

A Local Version of the Modified Moran's I

J. Chen^{1,*}, M. Jackson^{**}

American University, 4400 Massachusetts Ave NW, Washington, DC 20016

Abstract

Local indicators of spatial association (LISAs) are a class of tests that detect clusters and outliers in spatial data. Existing LISAs have statistical issues including no clear null distribution, fixed neighborhood size, and multiple testing issues. We derive and implement a new LISA - the Local Modified Moran's I (LMMI) - as a local version of Jackson et al.'s (2010) Modified Moran's I. We address many of the aforementioned issues by using Monte Carlo methods to construct an empirical distribution, using flexible neighborhood sizes, and applying Tango's (2000) multiple testing correction for global indicators to local indicators. To assess the performance of the LMMI relative to other tests, we simulate data with known clusters of various sizes and background populations. We find the LMMI performs better than Anselin's (1995) Local Moran's I when searching for larger clusters. The LMMI has both higher precision (detects a higher percentage of the simulated cluster) and lower false positive rate (false positives among all positives) in this scenario.

Keywords: Local indicators of spatial association, Local Moran's I, clustering, ESDA

1. Introduction

Tobler's First Law of Geography posits that, "Everything is related to everything else, but near things are more related than distant things" [4].

*Principal corresponding author

**Corresponding author

Email addresses: jess_chen@redlands.edu (J. Chen), monica@american.edu (M. Jackson)

¹Present Address: University of Redlands, 1200 E Colton Ave, Redlands, CA 92374

Spatial analysis is an array of statistical techniques that use geographic information to better understand the pattern of observed attribute values and the processes that generated the observed pattern. Simply put, it is the study of processes and relationships that occur over space - space that is typically interpreted as geographical, but has also been applied to more figurative notions such as social networks, socioeconomic similarity and interpersonal distance. The underlying motivation for spatial analysis is the possibility of spatial autocorrelation or spatial heterogeneity. Spatial autocorrelation occurs when a process transmits or “infects” from one data point to another across space. Spatial heterogeneity occurs when a process varies by spatial location due to some environmental heterogeneity. Spatial analysis has long been employed in applications such as cartography, geography, and epidemiology. More recently, increasing availability of micro-level spatial data, higher computing power and a wider range of programs to analyze spatial data (e.g. ArcGIS, spatial packages in R and Python) have garnered renewed interest and work to improve spatial analysis methods. This work has tended to evolve along separate strands of literature, including spatial econometrics (such as urban economics), epidemiology (such as disease mapping), and spatial statistics.

Spatial correlation in data requires a special class of statistical tests to detect spatial clusters. Spatial cluster tests span a wide swath of methods with five common characteristics [5]:

1. They quantify a spatial pattern.
2. They specify an alternative hypothesis for the spatial pattern the test is designed to detect.
3. They specify a null hypothesis of a spatial pattern when the alternative hypothesis is false.
4. They generate a reference distribution consistent with the null hypothesis.
5. They can be compared to the reference distribution to assess the probability of the null hypothesis being true.

Within spatial cluster tests, there is a distinction between global indices of spatial association and local indicators of spatial association. Global indices of spatial association identify the overall propensity of events in the entire study area to cluster, relative to the specified null hypothesis (e.g. constant risk of occurrence across the study area). Local indicators of spatial association (LISAs) are derived from global indices and identify a collection of

events that is more clustered than would be expected under the specified null hypothesis [3]. Both types of indices are used in many applications, including public health, the pure sciences, geography and social sciences such as economics and public policy. They have been used to understand regional trends of voting [6], to show the changing socio-economic distribution of states [7], and to determine environmental causes of diseases [8, 9, 10].

The test developed in this paper falls within the class of LISAs. There are several significant statistical issues regarding inference using current LISAs. Difficulties include no clear null distribution of the statistics, population heterogeneity problems [11], fixed neighborhood sizes, multiple testing issues [2], and spatial correlation of data and test statistics [12]. Many spatial statistics do not have a clear reference distribution (e.g. normal distribution) due to the idiosyncratic nature of spatial data. Heterogeneous populations across the regions of study complicate statistical inference for both count and rate data by making estimates unstable. If each region has a different number of people at risk, this yields different expectations and variances (and hence reference distributions) for each region. LISAs require specifying *a priori* the size of “neighborhood” with which to examine for clusters. This effectively fixes the cluster size in the alternative hypothesis, which is clearly limiting. The multiple testing problem arises when test statistics are derived for each spatial neighborhood, and inference must be made simultaneously on which one is the most likely cluster. The problem also arises when multiple tests are conducted using different neighborhood sizes in order to determine the most likely number of clusters along with the size of clusters. Finally, many LISAs rely on constructing test statistics from various neighborhoods of data points. To the extent that these neighborhoods overlap, local test statistics may be correlated across space.

We develop a local version of Jackson et al.’s (2010) Modified Moran’s I to address some of these issues, namely the problems of no clear null distribution, multiple testing, and fixed neighborhood sizes. Modified Moran’s I is a global indicator of spatial association and has been shown to have higher statistical power than existing global tests.² It improves upon Moran’s I by not only including the spatial weight function in the calculation of differences

²Jackson et al. [1] show that Modified Moran’s I offers a 5-10% power improvement over Moran’s I and Oden’s I^{pop} for evaluating global and local clustering patterns on geographic data with homogeneous populations.

in cases between regions, but also in the calculation of variance between regions. We implement a local version of the Modified Moran's I using Monte Carlo methods for inference. This addresses the problem of no clear distribution by simulating a separate probability distribution for each location. We also specify various neighborhood sizes rather than being limited to one fixed size. This leads to the problem of multiple testing. Tango [2] proposes an adjustment to his global test of clustering to correct for multiple testing when using variable neighborhood sizes. We extend his method to local tests. We use simulated data to test the performance of the Local Modified Moran's I and compare it to existing tests. The simulations show that the Local Modified Moran's I performs at least as well as existing tests for small clusters, and it performs better for larger clusters.

2. Local Moran's I

The LISA methodology was pioneered by Anselin [3], who developed a local version of Moran's I [13]. Moran's I is a workhorse in tests of global clustering, indicating the overall tendency of a study area to exhibit either a clustering or regular pattern. It can be thought of as a spatially-weighted version of Pearson's correlation coefficient. Specifically, Moran's I applies a set of spatial weights (w_{ij}) to the cross-product of absolute deviations ($y_i - \bar{y}$) for each data point:

$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_{j:i \neq j} w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\sum_i (y_i - \bar{y})^2} \quad (1)$$

Anselin [3] implements the Local Moran's I, which he categorized as a local indicator of spatial association. Anselin [3] defines LISAs according to two criteria:

1. “The LISA for each observation gives an indication of the extent of significant spatial clustering of similar values around that observation”; and
2. “The sum of LISAs for all observations is proportional to a global indicator of spatial association.”

The Local Moran's I assesses the contribution of each data point to the

global Moran's I:

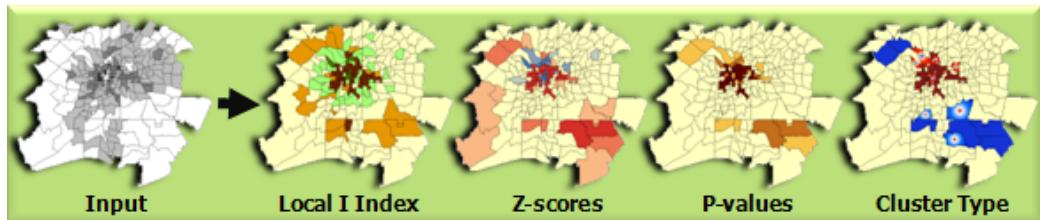
$$I_i = \frac{y_i - \bar{y}}{S_i^2} \sum_{j=1, j \neq i}^N w_{ij}(y_j - \bar{y}) \quad (2)$$

Where:

$$S_i^2 = \frac{\sum_{j=1, j \neq i}^N w_{ij}}{N-1} - \bar{y}^2 \quad (3)$$

Anselin [3] identifies the most likely clusters by comparing their contribution against a null distribution. Figure 1 illustrates an example of a cluster detected by Local Moran's I, using ArcGIS. Local Moran's I is calculated on regional data, with a separate statistic for each region. Z-scores and p-values are calculated based on the value of the statistic against a null distribution. The type of cluster is based on the Local Moran's I scores and the original value of the data point and that of surrounding regions.

Figure 1: Sample Cluster Detected by Local Moran's I



Source: Environmental Systems Research Institute [14]

Many tests to detect local clusters have been developed in the vein of Anselin's original test, some of which are classified as LISAs and some that take the approach of "scanning local rates." This latter class of tests identifies local clusters by constructing circles of varying radii, calculating rates within each circle, and selecting circles that exceed a defined threshold.

3. A Local Version of Modified Moran's I

We derive the local version of the Modified Moran's I and construct a procedure for inference.³ To assess the performance of the new LISA alongside

³R code for the Local Modified Moran's I and simulation study setup is available here: [Insert file Simulation_Code.ZIP here](#).

other tests, we simulate data under a host of cluster scenarios and compare the tests on several evaluation criteria.

The proposed test statistic is based on Jackson et al.'s (2010) improvement of Moran's I, which substitutes the variance component of Moran's I into a spatially-weighted covariance function. The Modified Moran's I is:

$$I_w = \frac{\sum_i \sum_{j:i \neq j} w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\sum_{i \leq i < j \leq N} w_{ij}(y_i - y_j)^2} \quad (4)$$

The weight function is not only included in the differences of regional cases (e.g. county cases) from the overall mean, but also in the calculation of the variance. Intuitively, local spatial relationships should not only contribute to the product of absolute deviations, but also to the variation between points. Jackson et al. [1] show through simulation that Modified Moran's I achieves higher power (around a 5-10% improvement) than Moran's I and Oden's I*pop for evaluating global and local clustering patterns on geographic data with homogeneous populations.

We derive and implement a local version of Modified Moran's I based on Anselin's (1995) Local Moran's I. The Local Modified Moran's I can be derived by breaking out the summation by each observation i :

$$I_{w,i} = \frac{(y_i - \bar{y}) \sum_j^N w_{ij}(y_j - \bar{y})}{\sum_j^N w_{ij}(y_i - y_j)^2} \quad (5)$$

This statistic yields the contribution of each data point to the global index of clustering. Local "hot spots" or clusters are those points with a statistically significant $I_{w,i}$. In the absence of a clear null distribution, we construct an empirical reference distribution for each location in order to conduct inference. The primary approach is to derive pseudo-significance levels [3]. This requires holding the value of y_i constant for each calculation of $I_{w,i}$ and randomizing the remaining values across observed locations (i.e. random labeling). Each iteration yields a value of $I_{w,i}$ that can be collected via Monte Carlo simulation to construct a distribution for each location's $II_{w,i}$.

The spatial weight matrix can be specified several ways, and results of cluster tests are very sensitive to the selection of the weight matrix. Thus, it is important to define the weight matrix according to the theory behind the spatial transmission mechanism. For the purposes of statistical testing, we use a row-standardized exponential distance decay matrix. A distance decay

matrix is more nuanced than a contiguity matrix, which only assumes spatial relationships between contiguous regions, or a k -nearest neighbor matrix, which assumes all k closest neighbors have the same impact on a region. Exponential distance decay allows neighboring regions to have exponentially more influence the closer they are located to the primary region. The distance decay matrix is defined as:

$$W_{ij} = \frac{\exp(-\frac{d_{ij}}{k})}{\sum_j^N \exp(-\frac{d_{ij}}{k})} \quad (6)$$

Where d_{ij} is the distance between the centroids of county i and county j . The matrix is row-standardized such that all spatial weights for each region sum to one. The weight matrix requires prior specification of the smoothing parameter k , which effectively determines the maximum size of detected clusters. Setting a fixed cluster size is clearly limiting. However, one cannot simultaneously test for the size of the clusters and the clusters themselves. This leads to a multiple testing problem where separate tests are conducted using different values of k , creating an inflated Type 1 error. We implement the method in Tango's MEET to correct for multiple testing. Tango [2] proposes an extended test statistic, the minimum of the profile p-value of C for k , with k varying from near zero to almost half the study area size. The proposed test statistic is:

$$P_{min} = \min PrC_k > c_k | H_0, k = \Pr C_k > c_k | H_0, k = k^* \quad (7)$$

Where c_k is the value of the test statistic as a function of k and k^* attains the minimum p-value of C_k , the global Modified Moran's I. In practice, Tango's method requires conducting a line search using discrete values of k and selecting the k that yields the lowest p-values of C_k . This effectively allows different sizes of clusters to be detected. By applying Tango's method to the Local Modified Moran's I, we allow for flexible neighborhood/cluster sizes for each place while dealing with multiple testing issues.

