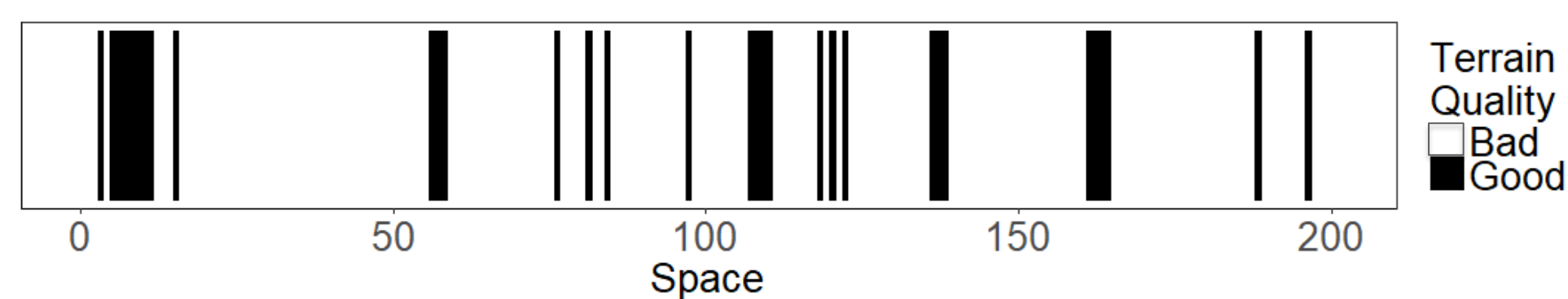


Figure 2. Example patchy terrain generated using the modified random clusters algorithm³, which allows for control of patch cohesion and the proportion of good terrain.

Acknowledgements

This research has been supported in part by the Mathematical Biosciences Institute and the National Science Foundation under grants DMS 1440386 and DMS 1757423.

