Searching for the best distance-based population density estimator

沈国春

华东师范大学

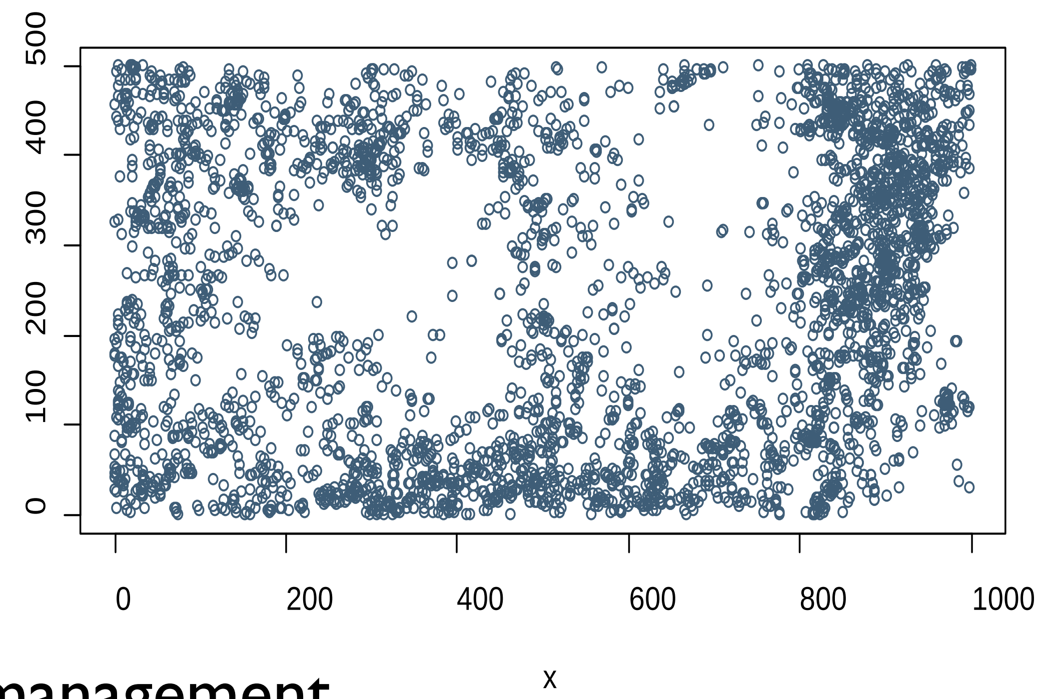
Outline

1. Our problem and objectives.
2. Three new general distance-based population density estimators.
3. Performance comparison

Problem and objectives

Problem---Population Size

- How many **individuals** of a plant species in an observed area?