4. Simulation Study

We conduct a simulation study to compare the performance of this test against that of other local cluster tests: Local Moran's I [3] and the Spatial Scan Statistic [12]. Local Moran's I assesses the contribution of each

data point to Moran's I, a global index of overall clustering that is essentially a spatially weighted version of Pearson's correlation coefficient. High values indicate local hot spots or clusters most inconsistent with the null hypothesis of constant risk of disease across the study area. The Spatial Scan Statistic, or SatScan, constructs circles of varying radii around data points to construct a likelihood ratio test indicating the most likely cluster. Test statistics are correlated since circles overlap, but independence across Monte Carlo simulations allow for inference.

To assess test performance under various scenarios, we examine different underlying populations, spatial weight matrices, and local cluster sizes:

- Heterogeneous and homogeneous underlying populations
- Varying the smoothing parameter k in the row-standardized exponential distance decay matrix ($k = 2, 5, 10, 25$)
- Various local cluster sizes (5, 10, 15, 20 and 30% of the national population)

We simulate count data with a fixed total number of national cases (5000) for representing data from the 3109 counties of the continental United States. The spatial weight matrix will be based on distance measures calculated using the latitude and longitude coordinates of the county centroids. Waller et al. [15] find that test performance varies with the local population at risk. Thus, we simulate count data using both homogeneous county populations of 5000 and heterogeneous populations using 2010 Census figures.

In order to assess test performance, we simulate datasets with known clusters. We generate regional counts from a multinomial distribution. For the homogeneous county population case, the alternative hypothesis is specified:

$$H_A : (y_1, y_2, \dots, y_{3109}) \text{ multinomial}(y_+, r_1 \frac{n_1}{n'_+}, r_2 \frac{n_2}{n'_+}, \dots, r_n \frac{n_N}{n'_+}) \quad (8)$$

Where $n'_+ = \sum_i r_i n_i$ and r_i is the relative risk at geographic unit i . We simulate clusters of different sizes, from 5-30% of the total number of counties. Counties in each cluster are given an increased relative risk ratio of 4 relative to non-cluster counties, which have a relative risk of 1 by default. For each cluster size, we simulate 1000 alternative datasets to test for robustness (following the number of simulations in Jackson et al. [16]). Figure A.2 of

the appendix shows simulated clusters of 5-30% of the total number of U.S. counties.

Following Anselin [3], we use pseudo-significance levels for inference. The null hypothesis under this approach is conditional randomization: for each county, the simulated count for that county is fixed while remaining counts are randomly assigned across remaining counties without replacement. In other words, simulated counts are randomized, not cases. This procedure is repeated 1000 times to construct the empirical distribution of the Local Modified Moran's I for that county. Thus, there is a separate null distribution for each county and simulated dataset. We use the 95th percentile of the null distribution as the critical value for determining if a county is a statistically significant cluster.

Global clustering tests are often evaluated on statistical power - the probability the test will detect an overall clustering pattern if there is one. Local cluster tests should be evaluated on statistical power for each regional unit, since the tests produce a separate result for each unit. We assess test performance based on two diagnostics: precision and true positive versus false positive rate. Precision indicates the percent of each simulated cluster detected. The higher the precision, the more cluster counties detected, and hence the more "powerful" the test.

While a test may have high statistical power, it may also have a high rate of false positives. There is a tradeoff between the true positive rate and the false positive rate: tests that are sensitive to the detection of positives may also incorrectly label negatives as positives. This tradeoff can be plotted on an x-y axis. The y-axis represents the true positive rate such that:

$$\text{True positive rate} = \frac{\text{Number of true positives}}{\text{Number of total positives}} \quad (9)$$

The x-axis represents the false positive rate such that:

$$\text{False positive rate} = \frac{\text{Number of false positives}}{\text{Number of total negatives}} \quad (10)$$

The ideal point is at (0, 1), which represents a true positive rate of 100% and a false positive rate of 0%.

Figures A.5 - A.9 of the appendix are histograms of precision results from 1,000 simulated datasets for each cluster size. The red line indicates the mean precision. Mean precision ranges from 0.852 - 0.997. As cluster

size increases, mean precision decreases. The test is also more precise for heterogeneous populations than for homogeneous populations.

We plot true and false positive rates from the 1000 simulations conducted under each cluster scenario, using a critical value of 5% (Figures A.10 - A.14 of appendix). Each simulation is plotted as a point to illustrate the robustness of the test. Since Tango's MEET searches for the value of k that satisfies the minimum p-value profile of a test statistic, it is straightforward that this approach will yield many false positives. Thus, we plot results from each value of k used as well as the minimum p-value ("minp") approach. These plots reveal several interesting results. Within simulations, results are fairly robust. The notable exception is for the 5% cluster scenario, where the true positive rate ranges from zero to one for the case where $k = 2$ (largest neighborhood size). However, the bulk of simulations hover closer to the upper limit of 100% true positive rate. Variation between simulations also is smaller under heterogeneous populations than under homogeneous populations. Moving along values of k , smaller neighborhood sizes (larger k) increases the false positive rate with no gain in the true positive rate. Using the k of the minimum of the p-value profile yields the highest false positive rate, also with no gain in the true positive rate. Heterogeneous populations yield higher true positive rates, *ceteris paribus*, with varying effects on the false positive rate depending on the value of k used. Moving along cluster sizes, the true positive rate decreases as cluster size increases for homogeneous populations but not for heterogeneous populations. This result is consistent with the precision charts.

The false positive rate should be interpreted in the context of the LISA statistical framework. As noted by Anselin [3], LISAs indicate the degree of spatial similarity between local neighboring regions. High p-values indicate spatial similarity, while low p-values indicate spatial dissimilarity. Identified clusters could represent clusters of high, low or moderate values - as long as the magnitude of values are similar within the cluster. Thus, the false positive rate is a misnomer in this context, since "false" positives may indicate clusters of moderate or low values, not just high values within the forced clusters. Further, setting the critical value at 5% is standard for most statistical studies, but the number of false positives can become meaningfully large as the number of regions increases (e.g. 155 false positives for 3109 counties).

We compare the results of the Local Modified Moran's I with that from Local Moran's I and SaTScan. Naturally, the Local Modified Moran's I is

the most comparable to Local Moran’s I. High values of both LISAs indicate spatial autocorrelation, while low values indicate counties that are outliers from their neighbors. We search for clusters of high Local Moran’s I values to identify spatial clusters. SatScan searches for clusters of similar values, but only for similar high rates or similar low rates. Since the simulated data force a cluster of elevated risk ratios, we specify SaTScan to search for high rate clusters. We specify a maximum cluster size of 50% of the population (the default in SaTScan) and use circular search windows (instead of elliptical ones).

Table 1 compares precision results for Local Modified Moran’s I (LMMI), Local Moran’s I (LMI) and SaTScan (SaT). The two LISAs perform very similarly, although the Local Modified Moran’s I tends to have slightly higher precision as cluster size increases. This is true for both homogeneous and heterogeneous populations. SaTScan outperforms both tests in terms of precision: precision is high across all cluster sizes, with a mean of 100% for heterogeneous populations and 99.9% for homogeneous populations.

Table 1: Precision Results

Cluster Size	Population	LMMI	LMI	SaT
5%	Homogeneous	97.5%	97.5%	99.9%
	Heterogeneous	99.7%	99.7%	100%
10%	Homogeneous	95.7%	95.7%	99.9%
	Heterogeneous	99.5%	99.4%	100%
15%	Homogeneous	93.6%	93.5%	99.9%
	Heterogeneous	99.4%	99.2%	100%
20%	Homogeneous	91%	90.4%	99.9%
	Heterogeneous	99%	98.8%	100%
30%	Homogeneous	85.2%	83.1%	100%
	Heterogeneous	98%	97.3%	100%

Table A.2 of the Appendix compares true positive and false positive rates for the three tests. The true positive rates are reflected in the precision results. The false positive rates are higher for the two LISAs than for SaTScan. As aforementioned, the interpretation of high LISA values can mean clusters of similar high values, moderate values or low values. The two LISAs may detect moderate value clusters or low value clusters and thus inflate the false positive rate. The LMMI tends to have a higher false positive rate than the

LMI, but this trend reverses as cluster size increases (for small values of k). By excluding clusters of moderate values, SaTScan minimizes its false positive rate. The Appendix shows precision results and true and false positive rates across simulations for each test.

We also illustrate the chance of each county being detected as inside a cluster. These maps are located in the Appendix as Figures A.15 - A.19 for the Local Modified Moran's I, A.30 - A.34 for the Local Moran's I, and A.45 - A.49 for SaTScan. Each map indicates the percent of times a county is selected as part of a cluster over 1000 simulations. For the LMMI and LMI, we map results using $k = 2$ in the spatial weight matrix. SaTScan, with a false positive rate close to zero, has the best chances of detecting true cluster counties without including non-cluster counties. It performs well for all cluster sizes. The LMMI has a high chance of detecting true cluster counties regardless of cluster size, but it also detects non-cluster counties in the western half of the United States. The LMI also has a high chance of detecting true cluster counties, but it detects non-cluster counties in the eastern half of the United States. While the LMI has a very low chance of detecting non-cluster counties around the true cluster (regardless of cluster size), the LMMI has a moderate chance of detecting these (the chance also decreases as cluster size increases).

5. Application

We apply the Local Modified Moran's I to real data, using the New York leukemia data introduced by Turnbull et al. [9] and used extensively in Waller and Gotway [11]. This dataset consists of the number of leukemia cases and underlying population at risk for eight counties (Broome, Cayuga, Chenango, Cortland, Madison, Onondaga, Tioga, and Tompkins) in upstate New York for the years 1978-1982. Leukemia cases are presented on the census tract level and includes a total of 592 cases among 1,057,673 people at risk. Figure A.3 of the Appendix presents the raw rates.

We calculate both Local Modified Moran's I (LMMI) and Local Moran's I (LMI) for the leukemia data to compare the results of the two tests. We use row-standardized adjacency weights, the same spatial weight matrix used by Waller and Gotway [11]. The Figure A.4 of the Appendix map the quintiles of the test statistics. The statistics produce somewhat similar results. The highest quintile of the test statistics (e.g. spatial similarity in leukemia rates) are for census tracts north of Binghamton, west of Ithaca, east of Cortland,

and in pockets along the northern edge of the study area. The Local Modified Moran's I detected slightly higher spatial similarity in tracts in the southwest and northwest sections of the study area. These results are consistent with findings in Waller and Gotway [11].

6. Conclusion

The selection of a local cluster test should depend on the researcher's objective. For identifying clusters of high or low values, SaTScan performs best among existing tests. For identifying clusters of similar values - high, moderate or low - a LISA is more appropriate. LISAs also allow for a user-defined spatial weight matrix, which allows theory to motivate empirics. These simulation studies show that the Local Modified Moran's I has at least the precision of the Local Moran's I for small clusters, and it has higher precision for larger clusters. The Local Modified Moran's I also implements several statistical improvements that affect performance and improve flexibility. Monte Carlo methods allow for an empirical distribution of the test statistic in the absence of a clear null distribution. Implementation of Tango's (2000) *minp* allows for variable neighborhood sizes in the spatial weight matrix while dealing with multiple testing problems.