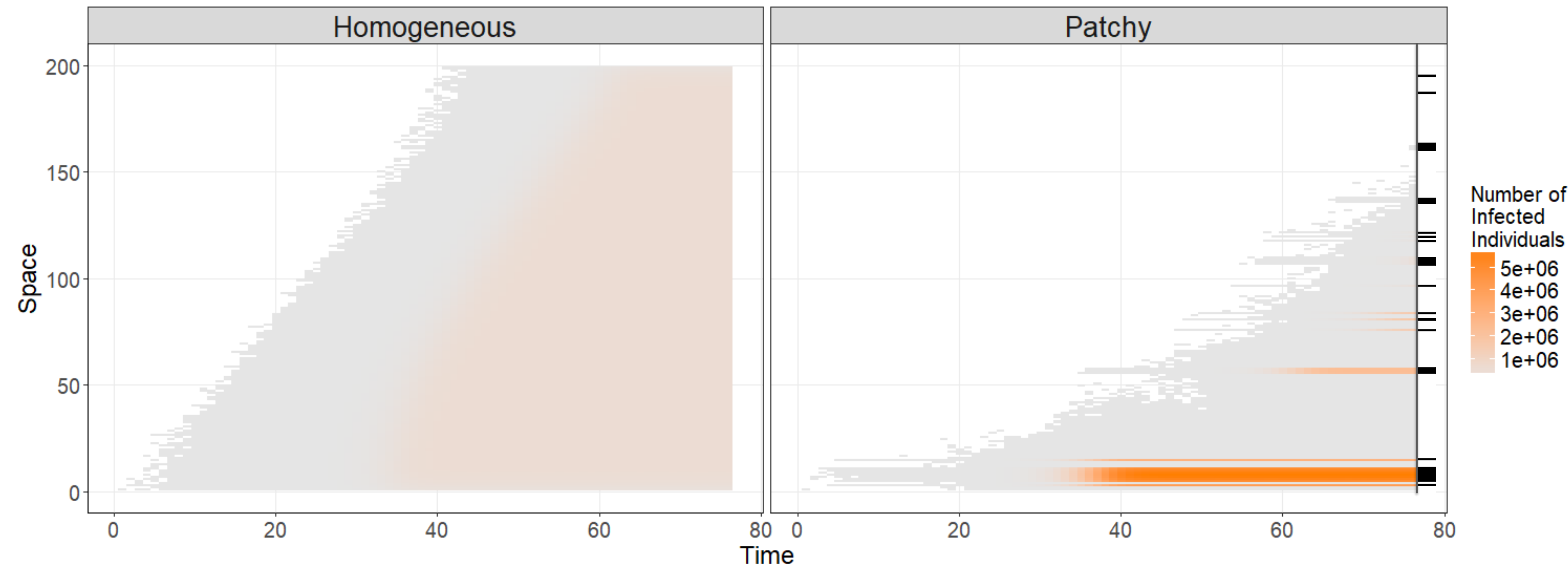


Figure 3. Simulation of population movement over a space of 200 nodes with homogeneous terrain (L) or patchy terrain (R). The epidemic spreads more slowly over patchy terrain and produces more concentrated pockets of infection but fewer total infected individuals.

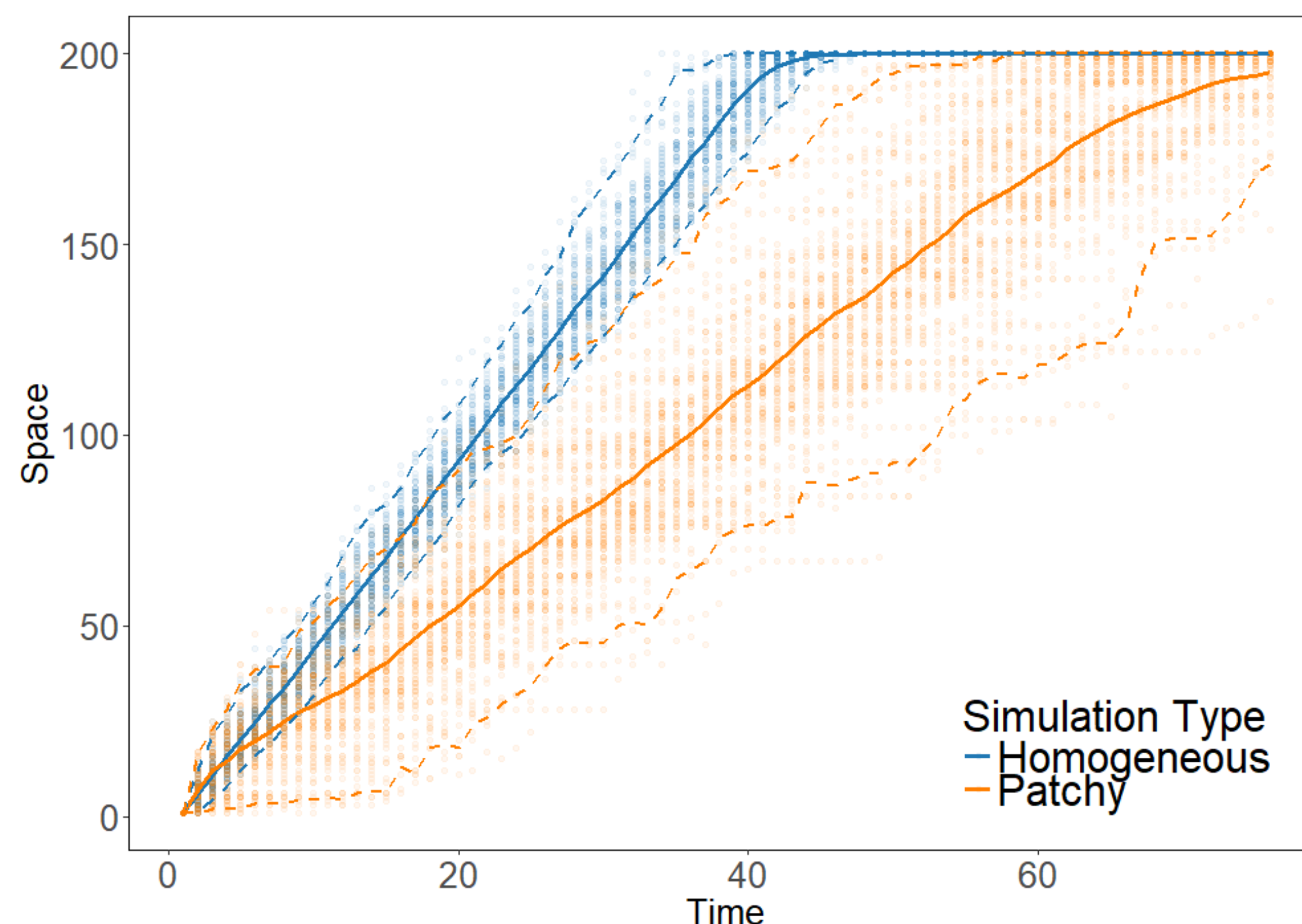


Figure 4. Furthest forward dispersers (FFDs) across 100 homogeneous simulations ($\beta = 0$) and 100 patchy landscape simulations ($\beta = 7$), with average FFDs (solid lines) and 95% quantiles (dotted lines). The more gradual slope and wider quantiles of the patchy simulations suggest that disease spread is slower and more erratic over patchy terrain.

