Quantifying Geometric Point Disorder in Geospatial Data with a Novel Algorithm

MTSU CBAS Scholar's Week 2020

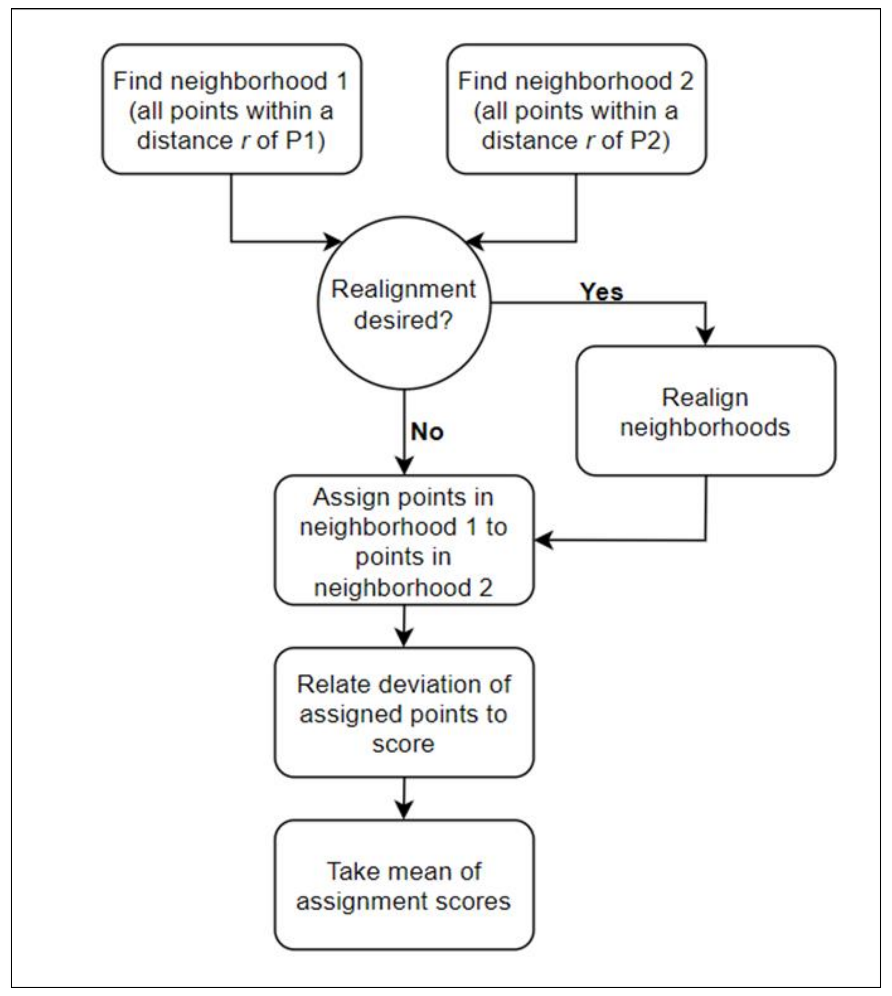
Sky Jones - *Middle Tennessee State University*

Dr. Henrique Momm - *Middle Tennessee State University*

Main result: The randomness of the arrangement of points can be quantified by comparing the change between the neighborhoods surrounding points, and this can be used to identify structures as deliberately or randomly placed.

BACKGROUND: Geometric spatial order in geographic data provides clues about possible anthropogenic provenance, but such disorder is understudied in geographic contexts. Current characterization schemes of geometric disorder are either limited to raster (pixel-based) data, do not provide an index that quantifies relative disorder, or make restrictive assumptions about what constitutes "order". We propose a new algorithm and related measure, the Index of Disorder (IoD), that accurately captures relative geometric disorder in coordinate-based data without making assumptions of underlying structure.