A drawback of the Local Modified Moran's I is a high false positive rate when neighborhood sizes are small (k is large). However, it has a smaller false positive rate than the Local Moran's I for larger clusters (for $k = 2, 5$). Nevertheless, when the research objective is to identify clusters of similar values (of any value), Local Modified Moran's I offers improvements over Local Moran's I.

Appendix A. Appendix

Appendix A.1. Summary

Figure A.2 describes the cluster sizes simulated for the Local Modified Moran's I and Local Moran's I. Larger clusters encompass the smaller clusters within them.

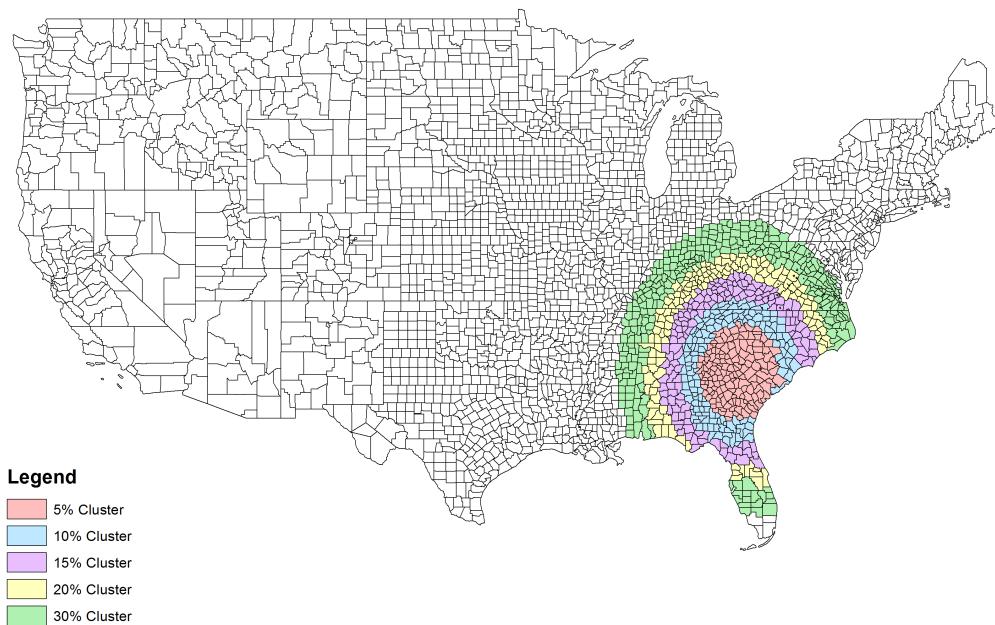


Figure A.2: Local Cluster Simulation for Local Modified Moran's, Various Sizes

Table A.2: True and False Positive Rates

Size	Pop	Rate	LMMI ($k = $)						LMI ($k = $)				SaTScan
			2	5	10	25	minp	2	5	10	25	minp	
5%	Homog	TPR	85.8%	97.5%	97.5%	97.5%	97.5%	97.5%	97.5%	97.5%	97.5%	97.5%	99.9%
		FPR	26.3%	47.5%	45.7%	23%	73.7%	10.4%	10.4%	10.4%	6.7%	11.1%	0%
	Heter	TPR	98.4%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	100%
		FPR	24%	47.2%	53.9%	47%	81.8%	8.4%	8.4%	8.4%	5.3%	8.9%	0%
10%	Homog	TPR	95.7%	95.7%	95.7%	95.7%	95.7%	95.7%	95.7%	95.7%	95.7%	95.7%	99.9%
		FPR	30.9%	49%	54.1%	44.1%	81.7%	18.8%	18.8%	18.7%	12.7%	20%	0%
	Heter	TPR	99.5%	99.5%	99.5%	99.5%	99.5%	99.4%	99.4%	99.4%	99.4%	99.4%	100%
		FPR	22.4%	47.7%	62.5%	62.8%	86.5%	17%	17%	16.9%	10.5%	18.2%	0%
15%	Homog	TPR	93.6%	93.6%	93.6%	93.6%	93.6%	93.5%	93.5%	93.5%	92.9%	93.5%	99.9%
		FPR	30.3%	50.1%	57.8%	54.8%	84.1%	24.9%	24.9%	24.8%	19.5%	26.4%	0%
	Heter	TPR	99.4%	99.4%	99.4%	99.4%	99.4%	99.2%	99.2%	99.2%	99.2%	99.2%	100%
		FPR	20.1%	49%	65.7%	71.2%	87.3%	29.3%	29.3%	29.2%	19.1%	31.2%	0%
20%	Homog	TPR	91%	91%	91%	91%	91%	90.4%	90.4%	90.4%	89.3%	90.4%	99.9%
		FPR	29.1%	51.2%	60.4%	60.2%	84.3%	30.7%	30.7%	30.5%	25.1%	32.2%	0%
	Heter	TPR	99%	99%	99%	99%	99%	98.8%	98.8%	98.8%	98.7%	98.8%	100%
		FPR	19.1%	50.8%	68.2%	75.5%	86.8%	43.4%	43.4%	43.4%	30.9%	45.8%	0%
30%	Homog	TPR	85.2%	85.2%	85.2%	85%	85.2%	83.1%	83.1%	83%	80.7%	83.1%	100%
		FPR	28%	56.3%	65.6%	68.6%	85.4%	47.5%	47.5%	47%	34.6%	48.6%	0%
	Heter	TPR	98%	98%	98%	97.9%	98%	97.3%	97.3%	97.3%	96.8%	97.3%	100%
		FPR	20.5%	57.9%	73.5%	81.7%	88.1%	66.2%	66.2%	66.1%	57.3%	68.3%	0%

Figure A.3: Raw rates of leukemia by census tract for eight counties in upstate New York, 1978-1982

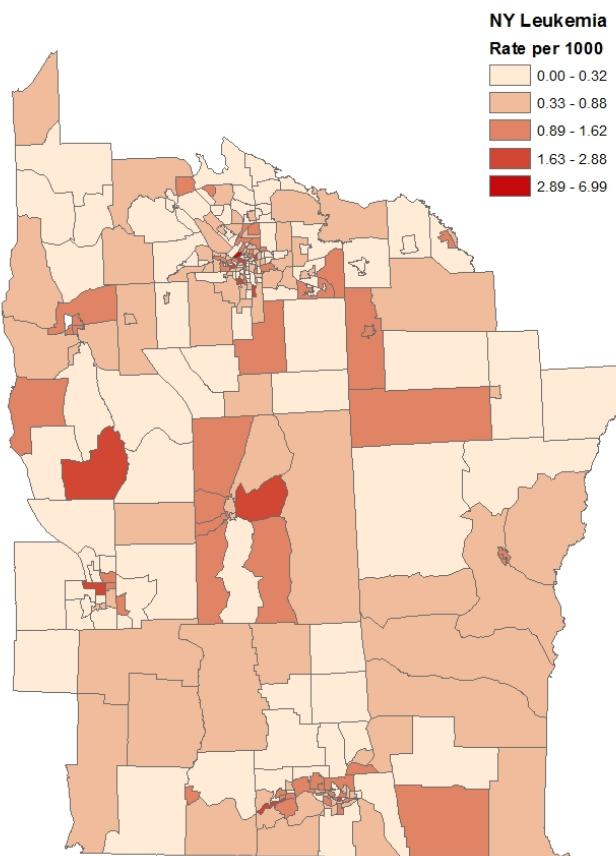
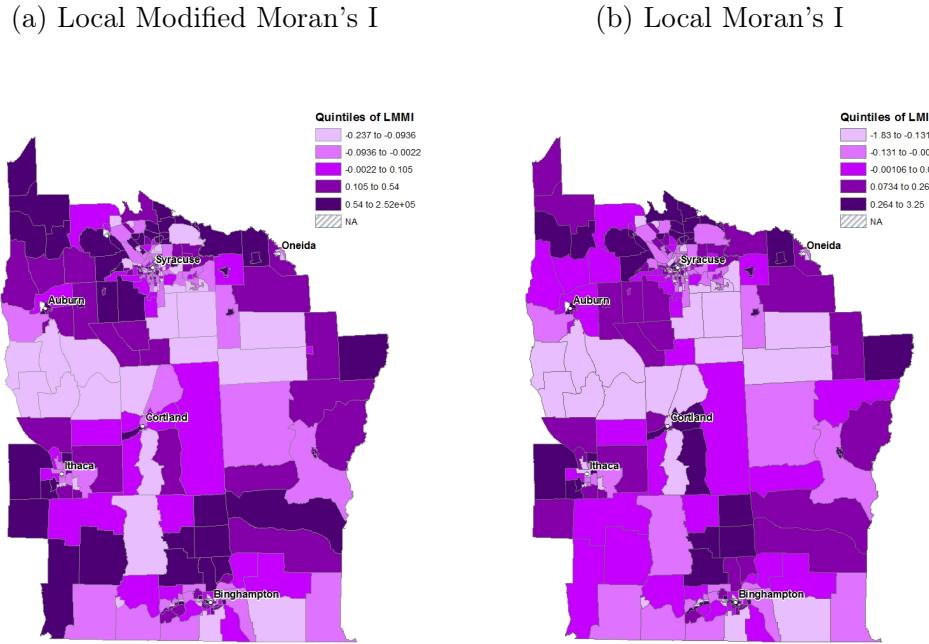


Figure A.4: Quintiles of Test Statistic, NY Leukemia



Appendix A.2. Detailed Results for Local Modified Moran's I

The plots in this section show the results for the Local Modified Moran's I. Figures A.5 - A.9 describes the precision of the Local Modified Moran's I for simulated data. Data are simulated by cluster size and homogeneous versus heterogeneous populations. The red line indicates the mean precision for each scenario.