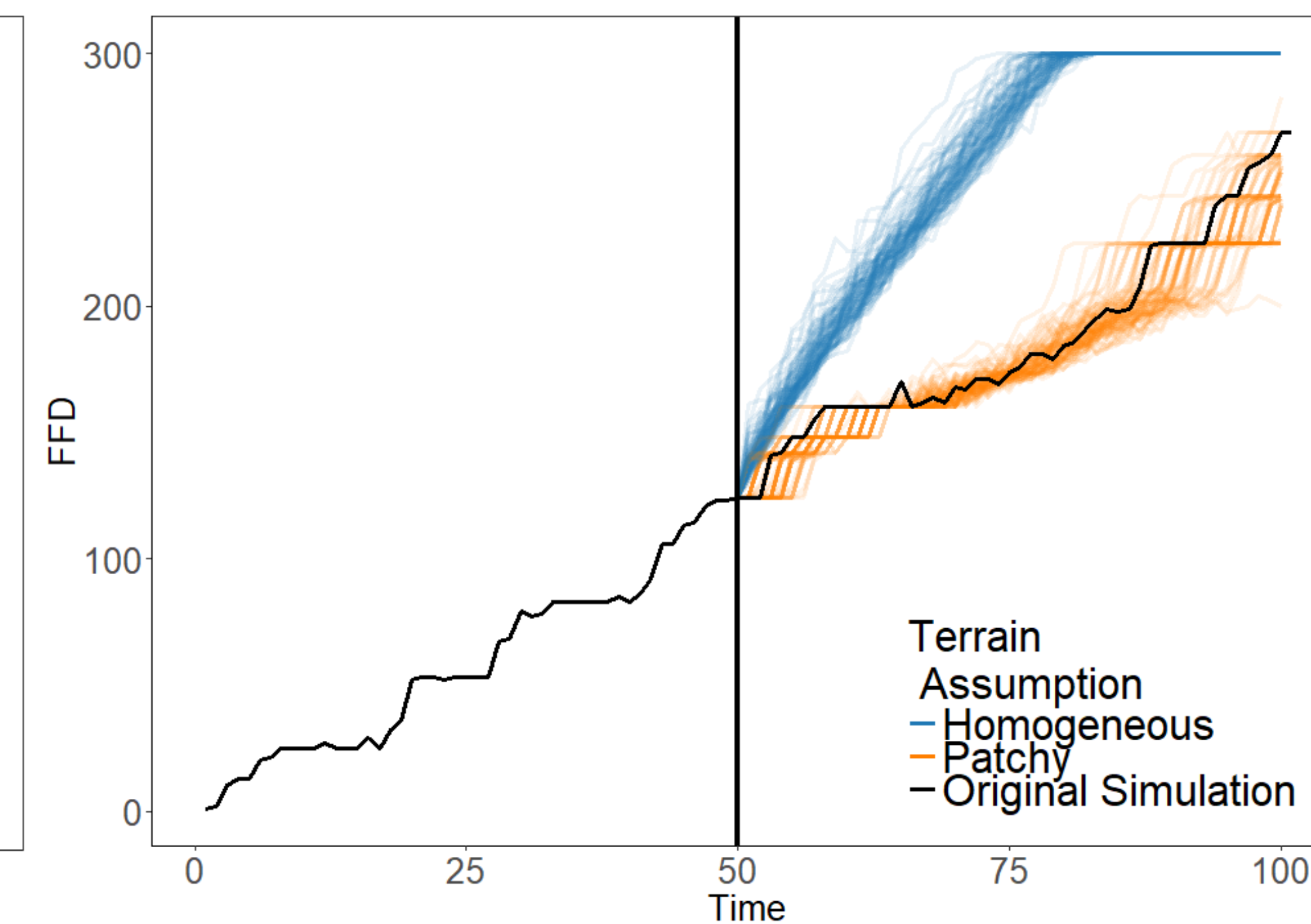


Figure 5. The first 50 time steps of the original simulation were used for parameter estimation under the assumptions of either homogeneous or patchy terrain. 100 simulations were run for each of the resulting parameter estimates to predict disease spread over the next 50 time steps. **Assuming homogeneous terrain leads to overestimating epidemic wave speed.**