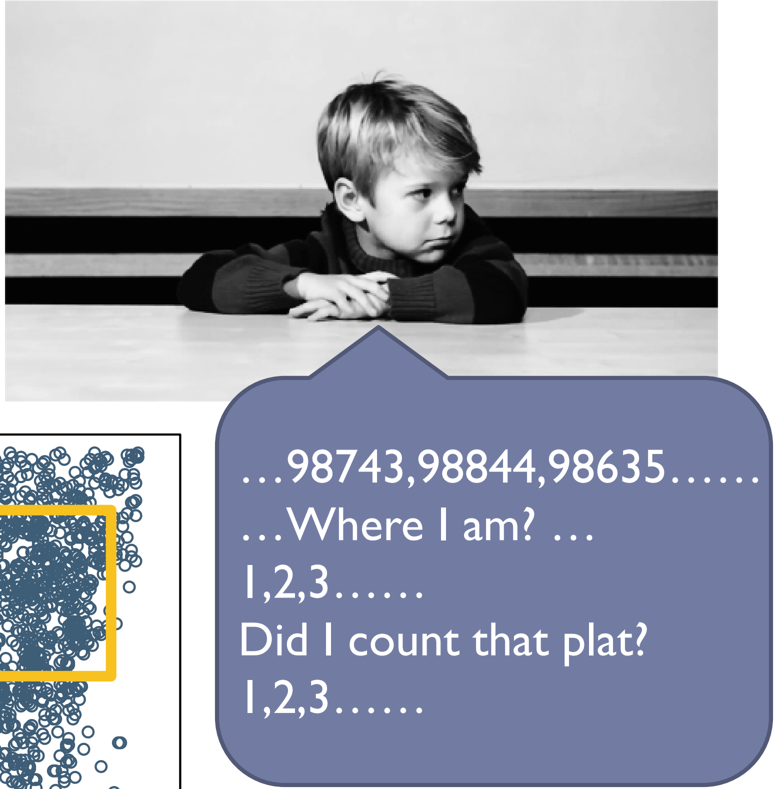


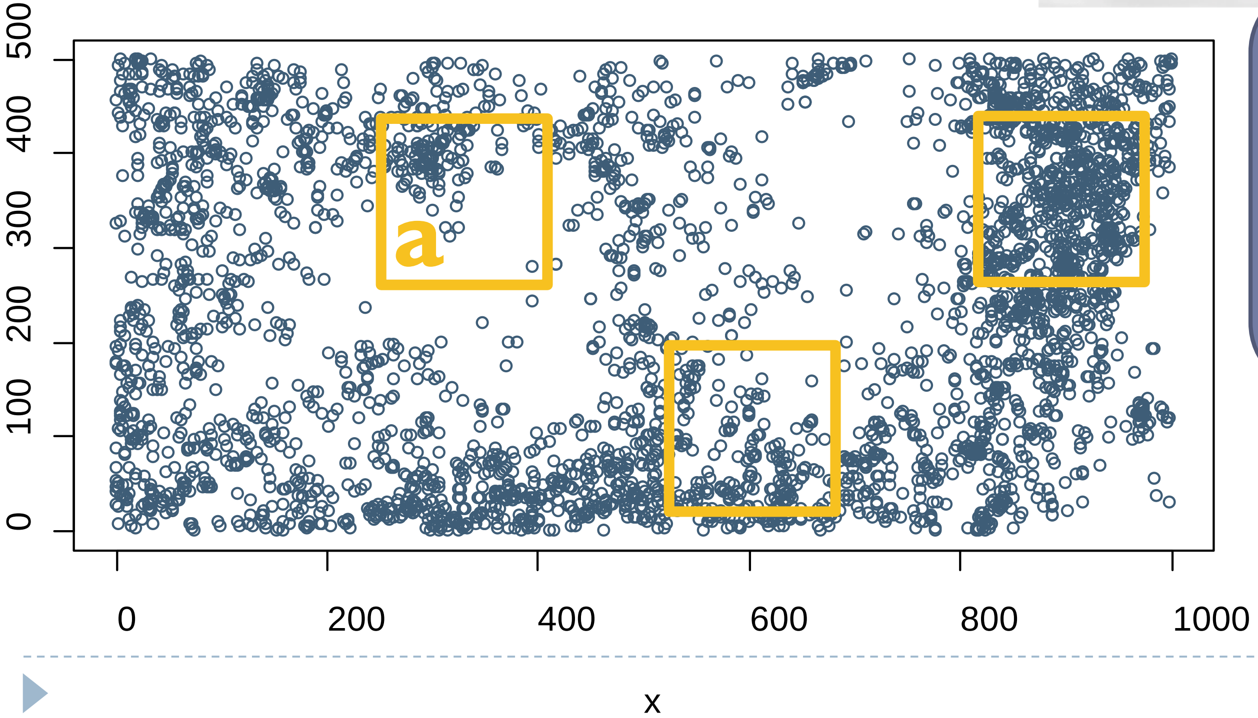
- Basic parameter
- Conservation
- Biological resource management