ALGORITHM OVERVIEW

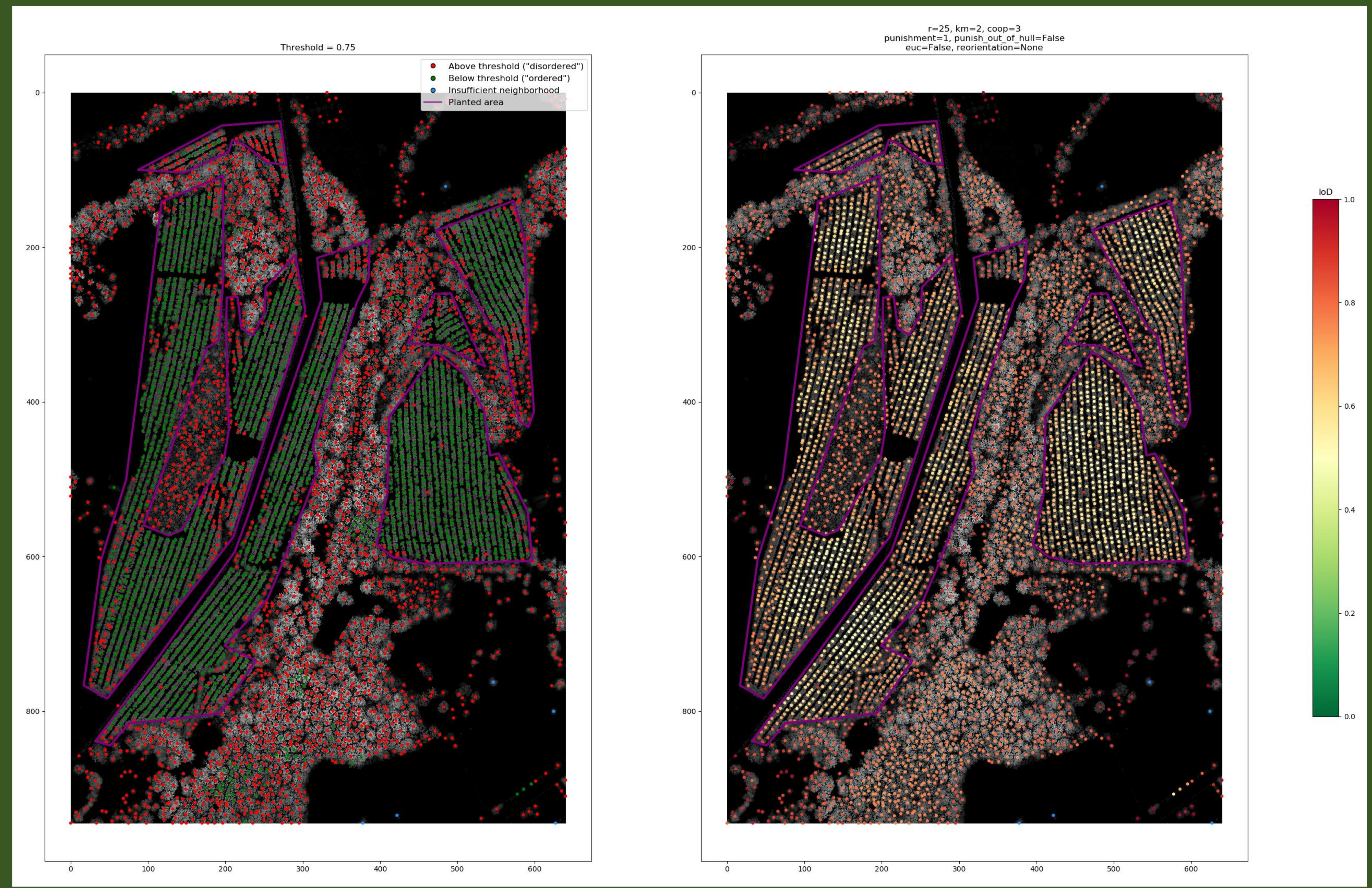


RESULTS

- The IoD is capable of distinguishing ordered and disordered points in synthetic datasets, though differentiation is greatest for grid-derived patterns
- The IoD on its own is sufficient to distinguish naturally occurring trees from planted stands such as orchards or reforested zones
- Geometric disorder may not be sufficient to classify buildings as "main" or "auxiliary", but the IoD can enrich such datasets for more comprehensive classification schemes
- The scale-dependent nature of the IoD allows quantification of pattern scale by using the IoD in reverse (i.e., optimizing IoD classification accuracy via parameter tuning)