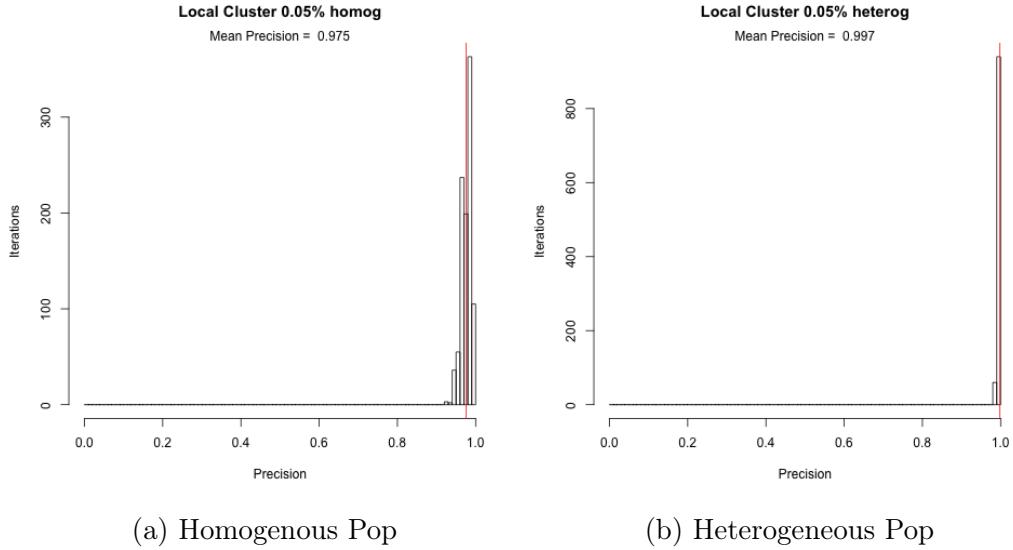


Figure A.5: Precision Results: 5% Cluster Size, Local Modified Moran's I

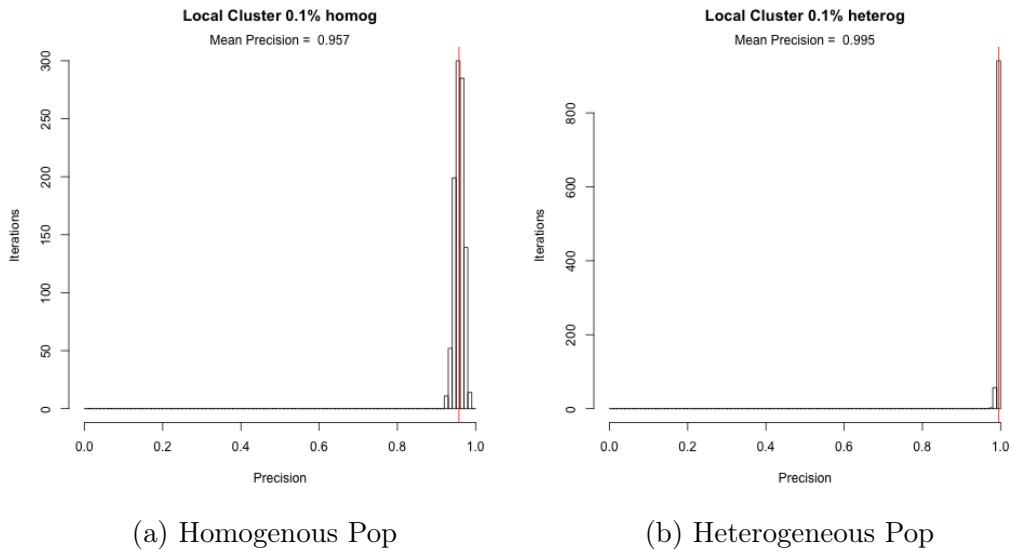


Figure A.6: Precision Results: 10% Cluster Size, Local Modified Moran's I

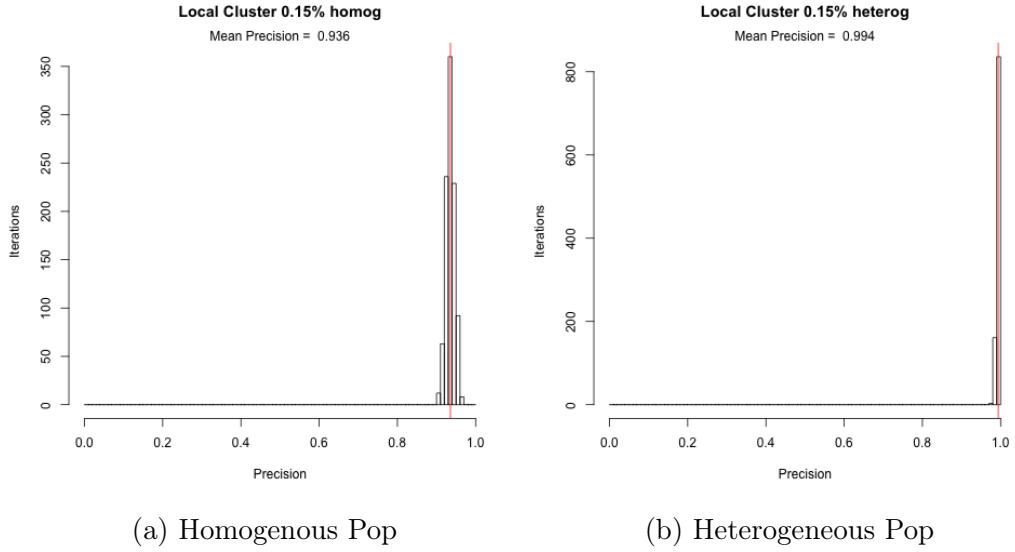


Figure A.7: Precision Results: 15% Cluster Size, Local Modified Moran's I

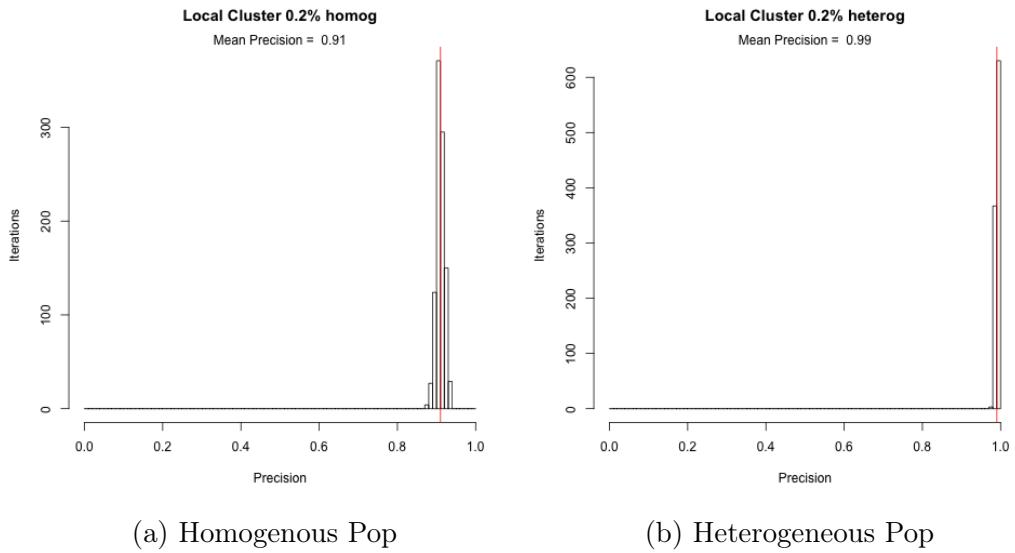


Figure A.8: Precision Results: 20% Cluster Size, Local Modified Moran's I

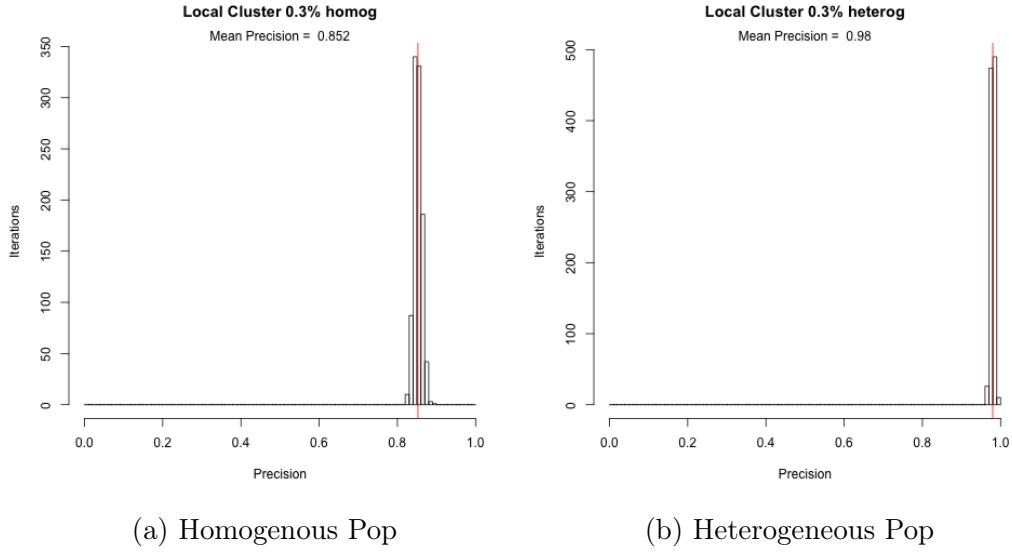


Figure A.9: Precision Results: 30% Cluster Size, Local Modified Moran's I

Figures A.10 - A.14 plot true positives versus false positives under the Local Modified Moran's I for each cluster scenario.