Quadrat Method---Population Size

- Quadrat method

$$\widehat{N}(A_0) = \frac{N_i}{n}\frac{A_0}{a}$$

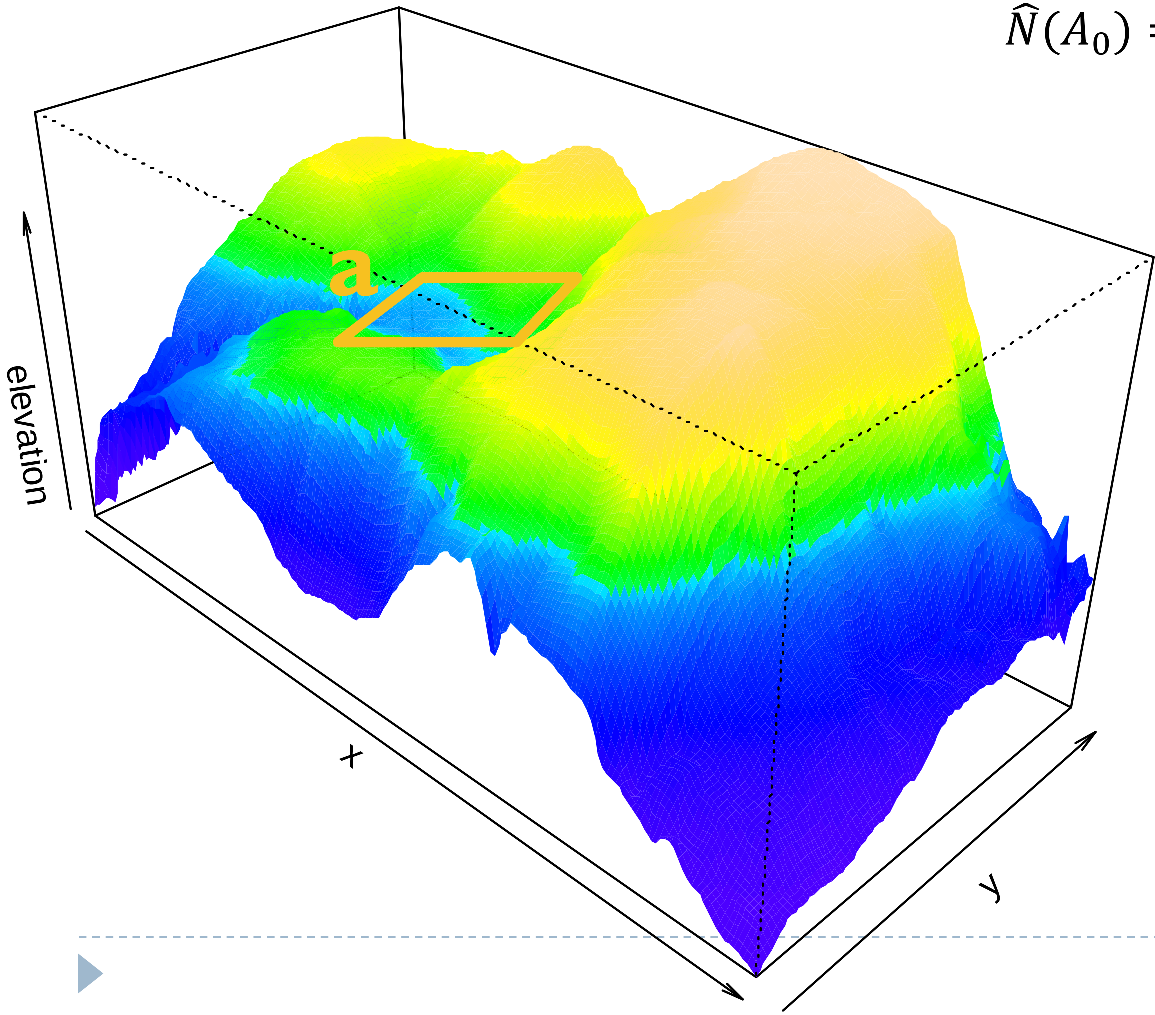




Quadrat Method---Population Size

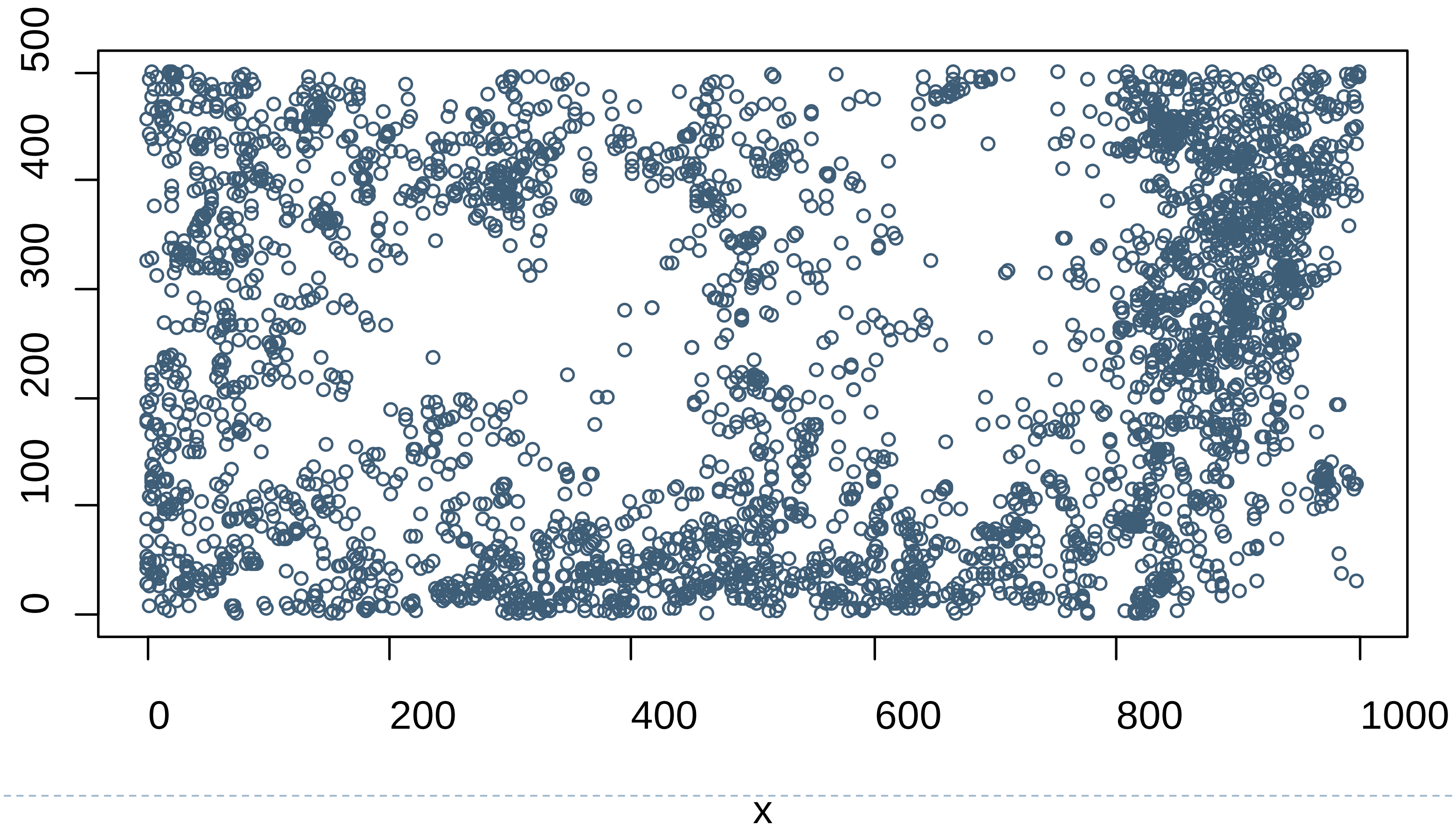
- Quadrat method

$$\hat{N}(A_0) = \frac{N_i}{n}\frac{A_0}{a}$$



Problem

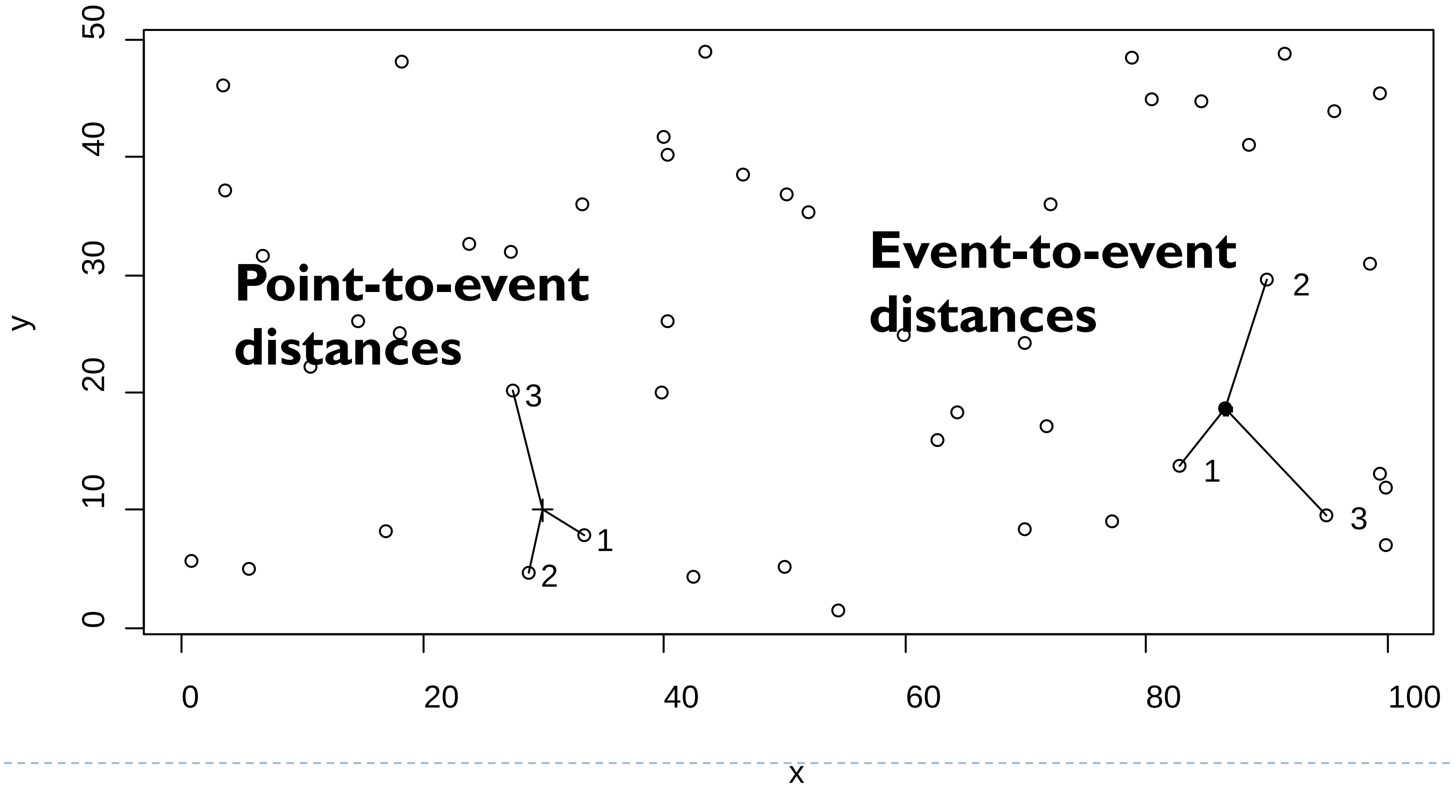
- Can we find a **efficient and robust** way to estimate the population density?