Study Area	Disordered Group	Ordered Group	Peak κ	Corresponding Accuracy
Rowell's Apple House near Crab Orchard, TN	Natural forest	Orchard	0.81	0.96
Reforested area near Atwell Airport, Mooresville, NC	Natural forest	Reforested zone	0.74	0.87
Lockeland Springs neighborhood in East Nashville, TN	Auxillary structures (sheds, detached garages)	Major buildings (homes and commerical buildings)	0.44	0.76

Note: Cohen's κ is used to measure inter-rater reliability for classification agreement that takes into account the possibility of the agreement occurring by chance. 0.81 = excellent agreement, 0.74 = substantial agreement, 0.44 = moderate agreement



Tree crowns in Mooresville, NC. Purple polygons indicate replanted (anthropogenic) tree stands. The IoD was used to quantify the relative geometric disorder of tree placement (right), and trees were classified as being natural/disordered or planted/ordered depending on whether their IoD exceeded a threshold value of 0.75.

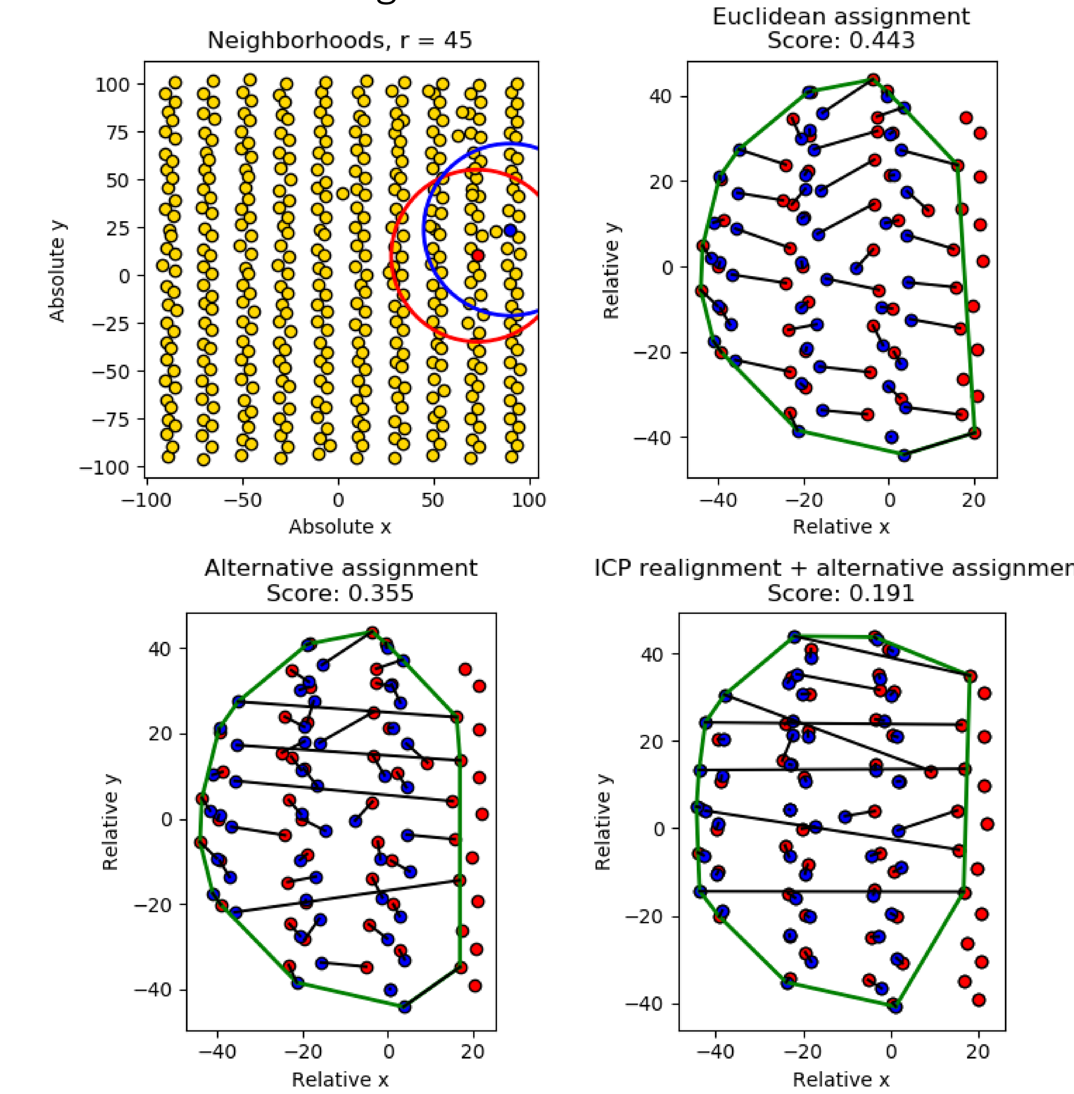
METHODS:

The IoD was applied to 12 synthetic data sets. Each of these synthetic datasets were generated by perturbing a pattern with noise of increasing strength outside an arbitrary distance from the origin. Thus, the order of the pattern is entirely preserved within the arbitrary radius, but the system becomes increasingly disordered beyond that radius. A variety of patterns were created, including simple grids, grids transformed with sinusoidal signals, overlapping grids, radial patterns, and frieze (wallpaper) groups. The IoD was applied to each dataset to quantify relative geometric disorder.

LiDAR (Light Detection and Ranging) point cloud datasets were for the area surrounding an orchard in Crab Orchard, TN and a replanted forest in Mooresville, NC. Digital elevation products were derived from these data, and from these elevation products tree crowns were extracted using a blob detection algorithm. The IoD was applied to these tree crown datasets and trees were classified as being natural/disordered or planted/ordered depending on their IoD value. Algorithm parameters were subjected to sensitivity testing, and the classification accuracy and Cohen's κ were calculated to quantify classification quality.

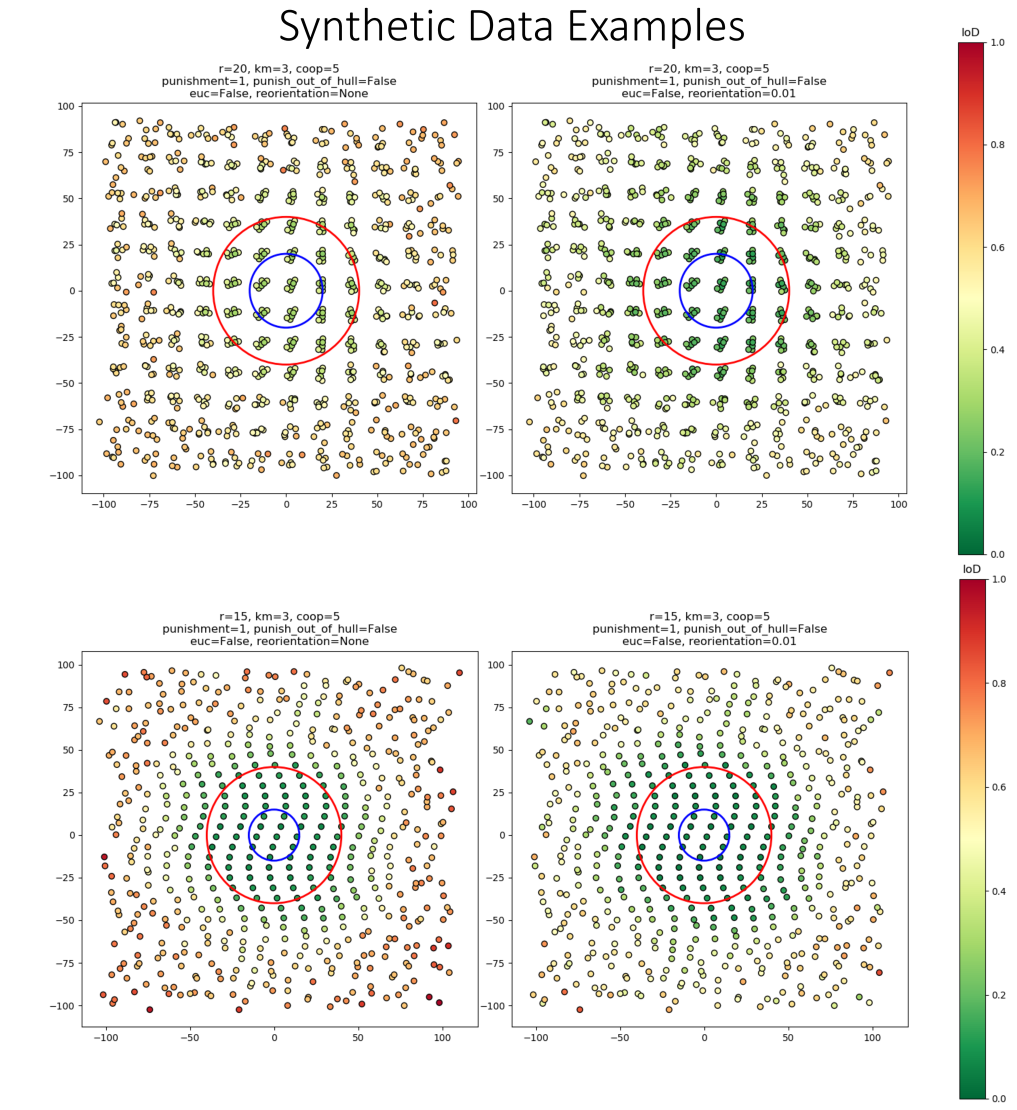
Building footprints were obtained for the metropolitan Nashville. The IoD was calculated for extracted centroids, and buildings were classified as "disordered" if their IoD was above an arbitrary IoD threshold and "ordered" if below the threshold. "Ordered" buildings were assumed to correspond to "main" buildings, while disordered buildings were assumed to correspond to auxiliary structures. Algorithm parameters were subjected to sensitivity testing, and the classification accuracy and Cohen's κ were calculated to quantify classification quality.