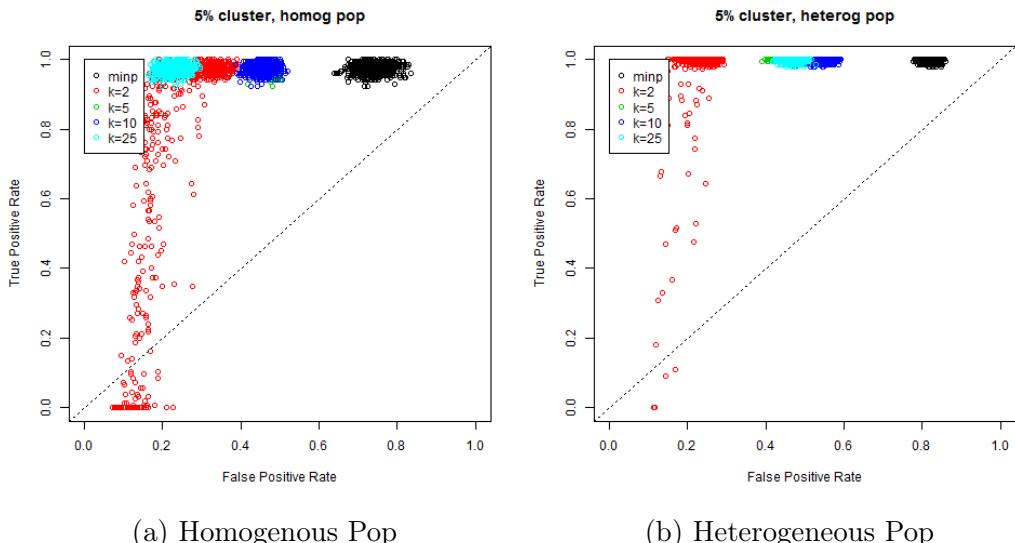


Figure A.10: ROC Curves: 5% Cluster Size, Local Modified Moran's I

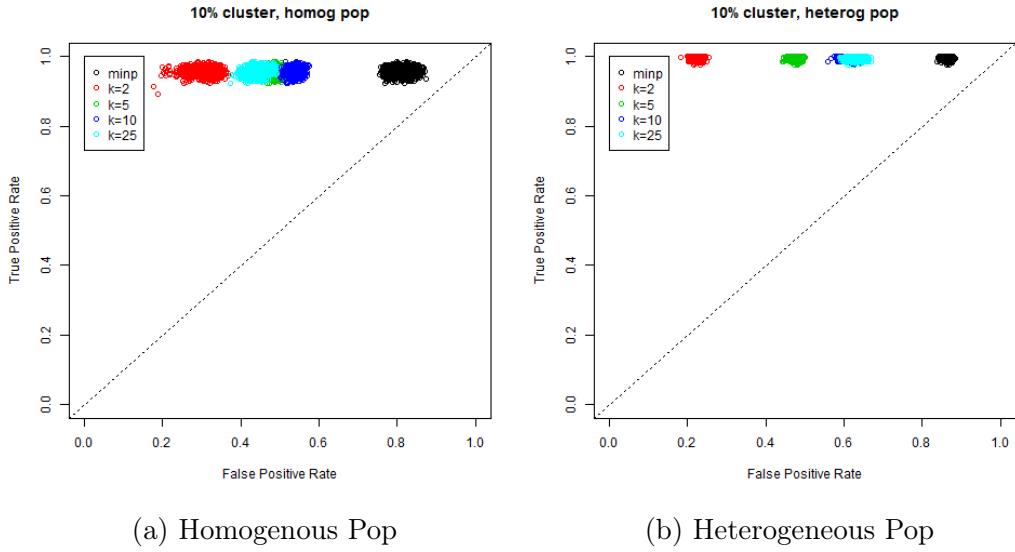


Figure A.11: ROC Curves: 10% Cluster Size, Local Modified Moran's I

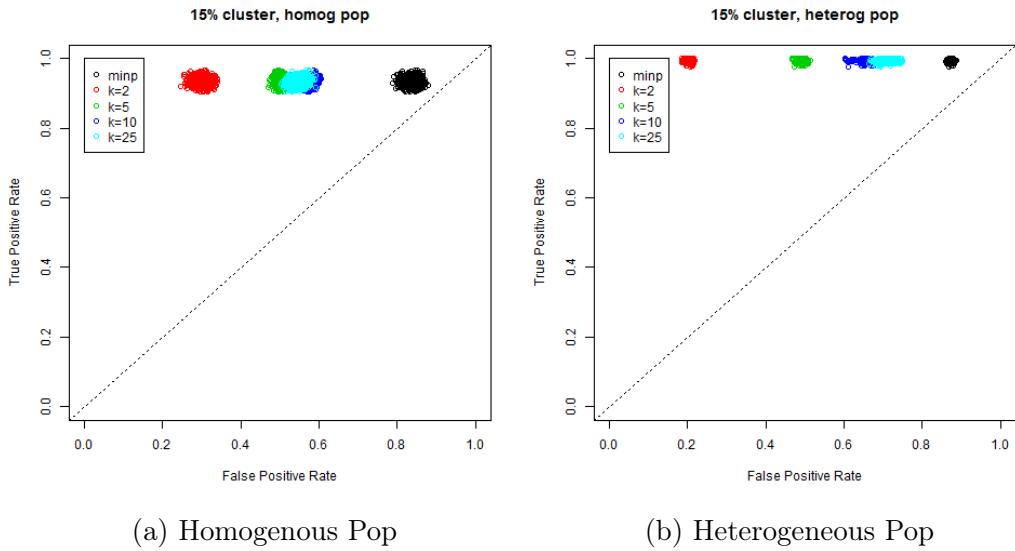


Figure A.12: ROC Curves: 15% Cluster Size, Local Modified Moran's I

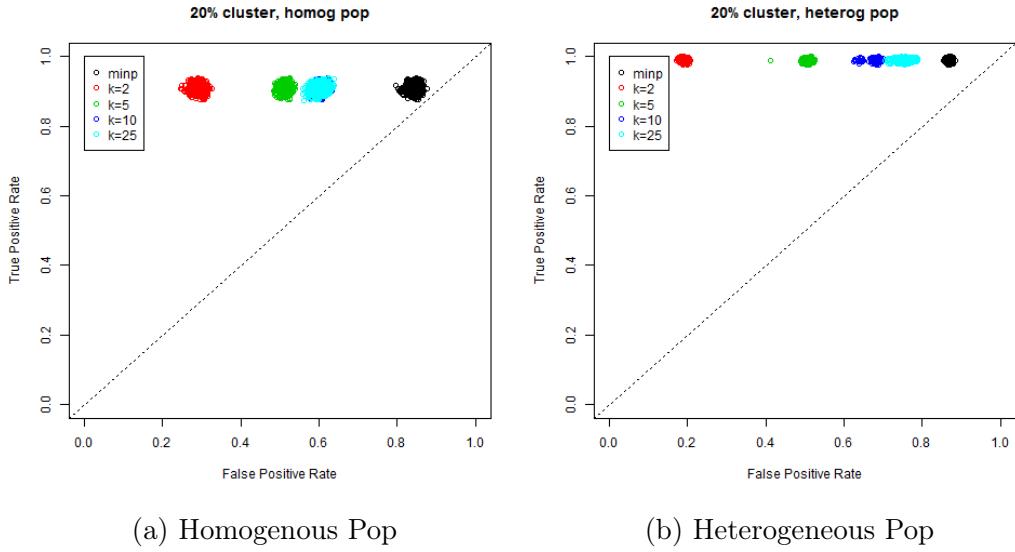


Figure A.13: ROC Curves: 20% Cluster Size, Local Modified Moran's I

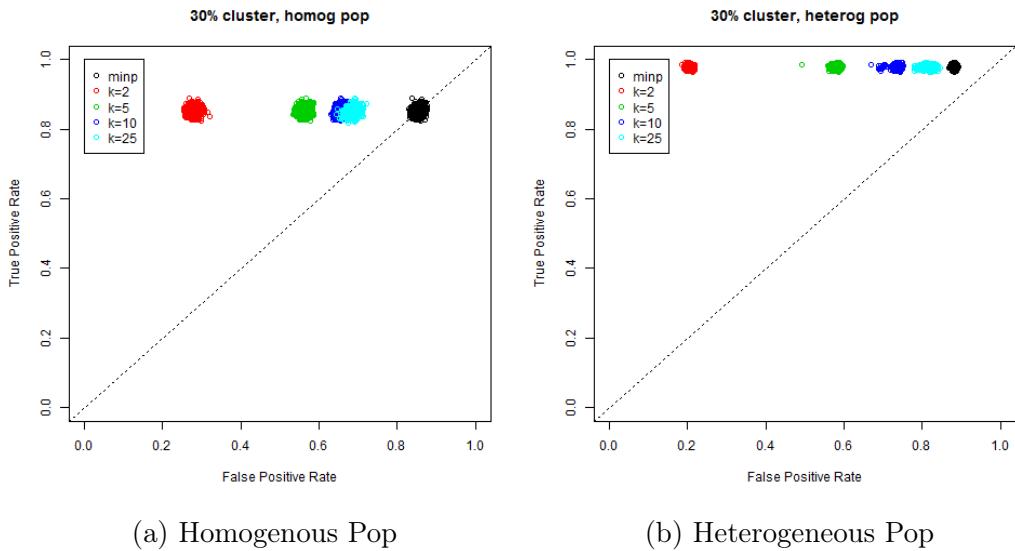


Figure A.14: ROC Curves: 30% Cluster Size, Local Modified Moran's I

Figures A.15 - A.19 shows the chance of each county being detected as clusters (the percent of times a county was selected as part of a cluster over 1000 simulations).