Distance Method---Population Size

Distance-based population density estimator (DPDE)

$$\widehat{N}(A_0) = \lambda A_0 \leftarrow f(\{r\})$$



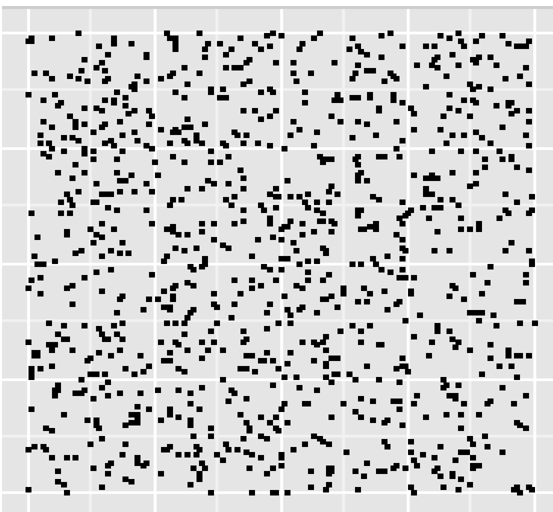
Existed DPDEs

- At least 30 DPDEs existed.
- Which one we can use?

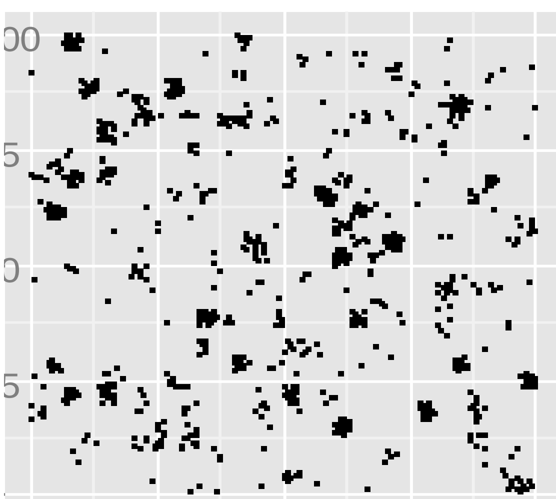
TABLE 1. Summary of the density estimators used in the simulations, their formulae, and the primary references.

Description*	Formula†	References
Basic Distance (BD) estimators		
1. Closest individual (CI)	$BDCI = 1/(4[\Sigma R_{(1)i}/N]^2)$	Cottam et al. 1953, Cottam and Curtis 1956, Kendall and Moran 1963, Pollard 1971
2. Nearest neighbor (NN)	$BDNN = 1/(2.778[\Sigma H_{(1)i}/N]^2)$	Cottam and Curtis 1956
3. Second nearest neighbor (2N)	$BD2N = 1/(2.778[\Sigma H_{(2)i}/N]^2)$	Cottam and Curtis 1956
4. Compound	BDAV2 = (BDCI + BDNN)/2	Diggle 1975
5. Another compound	BDAV3 = (BDCI + BDNN + BD2N)/3	This paper
Batcheler-Bell (BB) estimators		
6. Closest individual (CI)	$BBCI = p/\pi[\Sigma R^2_{(1)i} + (N - p)R^2]$	Batcheler and Bell 1970
7. Nonrandomness (NR) corrected	BBNR (see reference)	Batcheler and Bell 1970
Non-parametric (NP) estimators		
8. Original bias reduced (i.e., general form [GF])	$NPGF = (N^{1/2} - 1)/NA_{([N^{1/2}])}$	Patil et al. 1979
9. Interpolated original general form (IG)	$NPIG = (N^{1/2} - 1)/NA^*_{(N^{1/2})}$	This paper
10. Optimal form (OF)	$NPOF = (N^{2/3} - 1)/NA_{([N^{2/3}])}$	Patil et al. 1982
11. Interpolated optimal form (IO)	$NPIO = (N^{2/3} - 1)/NA^*_{(N^{2/3})}$	This paper
Kendall-Moran (KM) estimators		
12. CI and NN search areas pooled (P)	$KMP = \{[\Sigma(p_i + n_i)] - 1\}/\Sigma B_i$	Kendall and Moran 1963, James 1971
13. CI, NN, 2N search areas pooled (i.e., pooled with search area to 2N [2P])	$KM2P = \{[\Sigma(p_i + n_i + m_i)] - 1\}/\Sigma C_i$	Kendall and Moran 1963
T-Square (TS) estimators		
14. Basic T² estimator (BA)	$TSBA = 2N/[\pi \Sigma R_{(1)i}^2 + 0.5\pi \Sigma T_i^2]$	Diggle 1975
15. Reduced bias (RB) in aggregated populations	$TSRB = N/\pi[(\Sigma R_{(1)i}^2)(0.5 \Sigma T_i^2)]^{1/2}$	Diggle 1975
16. Robust (Byth [B])	$TSB = N^2/[(2 \Sigma R_{(1)i})(\sqrt{2})(\Sigma T_i)]$	Byth 1982
Ordered Distance (OD) estimators		
17. Closest individual	$ODCI = (N - 1)/\pi \Sigma(R_{(1)i})^2$	Morisita 1957, Pollard 1971
18. Second closest individual (2C)	$OD2C = (2N - 1)/\pi \Sigma(R_{(2)i})^2$	Morisita 1957, Pollard 1971
19. Third closest individual (3C)	$OD3C = (3N - 1)/\pi \Sigma(R_{(3)i})^2$	Morisita 1957, Pollard 1971
Angle-Order (AO) estimators		
20. Point-centered-quarter (i.e., 1 observation per quadrant [1Q])	$AO1Q = 12N/\pi \Sigma 1/R_{(1)ij}^2$	Stearns 1949, Cottam et al. 1953, Cottam and Curtis 1956, Morisita 1957, Pollard 1971
21. Second closest individual in each quadrant (2Q)	$AO2Q = 28N/\pi \Sigma 1/R_{(2)ij}^2$	Morisita 1957, Pollard 1971
22. Third closest individual in each quadrant (3Q)	$AO3Q = 44N/\pi \Sigma 1/R_{(3)ij}^2$	Morisita 1957, Pollard 1971
23. Third closest individual in each quadrant (3)	$AO3 = [2/\pi N] \Sigma \Sigma(1/R_{(3)ij})$	Morisita 1971