Results

- Ignoring terrain patchiness in models of population movement and disease spread leads to misleading conclusions:
 - Disease spreads more slowly over patchy terrain
 - Infected individuals are much more clustered than would be expected in homogeneous terrain
 - The estimated dispersal kernel parameters are heavier tailed (see Table 1)
- Larger β values (which determine the difference between good and bad terrain) slow spread
- If patchy terrain is ignored, kernel parameters are most poorly estimated when there is little good terrain and β is large (strong preference for good terrain)

| True Terrain | Terrain Assumption | Parameter | True Value | Mean | SD |
|--------------|--------------------|-----------|------------|----------|---------|
| H | Homogeneous | α | 7.00 | 7.00 | 2.07e-4 |
| | | c | 2.00 | 2.00 | 8.62e-5 |
| | Patchy | α | 7.00 | 7.00 | 6.84e-4 |
| | | c | 2.00 | 2.00 | 2.96e-4 |
| P | Homogeneous | β | 0.00 | -2.02e-5 | 1.53e-4 |
| | | α | 7.00 | 3.01 | 9.45e-1 |
| | Patchy | c | 2.00 | 1.05 | 1.87e-1 |
| | | α | 7.00 | 7.00 | 5.03e-4 |
| | | β | 7.00 | 7.00 | 9.34e-4 |

Table 1. The effect of terrain assumption on parameter estimation for 100 homogeneous (H) simulations and 100 patchy (P) simulations. Highlighted box shows biased estimation.

Simulation Framework

- Space comprised of k nodes in a line
- Some nodes “good” terrain, others “bad”
- Initial conditions: One infected individuals at node 1; $n = \{n_1, n_2, \dots, n_k\}$ is stationary distribution of movement model

1. Disease Transmission within Each Node:

Euler approximation to Susceptible-Infected model

$$\tilde{I}_i(t+1) = I_i(t) + \frac{R_0 S_i(t) I_i(t)}{n_i}$$

- $S_i(t)$ = number of susceptible individuals at node i at time t
- $I_i(t)$ = number of infected individuals at node i at time t
- $n_i = S_i(t) + I_i(t)$; total number of individuals at node i
- $\tilde{I}_i(t+1)$ is the number of infected individuals at node i at time $t+1$ after transmission and before movement
- R_0 is the basic reproduction number of the disease, a measure of disease transmissibility

2. Population Movement:

$$X_i \sim \text{Multi}(\tilde{I}_i(t+1), p_i)$$

$$I(t+1) = \sum_{i=1}^k X_i$$

$I_i(t+1)$ is the number of infected individuals at node i at time $t+1$ after both transmission and movement

Maximum Likelihood Estimation of α , c , and β

Our likelihood function is given by:

$$L(X|\alpha, c, \beta) = \sum_{t=1}^T \sum_{j=1}^k -\log(p(X(t)_{\cdot,j}|\alpha, c, \beta))$$

- X is a list of T $k \times k$ matrices such that the j^{th} column of its t^{th} element ($X(t)_{\cdot,j}$) represents the j^{th} random vector from step 2 of the simulation at time step t

$$p(X(t)_{\cdot,j}|\alpha, c, \beta) = \binom{n_j}{X(t)_{\cdot,j}} \prod_{i=1}^k (M_{j,i})^{X(t)_{i,j}}$$

α , c , and β are estimated by maximizing $L(X|\alpha, c, \beta)$ numerically.

References

1. Dhanushi A Wijeyakulasuriya, Ephraim M Hanks, Benjamin A Shaby, and Paul C Cross. Extreme Value Based Methods for Modeling Elk Dispersal. *In review*.
2. James S. Clark. Why Trees Migrate So Fast: Confronting Theory with Dispersal Biology and the Paleorecord. *The American Naturalist*, 152(2):204-224, 1998.
3. S. Saura and J. Martinez-Millan. Landscape patterns simulation with a modified random clusters method. *Landscape Ecology*, 15(7):661-678, 2000.