Figure A.15: 5% Cluster Size, Heterogeneous Population

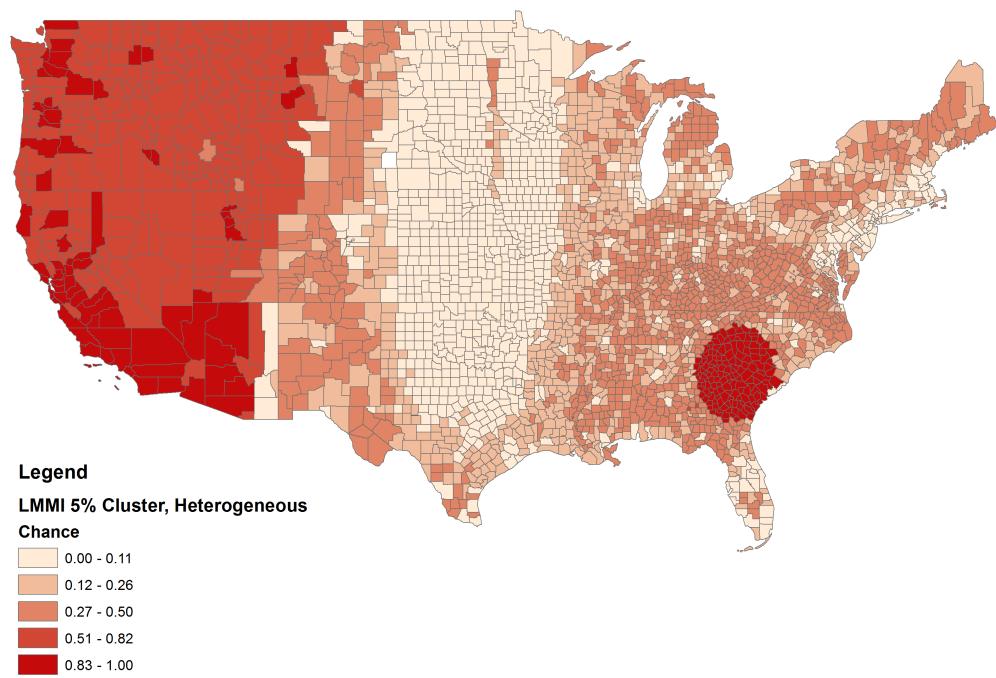


Figure A.16: 10% Cluster Size, Heterogeneous Population

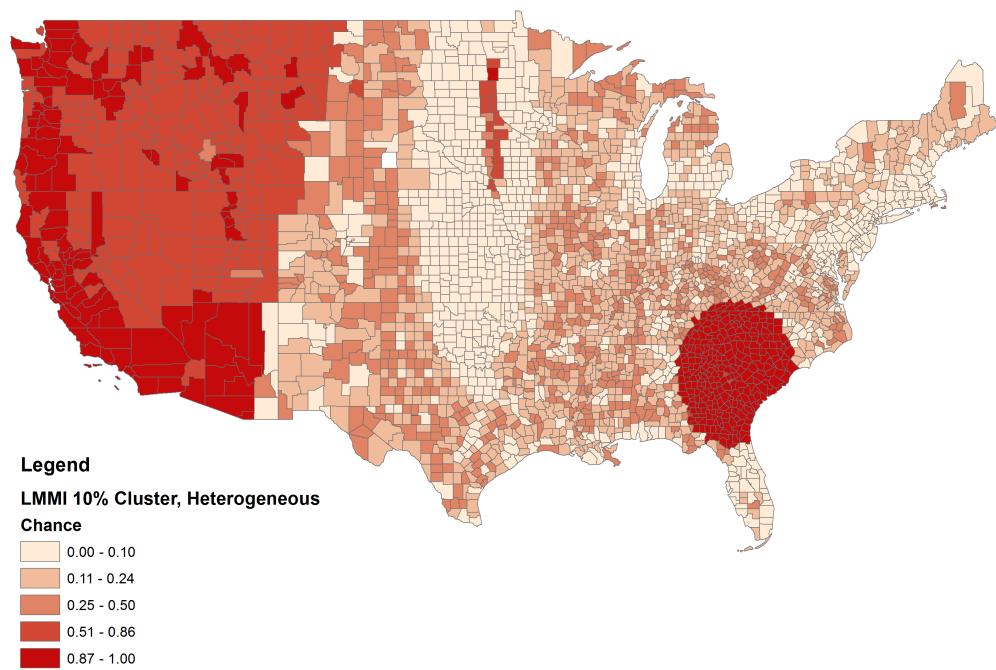


Figure A.17: 15% Cluster Size, Heterogeneous Population

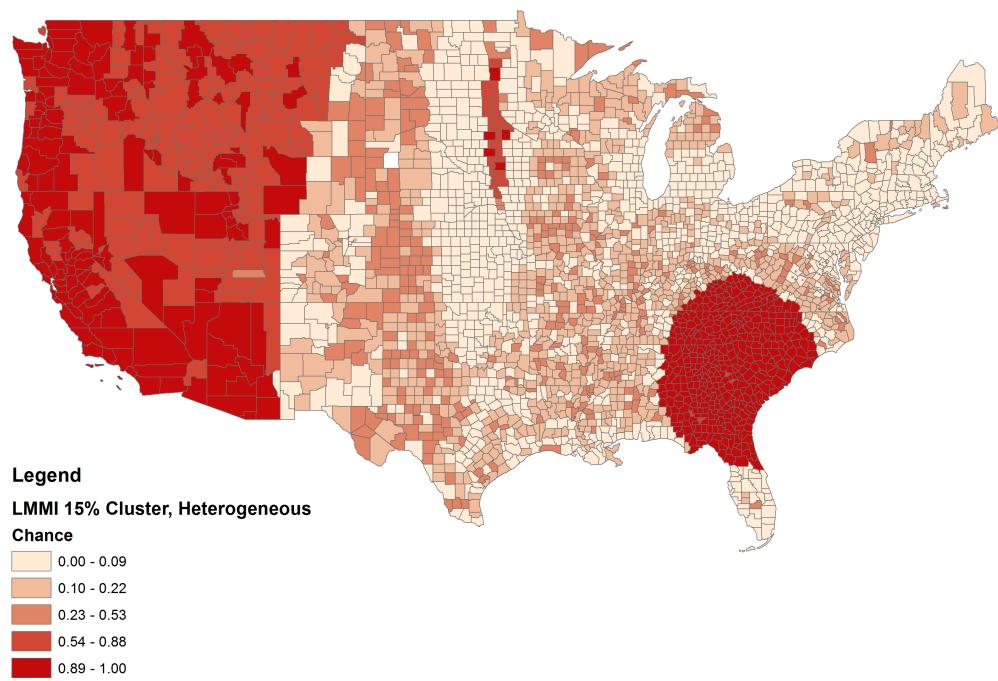


Figure A.18: 20% Cluster Size, Heterogeneous Population

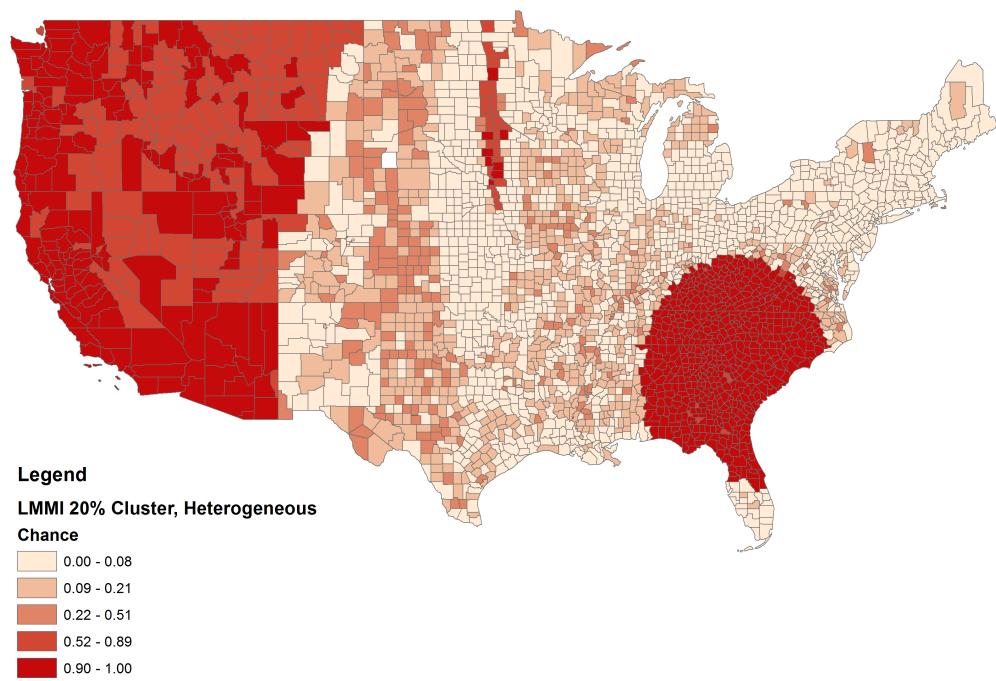
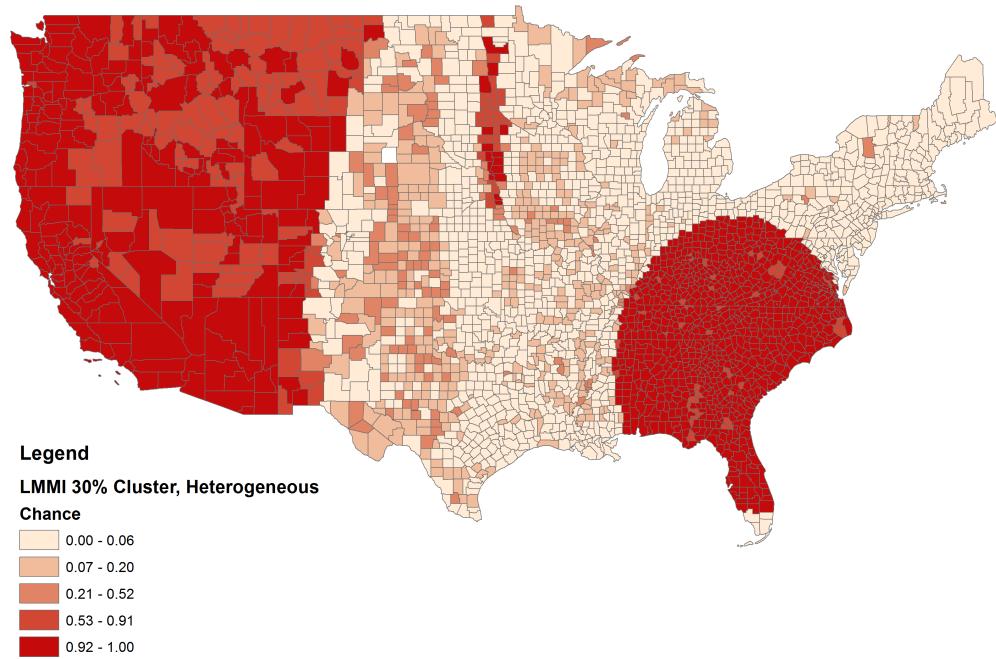


Figure A.19: 30% Cluster Size, Heterogeneous Population



Appendix A.3. Detailed Results for Local Moran's I

The plots in this section show the results for Local Moran's I, a comparison test. Figures A.20 - A.24 plot the precision of the Local Moran's I under various cluster scenarios.