Existed DPDEs



Assumption:
Random
distribution



Population distributions of most species in nature are **aggregated**.

TABLE 1. Summary of the density estimators used in the simulations, their formulae, and the primary references.

Description*	Formula†	References
Basic Distance (BD) estimators		
1. Closest individual (CI)	$BDCI = 1/(4[\Sigma R_{(1)i}/N]^2)$	Cottam et al. 1953, Cottam and Curtis 1956, Kendall and Moran 1963, Pollard 1971
2. Nearest neighbor (NN)	$BDNN = 1/(2.778[\Sigma H_{(1)i}/N]^2)$	Cottam and Curtis 1956
3. Second nearest neighbor (2N)	$BD2N = 1/(2.778[\Sigma H_{(2)i}/N]^2)$	Cottam and Curtis 1956
4. Compound	$BDAV2 = (BDCI + BDNN)/2$	Diggle 1975
5. Another compound	$BDAV3 = (BDCI + BDNN + BD2N)/3$	This paper
Batcheler-Bell (BB) estimators		
6. Closest individual (CI)	$BBCI = p/\pi[\Sigma R^2_{(1)i} + (N - p)R^2]$	Batcheler and Bell 1970
7. Nonrandomness (NR) corrected	BBNR (see reference)	Batcheler and Bell 1970
Non-parametric (NP) estimators		
8. Original bias reduced (i.e., general form [GF])	$NPGF = (N^{1/2} - 1)/NA_{([N^{1/2}])}$	Patil et al. 1979
9. Interpolated original general form (IG)	$NPIG = (N^{1/2} - 1)/NA^*_{(N^{1/2})}$	This paper
10. Optimal form (OF)	$NPOF = (N^{1/2} - 1)/NA_{([N^{1/2}])}$	Patil et al. 1982
11. Interpolated optimal form (IO)	$NPIO = (N^{1/2} - 1)/NA^*_{(N^{1/2})}$	This paper
Kendall-Moran (KM) estimators		
12. CI and NN search areas pooled (P)	$KMP = \{[\Sigma(p_i + n_i)] - 1\}/\Sigma B_i$	Kendall and Moran 1963, James 1971
13. CI, NN, 2N search areas pooled (i.e., pooled with search area to 2N [2P])	$KM2P = \{[\Sigma(p_i + n_i + m_i)] - 1\}/\Sigma C_i$	Kendall and Moran 1963
T-Square (TS) estimators		
14. Basic T^2 estimator (BA)	$TSBA = 2N/[\pi \Sigma R_{(1)i}^2 + 0.5\pi \Sigma T_i^2]$	Diggle 1975
15. Reduced bias (RB) in aggregated populations	$TSRB = N/\pi[(\Sigma R_{(1)i}^2)(0.5 \Sigma T_i^2)]^{1/2}$	Diggle 1975
16. Robust (Byth [B])	$TSB = N^2/[(2 \Sigma R_{(1)i})(\sqrt{2})(\Sigma T_i)]$	Byth 1982
Ordered Distance (OD) estimators		
17. Closest individual	$ODCI = (N - 1)/\pi \Sigma(R_{(1)i})^2$	Morisita 1957, Pollard 1971
18. Second closest individual (2C)	$OD2C = (2N - 1)/\pi \Sigma(R_{(2)i})^2$	Morisita 1957, Pollard 1971
19. Third closest individual (3C)	$OD3C = (3N - 1)/\pi \Sigma(R_{(3)i})^2$	Morisita 1957, Pollard 1971
Angle-Order (AO) estimators		
20. Point-centered-quarter (i.e., 1 observation per quadrant [1Q])	$AO1Q = 12N/\pi \Sigma 1/R_{(1)ij}^2$	Stearns 1949, Cottam et al. 1953, Cottam and Curtis 1956, Morisita 1957, Pollard 1971
21. Second closest individual in each quadrant (2Q)	$AO2Q = 28N/\pi \Sigma 1/R_{(2)ij}^2$	Morisita 1957, Pollard 1971
22. Third closest individual in each quadrant (3Q)	$AO3Q = 44N/\pi \Sigma 1/R_{(3)ij}^2$	Morisita 1957, Pollard 1971
23. Third closest individual in each quadrant (3)	$AO3 = [2/\pi N] \Sigma \Sigma(1/R_{(3)ij})$	Morisita 1971

Performance of existed DPDEs are low.

Main Objectives

- Propose two estimators to generalize most existing distance-based population density estimators.
- Propose a new estimator based on non-randomly distributed populations
- Test the new estimators by simulation and real forest data.

