

# A robust stepwise clustering approach to detect individual trees in temperate hardwood plantations with varied stand densities

Gang Shao <sup>1,\*</sup>, Songlin Fei <sup>2</sup> and Guofan Shao <sup>2</sup>

<sup>1</sup> Libraries and School of Information Studies, Purdue University; West Lafayette, IN, 47907, United States

<sup>2</sup> Forestry and Natural Resources, Purdue University; West Lafayette, IN, 47907, United States

\* Correspondence: gshao@purdue.edu

## Supplementary materials

### Supplement figures

Figure S1. Linear correlation between tree number and LiDAR metrics in northern red oak stand (NT: tree number; PtsD: average point density metric; MeanH: mean height metric; MaxH: maximum height metric; Gap: gap fraction metric).

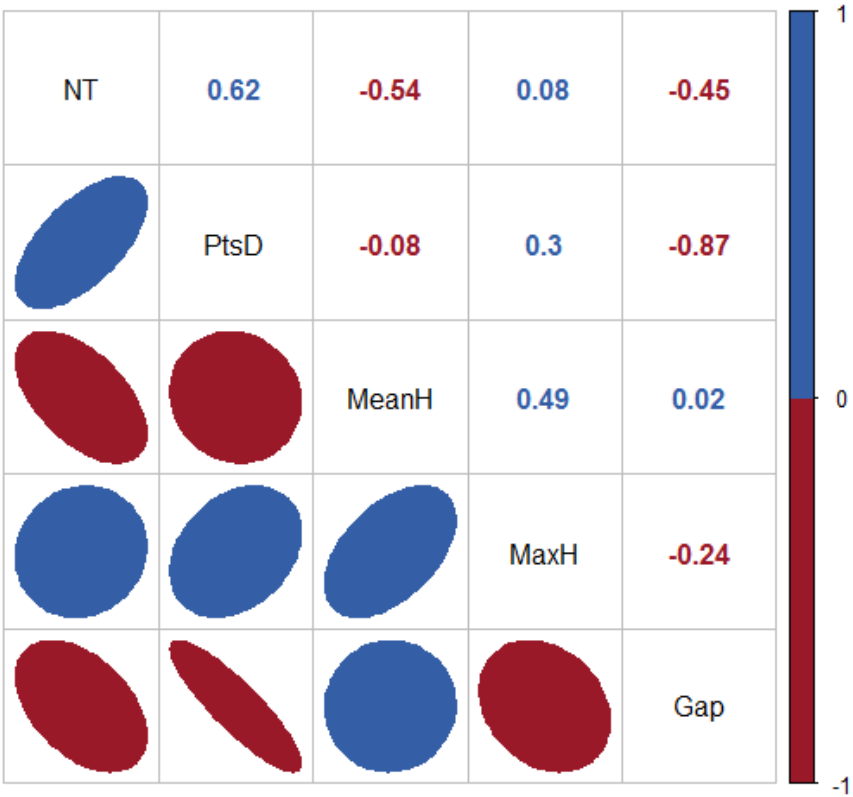


Figure S2. The boxplot shows tree cluster estimation distribution over 100 iterations in each block. Subplot 1 (on top) presents the results using constant kernel HMeanShift clustering; Subplot 2 (at bottom) shows the results using dynamic kernel HMeanShifter clustering.