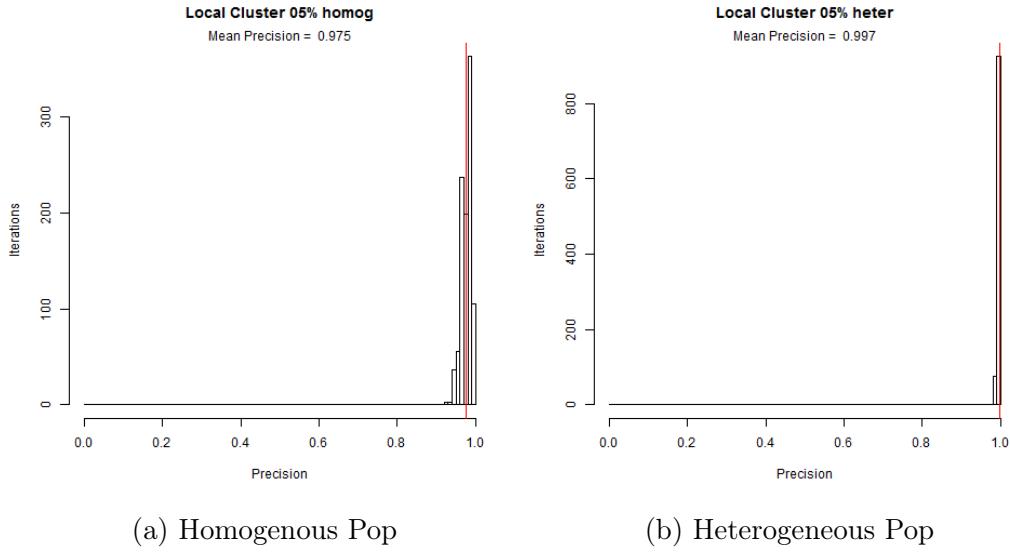


Figure A.20: Precision Results: 5% Cluster Size, Local Moran's I

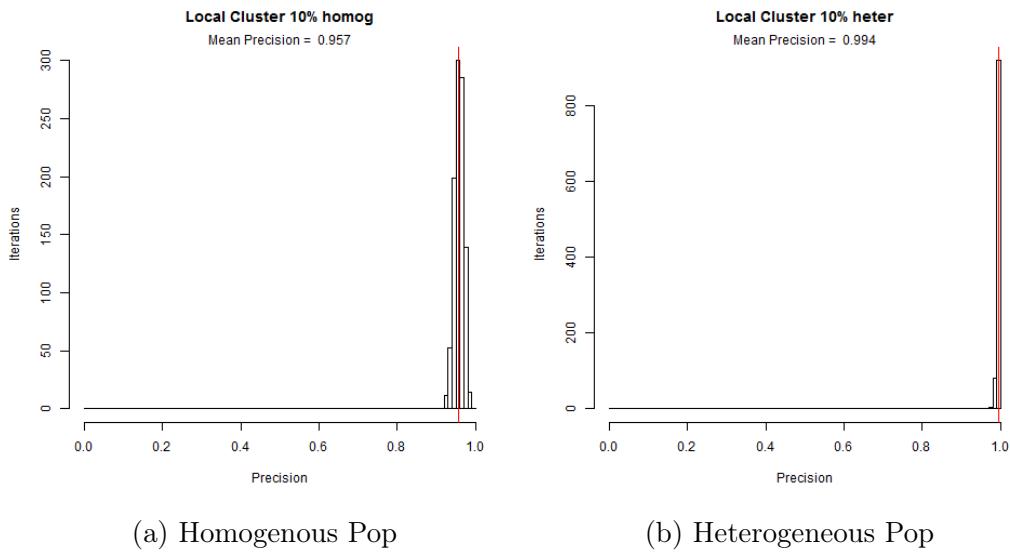


Figure A.21: Precision Results: 10% Cluster Size, Local Moran's I

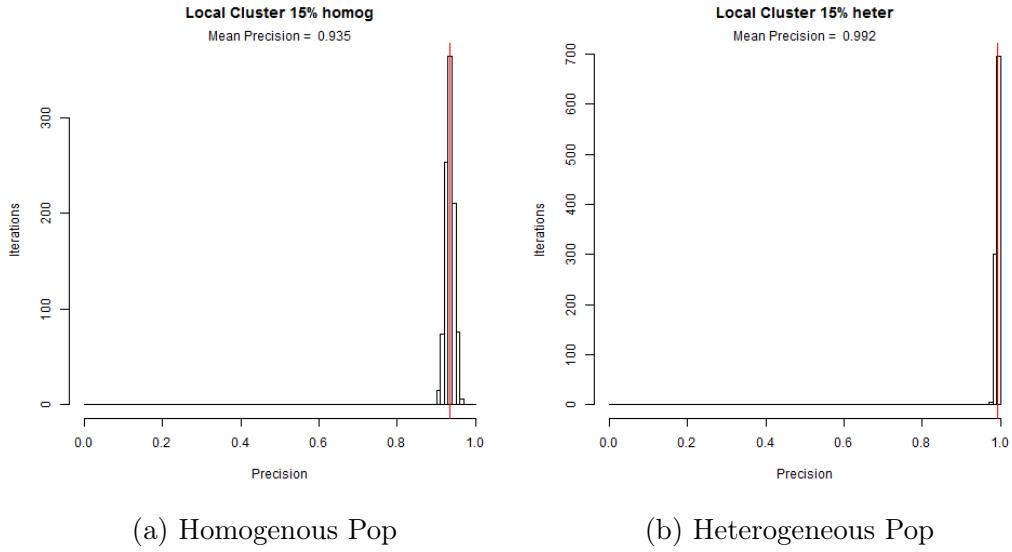


Figure A.22: Precision Results: 15% Cluster Size, Local Moran's I

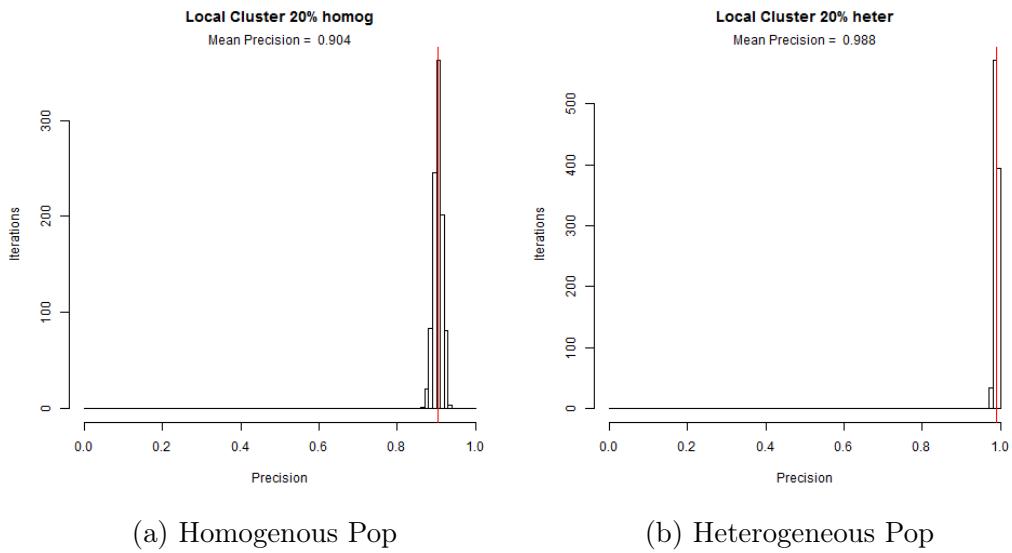


Figure A.23: Precision Results: 20% Cluster Size, Local Moran's I

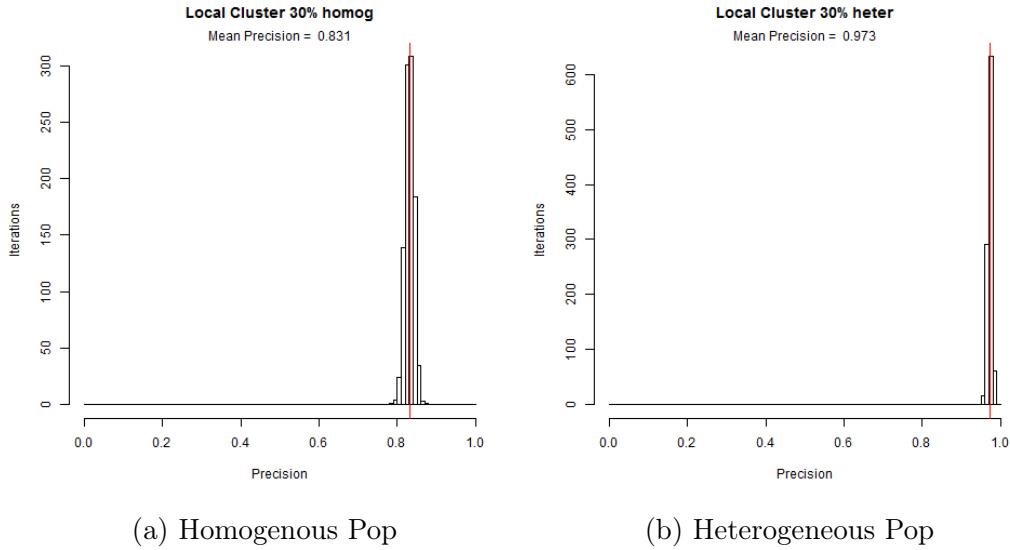


Figure A.24: Precision Results: 30% Cluster Size, Local Moran's I

Figures A.25 - A.29 plot true positives versus false positives for the Local Moran's I under each cluster scenario.

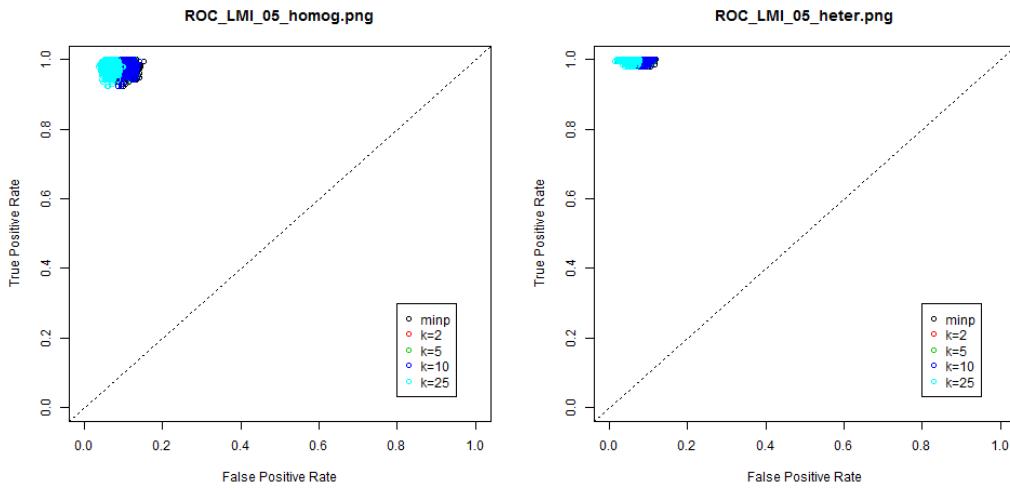


Figure A.25: ROC Curves: 5% Cluster Size, Local Moran's I

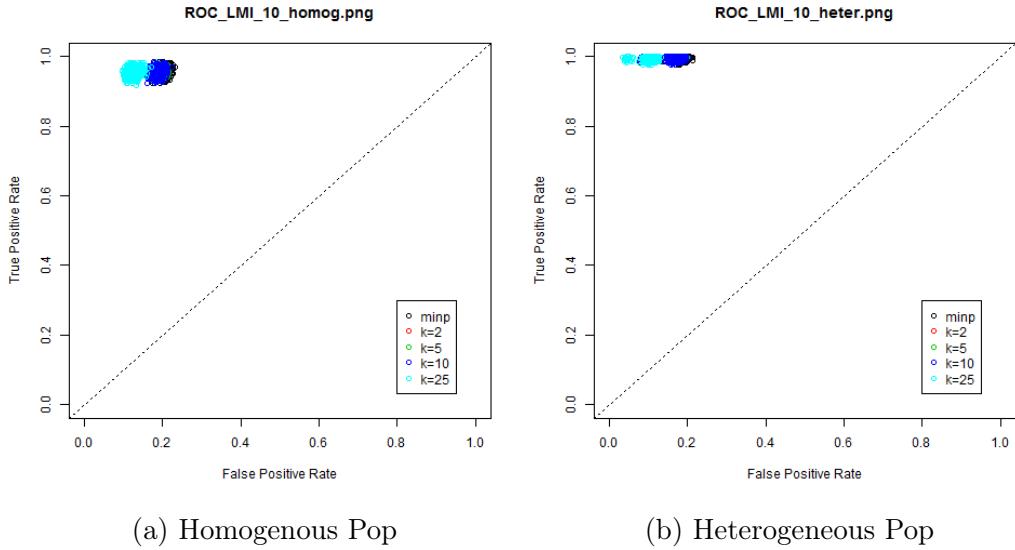


Figure A.26: ROC Curves: 10% Cluster Size, Local Moran's I

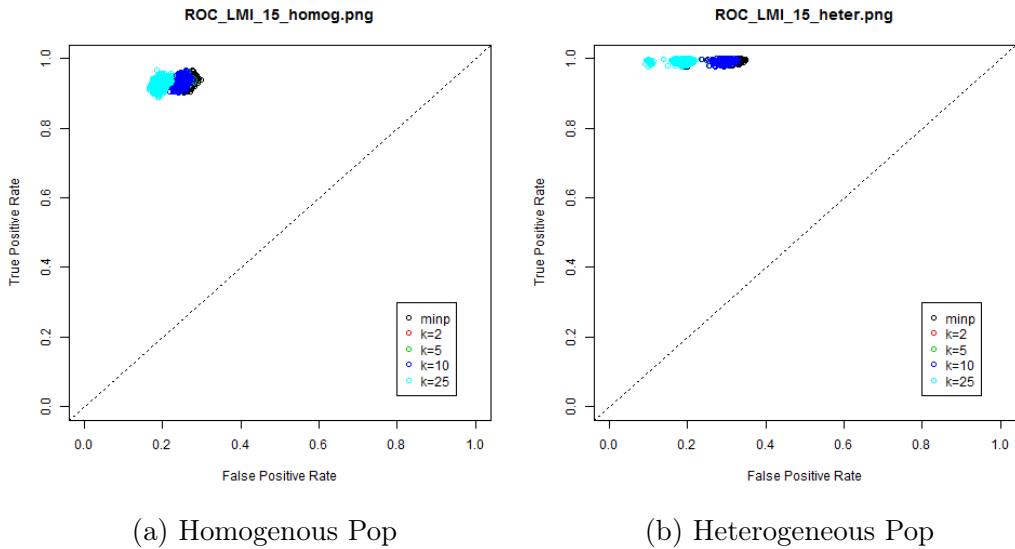


Figure A.27: ROC Curves: 15% Cluster Size, Local Moran's I

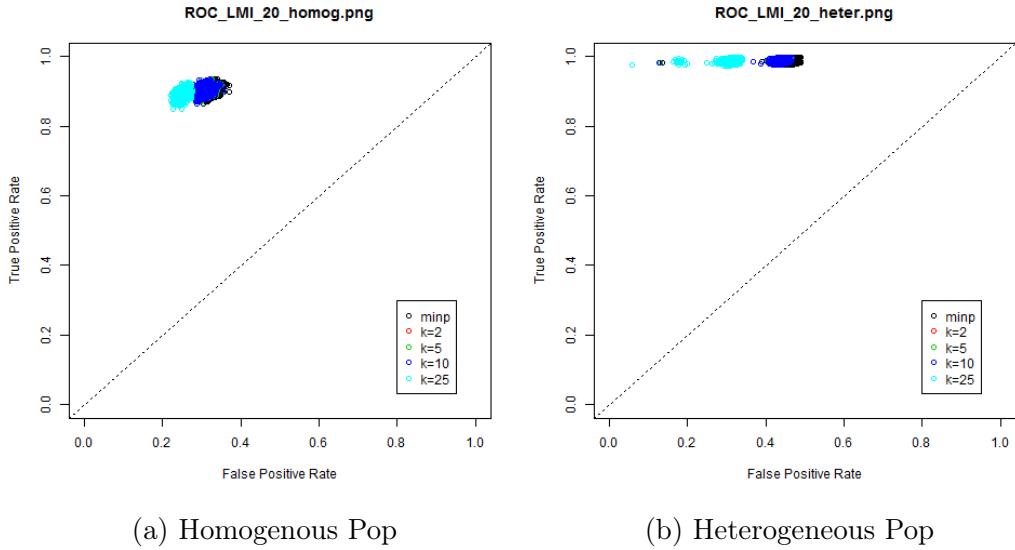


Figure A.28: ROC Curves: 20% Cluster Size, Local Moran's I

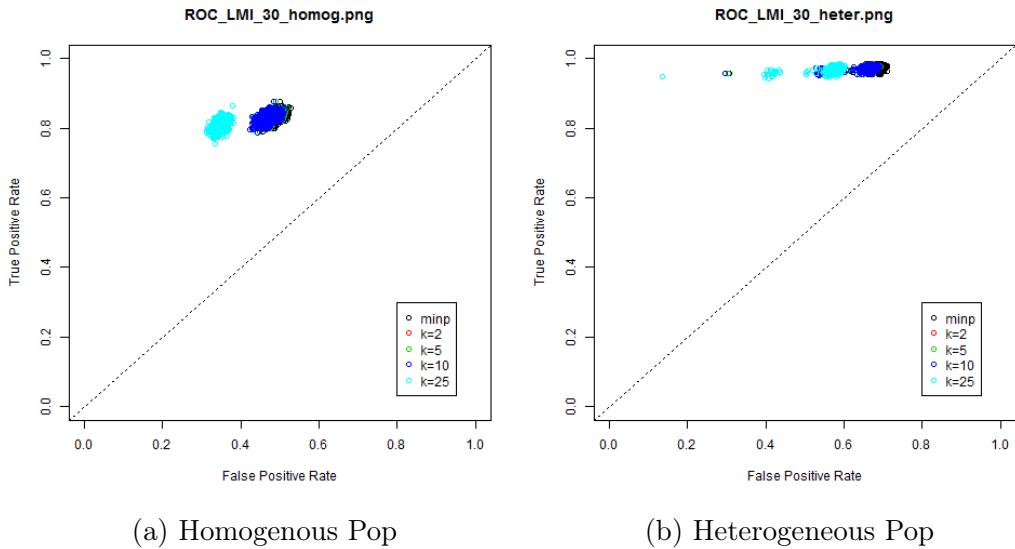


Figure A.29: ROC Curves: 30% Cluster Size, Local Moran's I

Figures A.30 - A.34 shows the chance of each county being detected as clusters.