Three new general population density estimators

The 1st general DPDEs

I. The general simple DPDE

$$\hat{\lambda}_1 = \frac{q(\mathrm{nqk}-1)}{\pi \sum_{i=1}^{n} \sum_{j=1}^{q} r_{kij}^2}$$

i: i^{th} focal point/event;

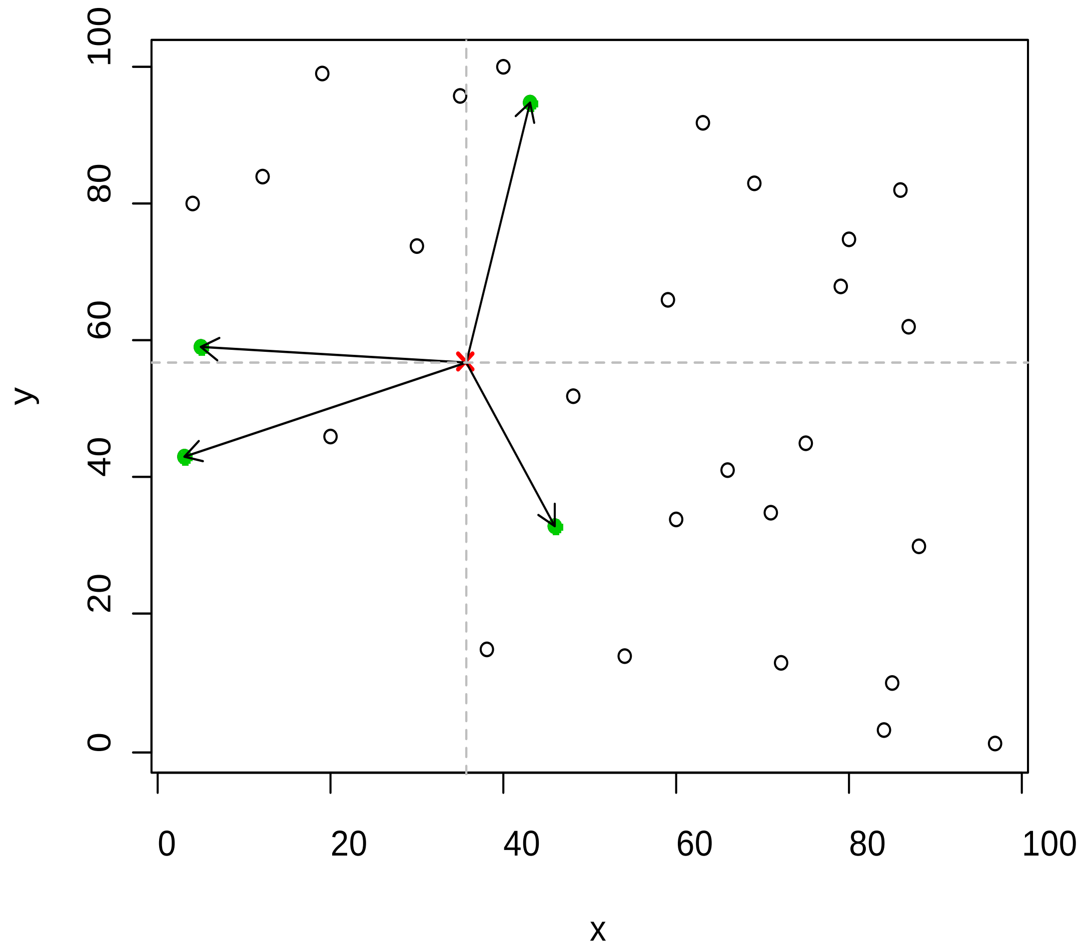
n: total # focal point/e;

j: j^{th} equal angle sector;

q: total # sectors centered at one focal point/event;

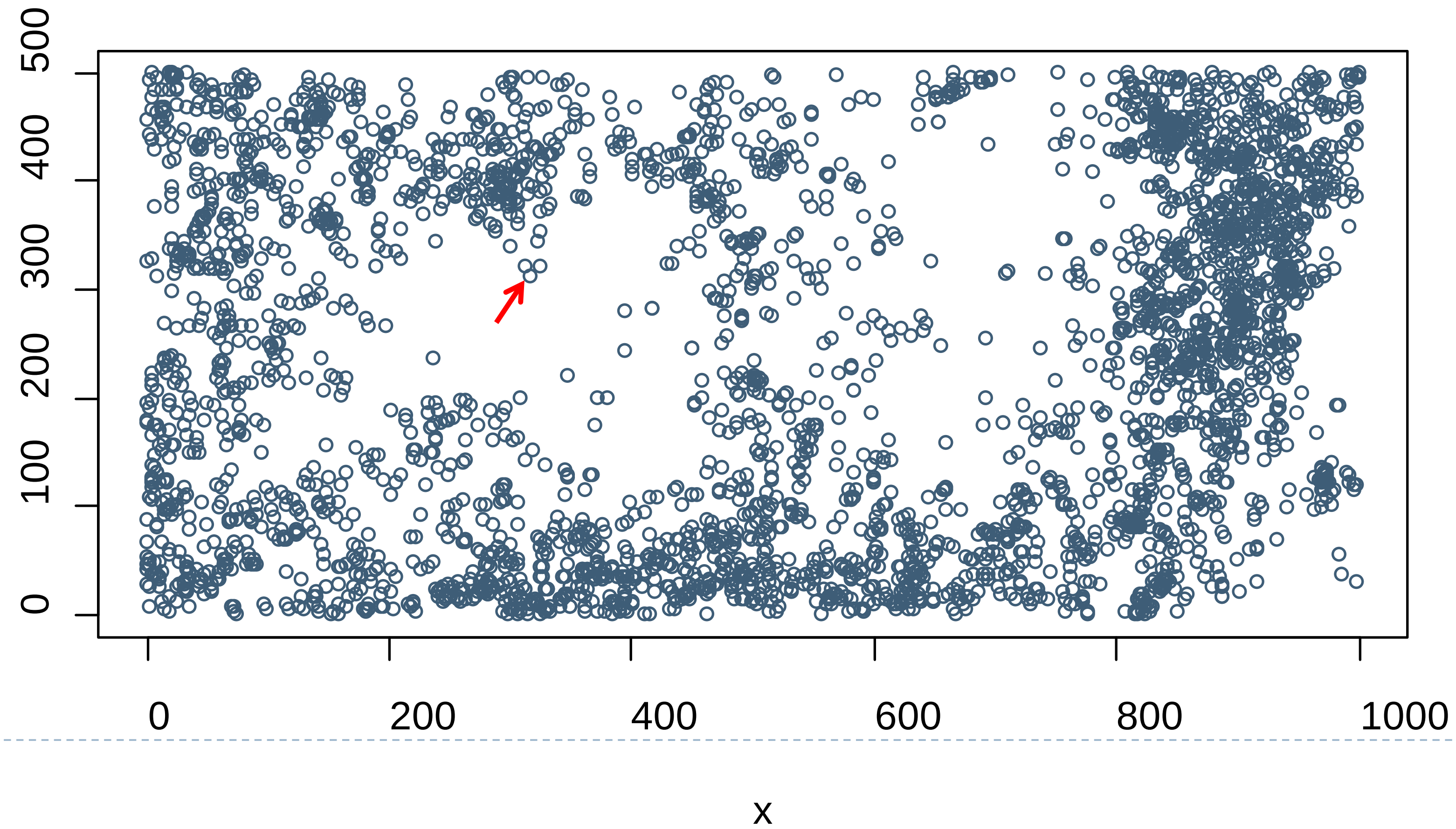
k: k^{th} nearest neighbor;