Algorithm Visualization



Visualization of the neighborhood comparison portion of the IoD algorithm. The neighborhood radius is shown for both the red and blue points in the upper left figure. The remaining three figures explore various point assignment methods and the calculated IoD. Each assignment subfigure shows the neighborhood corresponding to the red point in red and the neighborhood corresponding to the blue point in blue. Black lines connect a red neighborhood point to the blue neighborhood point it is assigned to under the applied assignment scheme. Euclidean assignment attempts to minimize the sum of the lengths of the black lines. The alternative assignment scheme attempts to minimize the sum of the scores of the assignment lengths using a sigmoidal scoring function. The ICP (iterative closest point) realignment iteratively rotates, scales and translates the blue neighborhood to attempt to minimize the assignment cost. The green polygon represents the convex hull of the assigned points. In this example, unassigned points outside the convex hull are ignored for the purpose of scoring while unpaired points within the convex hull are penalized with a score of 1; paired points are scored using the function from Figure 1 regardless of the assignment scheme used.

Synthetic Data Examples



Results from applying two versions of the IoD algorithm without and with realignment (left and right respectively) to synthetic patterns. The pattern is perturbed with random noise that increases in strength as a function of distance beyond the red circle, and the blue circle represents the neighborhood size. Realignment enables better detection of discrete patterns, such as the top pattern, but often reduces differentiation for continuous patterns, such as the pattern on the bottom.

Acknowledgements

This project was made possible thanks to grants from the National Resources Conservation Service as well as MTSU's Undergraduate Research Experience and Creative Activity (URECA) program.

References

R. M. Haralick, K. Shanmugam and I. Dinstein, "Textural Features for Image Classification," IEEE Transactions on Systems, Man, and Cybernetics, Vols. SMC-3, no. 6, pp. 610 - 621, 1973.

D. Phiri and J. Morgenroth, "Developments in Landsat Land Cover Classification Methods: A Review," Remote Sensing, vol. 9, no. 9, 2017.

S. Aksoy, I. Z. Yalniz and K. Tasdemir, "Automatic Detection and Segmentation of Orchards Using Very High Resolution Imagery," IEEE Transactions on Geoscience and Remote Sensing, vol. 50, no. 8, pp. 3117-3131, 2012.

NRCS
Natural Resources Conservation Service

MIDDLE TENNESSEE STATE UNIVERSITY