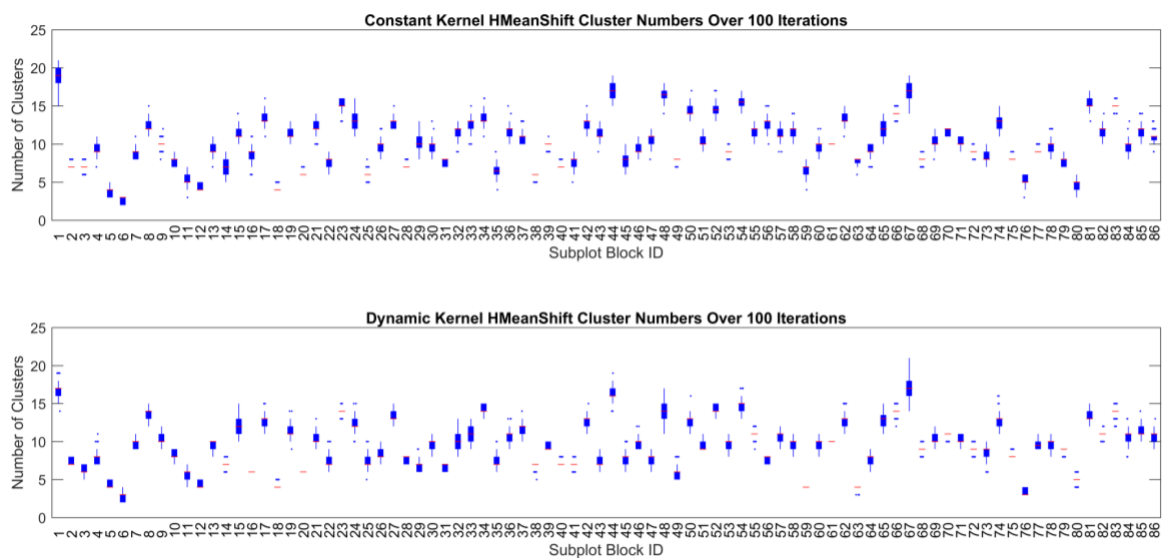


Figure S3. Individual tree maps in the northern red oak stand.

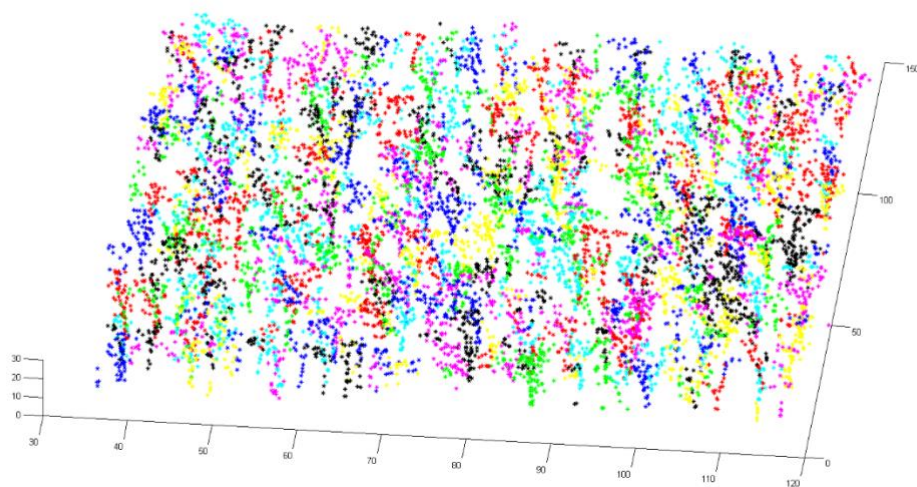
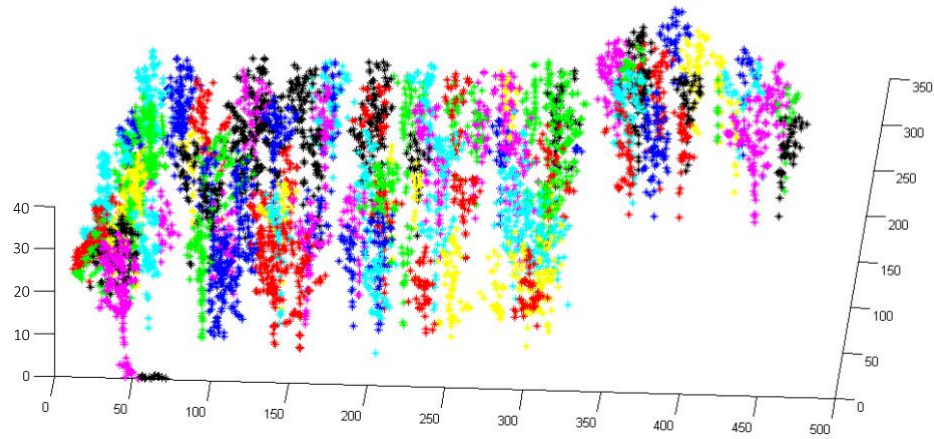


Figure S4. Individual tree maps in the black walnut stand.



### Supplement tables

Table S1. Tree detection accuracy comparison table.