$$\mathrm{Var}(\hat{\lambda}_1) = \frac{\lambda^2}{nqk-2}$$



The 1^{st} general DPDEs

- Point-to-event distances used in $\hat{\lambda}_1$ tend to be increased by aggregation and decreased by regularity in the underlying pattern; while the reverse is true for event-to-event distances.



The 2nd general DPDEs

2. The general composite DPDE

$$\hat{\lambda}_c = \sqrt{\hat{\lambda}_e \hat{\lambda}_p}$$

Substitute by $\hat{\lambda}_1$ and corrected it to a unbiased form

$$\hat{\lambda}_c = \frac{q\Gamma^2(\frac{1}{2}nqk)}{\pi\Gamma^2(\frac{1}{2}nqk - \frac{1}{2})\sqrt{\sum_{i=1}^{qn/2} \mathrm{r}_{kie}^2 \sum_{j=1}^{qn/2} \mathrm{r}_{kjp}^2}}$$

r_{kie} and r_{kjp} are distance from event to event and point to event respectively.

The 2nd general DPDEs

$$\mathrm{Var}\left(\hat{\lambda}_{c1}\right) = \lambda^2 \left(\frac{\Gamma^4\left(\frac{1}{2}nqk\right)}{\left(\frac{1}{2}nqk - 1\right)^2 \Gamma^4\left(\frac{1}{2}nqk - \frac{1}{2}\right)} - 1 \right)$$

- However, the first general simple DPDE and the second general composite DPDE are both derived from random population.

The 3rd general DPDE

- If we assume number of individuals in each sector with radii r_{kij} follows negative binomial distribution, then we can estimated population density λ by maximizing the following likelihood function:

$$\hat{\lambda}_{n\_ptoe} = \frac{q(2k-1)\sum_{m=1}^{nq} r_{m\_ptoe}^{-1}}{\pi \sum_{m=1}^{nq} r_{m\_ptoe}} - \frac{nq^2 k}{\pi \sum_{m=1}^{nq} r_{m\_ptoe}^2}$$

Performance criterion

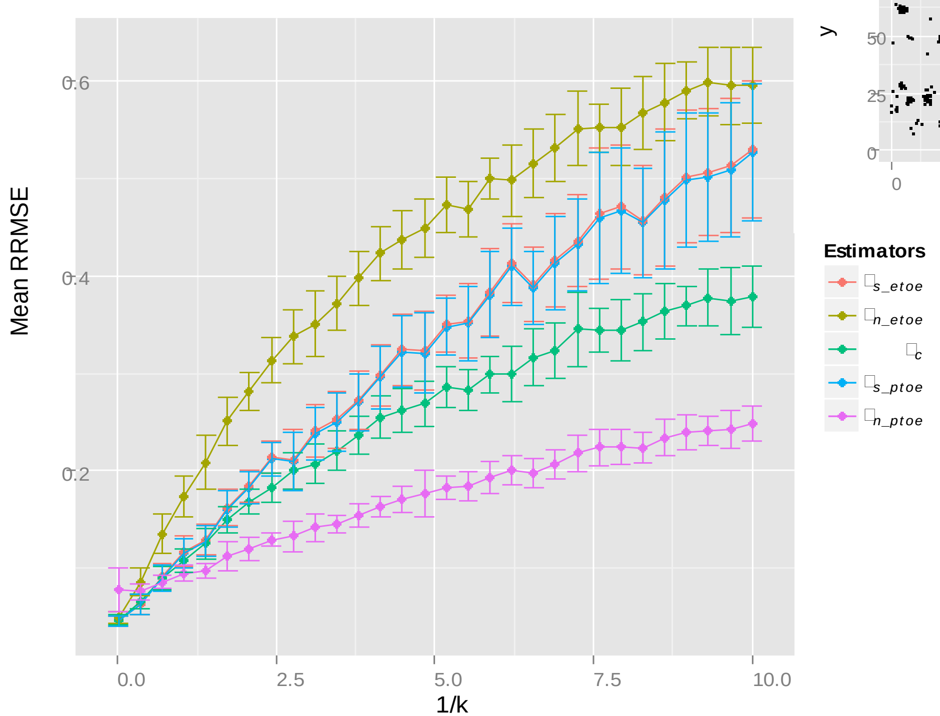
Relative root-mean-squared-error (RRMSE)

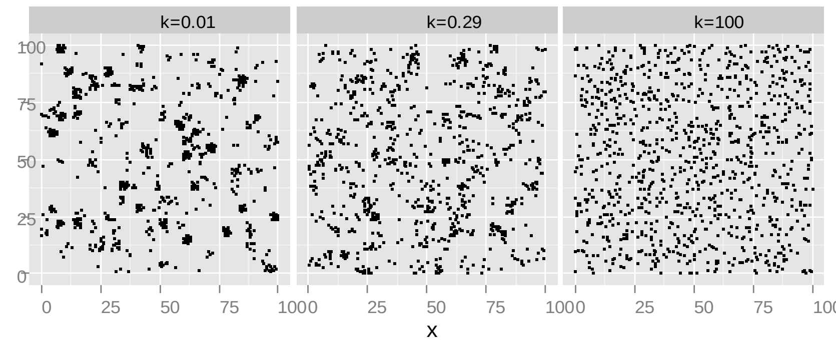
$$\text{RRMSE} = \sqrt{\frac{\sum_{\text{i}}^{\text{c}}\left(\lambda - \hat{\lambda}_{\text{i}}\right)^2}{c\lambda^2}}$$

where c is the total number of estimates.

Performance comparison

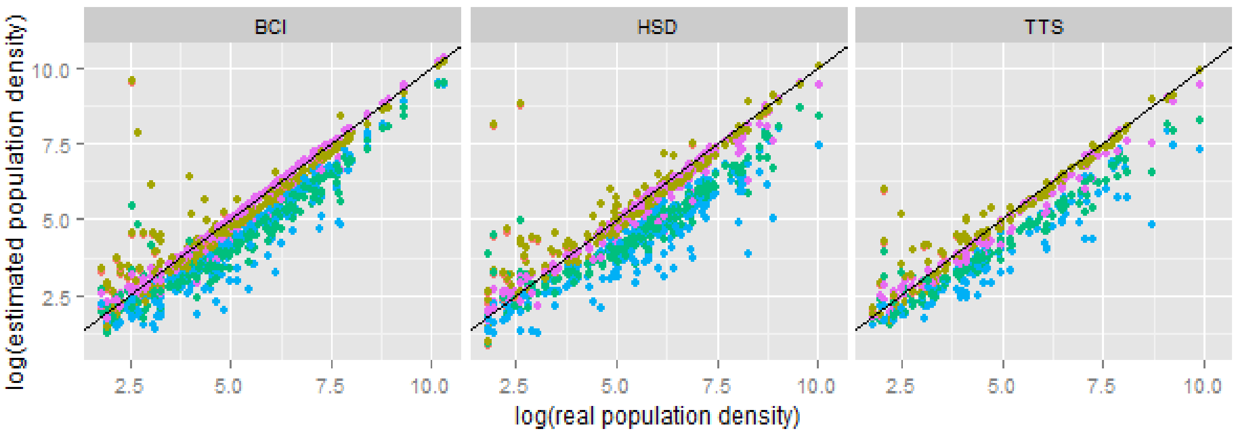
Performance Comparison by simulation

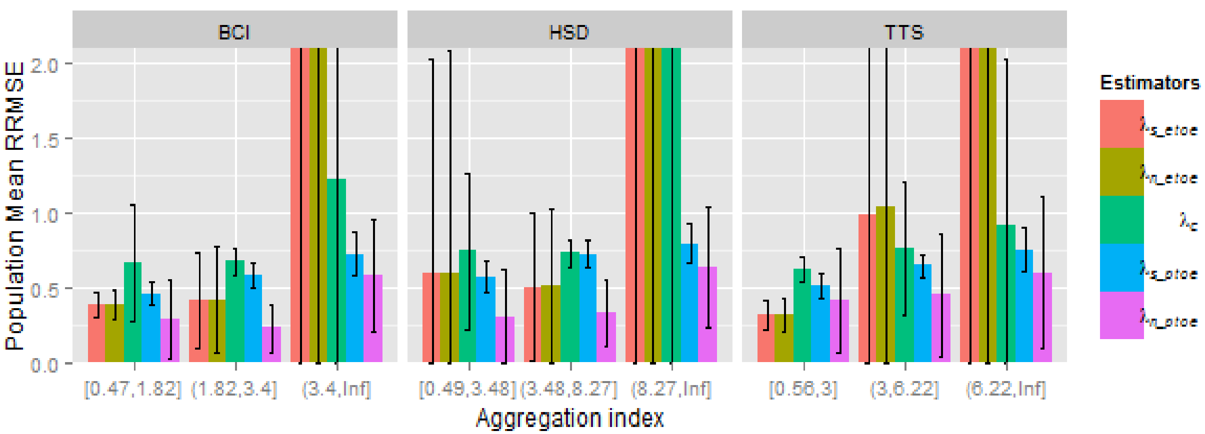




The **nonrandom estimator based on point to event distances** has the highest performance.

Performance Comparison by real data



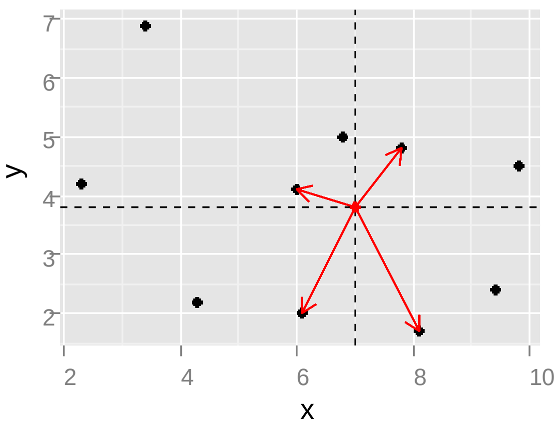


The **nonrandom estimator based on point to event distances** has the highest and most consistent performance.

Estimation of population density

Point-to-event distance samples

The general nonrandom distance-based population density estimator



$$\hat{\lambda}_{n\_ptoe} = \frac{q(2k-1)\sum_{m=1}^{nq} r_{m\_ptoe}^{-1}}{\pi \sum_{m=1}^{nq} r_{m\_ptoe}} - \frac{nq^2 k}{\pi \sum_{m=1}^{nq} r_{m\_ptoe}^2}$$

Efficient & robust

Thanks!