**C-S:** Constant-kernel Stepwise Approach; **D-S:** Dynamic-kernel Stepwise Approach;

**C-H:** Constant-kernel Single-step HMeanShift; **D-H:** Dynamic-kernel Single-step HMeanShift;

**T-W:** Treetop-based Watershed Approach; **P-W:** Point-density-based Watershed Approach.

	Detection Rate						Detection Accuracy						Estimated Percentage					
	C-S	D-S	C-H	D-H	T-W	P-W	C-S	D-S	C-H	D-H	T-W	P-W	C-S	D-S	C-H	D-H	T-W	P-W
H-oak	0.93	<b>0.94</b>	0.86	0.91	0.78	0.89	0.95	<b>0.96</b>	0.92	0.94	0.81	0.90	0.95	0.95	0.88	0.94	0.92	<b>0.98</b>
M-oak	<b>0.91</b>	0.90	0.87	<b>0.91</b>	0.82	0.88	<b>0.94</b>	0.93	0.90	0.91	0.82	0.87	0.93	0.92	0.94	<b>1.00</b>	<b>1.00</b>	1.02
L-oak	<b>0.93</b>	0.92	0.92	0.93	0.96	0.90	0.89	<b>0.90</b>	0.77	0.81	0.75	0.82	1.08	<b>1.04</b>	1.37	1.30	1.56	1.19
Oak	<b>0.92</b>	0.91	0.88	0.91	0.84	0.88	<b>0.93</b>	<b>0.93</b>	0.88	0.90	0.80	0.87	0.97	0.95	<b>1.01</b>	1.04	1.09	1.04
Walnut	0.94	<b>0.96</b>	0.94	<b>0.96</b>	0.91	0.97	<b>0.96</b>	0.94	0.93	0.95	0.91	0.97	0.96	1.03	1.03	1.03	<b>1.00</b>	<b>1.00</b>
Total	<b>0.92</b>	<b>0.92</b>	0.89	<b>0.92</b>	0.85	0.90	<b>0.94</b>	0.93	0.89	0.91	0.82	0.89	0.97	0.97	<b>1.01</b>	1.04	1.07	1.03

Table S2 t-Test for tree heights between single-step HMeanShift using constant (C-H) and dynamic kernels (D-H).

	C-H	D-H
Mean	19.93	19.94
Variance	8.57	8.94
Observations	58	58
Pearson Correlation	0.99	

Hypothesized Mean Difference	0
df	57
t Stat	-0.20
P(T<=t) one-tail	0.42
t Critical one-tail	1.67
P(T<=t) two-tail	0.84
t Critical two-tail	2.00

Table S3 t-Test for tree heights between stepwise approach (C-S) and single-step HMeanShift (C-H) using constant kernel.

	C-S	C-H
Mean	19.88	19.93
Variance	8.71	8.57
Observations	58	58
Pearson Correlation	0.99	
Hypothesized Mean Difference	0	
df	57	
t Stat	-1.05	
P(T<=t) one-tail	0.15	
t Critical one-tail	1.67	
P(T<=t) two-tail	0.30	
t Critical two-tail	2.00	

Table S4 t-Test for tree heights between stepwise approach (D-S) and single-step HMeanShift (D-H) using dynamic kernel.

	D-S	D-H
Mean	19.88	19.94
Variance	8.71	8.94
Observations	58	58
Pearson Correlation	0.99	
Hypothesized Mean Difference	0	
df	57	
t Stat	-1.25	
P(T<=t) one-tail	0.11	
t Critical one-tail	1.67	
P(T<=t) two-tail	0.22	
t Critical two-tail	2.00	

Table S5 t-Test for crown size between single-step HMeanShift using constant (C-H) and dynamic kernels (D-H).

	C-H	D-H
Mean	3.71	3.79
Variance	3.04	3.35
Observations	58	58
Pearson Correlation	0.94	
Hypothesized Mean Difference	0	
df	57	
t Stat	-0.97	
P(T<=t) one-tail	0.17	
t Critical one-tail	1.67	
P(T<=t) two-tail	0.34	
t Critical two-tail	2.00	

Table S6 t-Test for crown size between stepwise approach (C-S) and single-step HMeanShift (C-H) using constant kernel.

	C-S	C-H
Mean	4.05	3.71
Variance	3.56	3.04
Observations	58	58
Pearson Correlation	0.88	
Hypothesized Mean Difference	0	
df	57	
t Stat	2.91	
P(T<=t) one-tail	0.003	
t Critical one-tail	1.67	
P(T<=t) two-tail	0.01	
t Critical two-tail	2.00	

Table S7 t-Test for crown size between stepwise approach (D-S) and single-step HMeanShift (D-H) using dynamic kernel.

	D-S	D-H
Mean	4.05	3.79
Variance	3.56	3.35
Observations	58	58
Pearson Correlation	0.87	

Hypothesized Mean Difference	0
df	57
t Stat	2.11
P(T<=t) one-tail	0.02
t Critical one-tail	1.67
P(T<=t) two-tail	0.04
t Critical two-tail	2.00

Table S8 Time efficiency analysis of point cloud based mean shift clustering.

Area (Acre)	Window Size (foot)	Points	NO. of Trees	Time (second)	Iterations for 5165 acres	Total time (second)
0.092	200	6543	103	0.6	5625	3375
3.673	400	27574	367	9.83	1444	14194.52
8.264	600	45809	668	33.49	625	20931.25
14.692	800	81030	1070	137.14	361	49507.54
22.957	1000	109383	1379	279.05	225	62786.25