Figure A.30: 5% Cluster Size, Heterogeneous Population

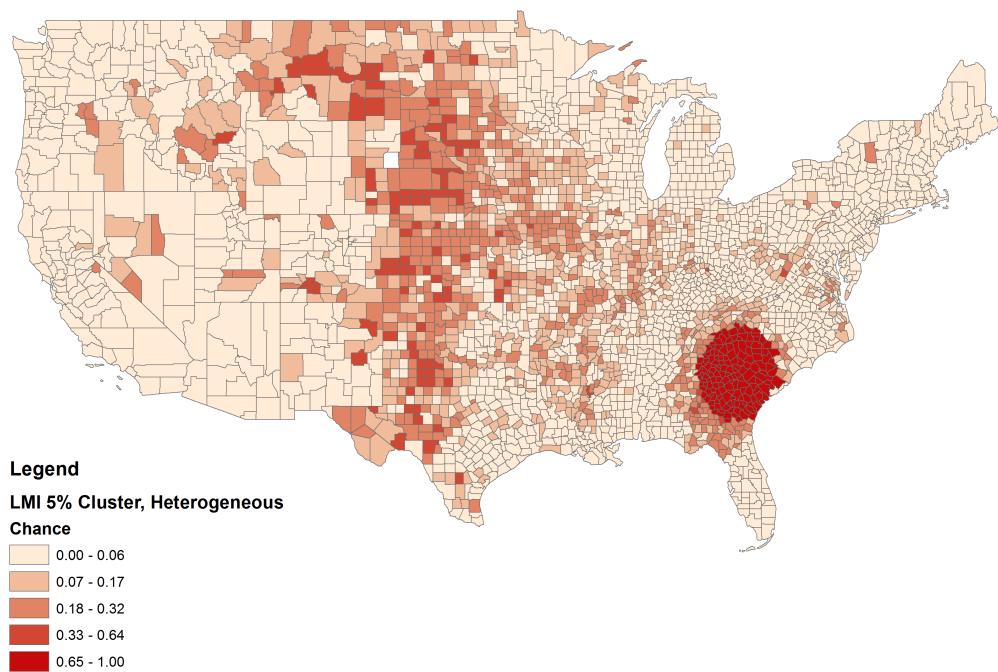


Figure A.31: 10% Cluster Size, Heterogeneous Population

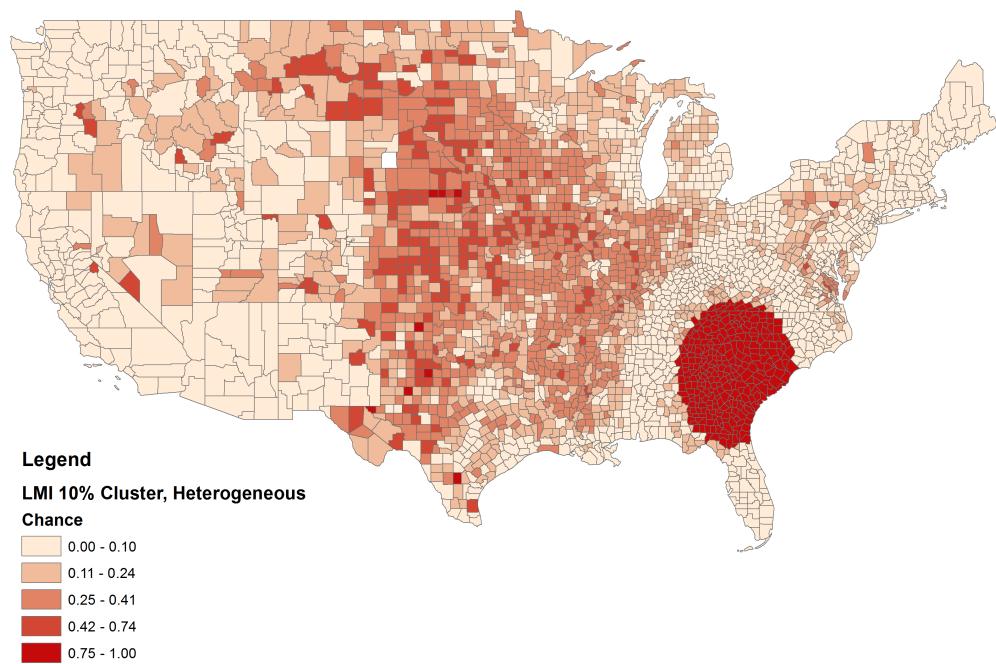


Figure A.32: 15% Cluster Size, Heterogeneous Population

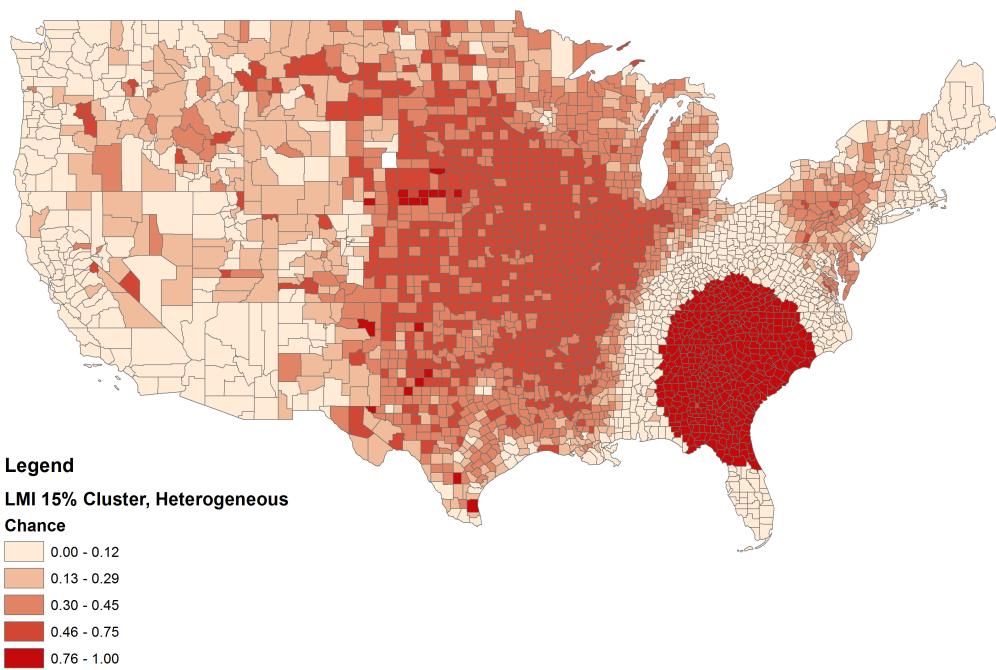


Figure A.33: 20% Cluster Size, Heterogeneous Population

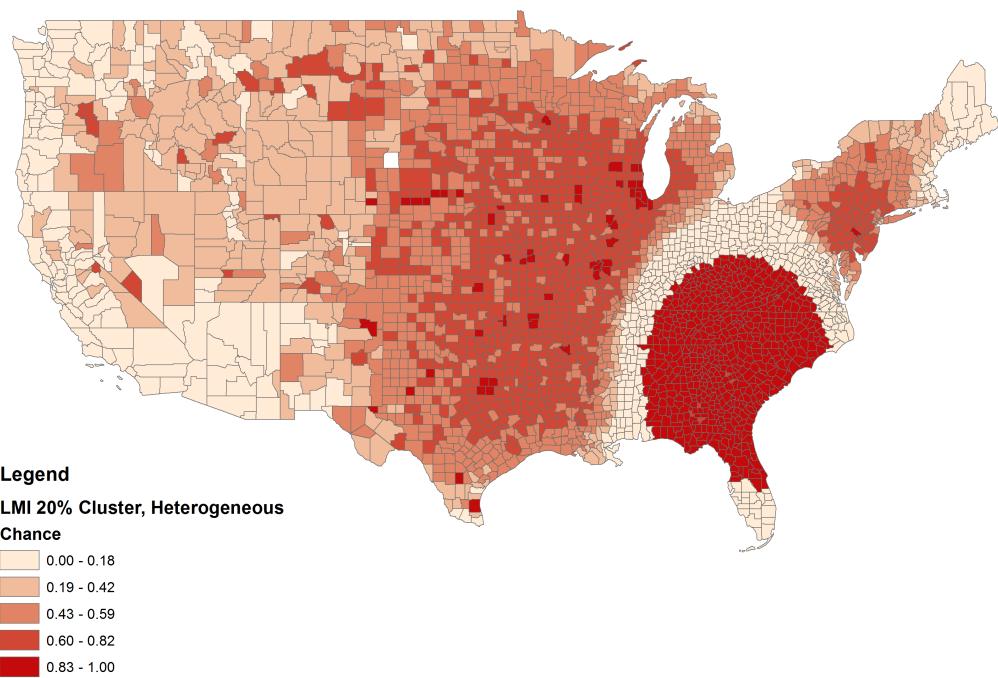
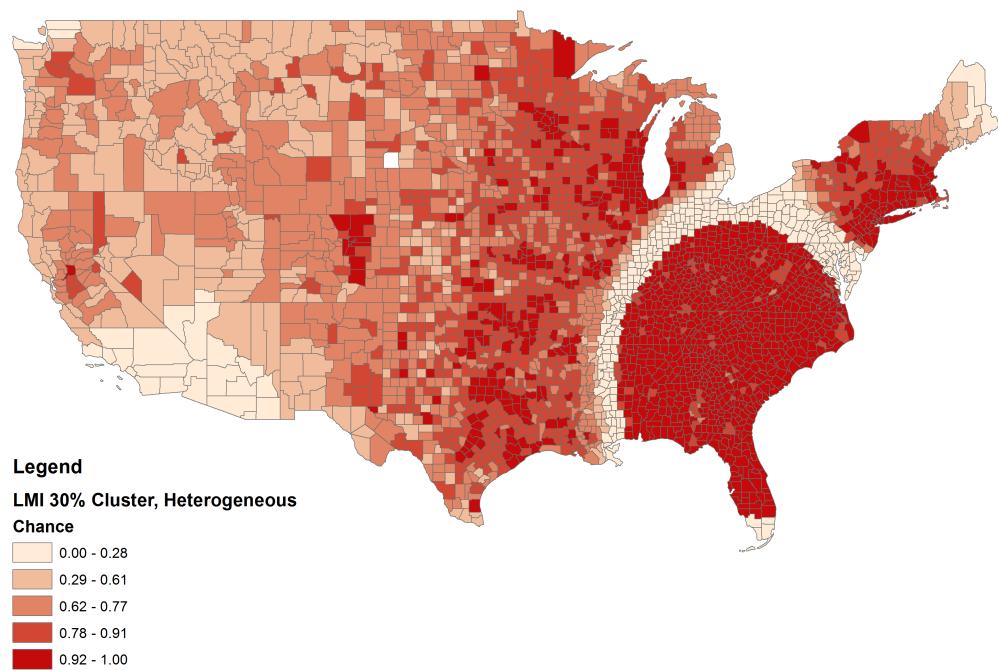


Figure A.34: 30% Cluster Size, Heterogeneous Population



Appendix A.4. Detailed Results for SaTScan

The plots in this section show the results for SaTScan, a comparison test. Figures A.35 - A.39 plot the precision of SaTScan under various cluster scenarios.

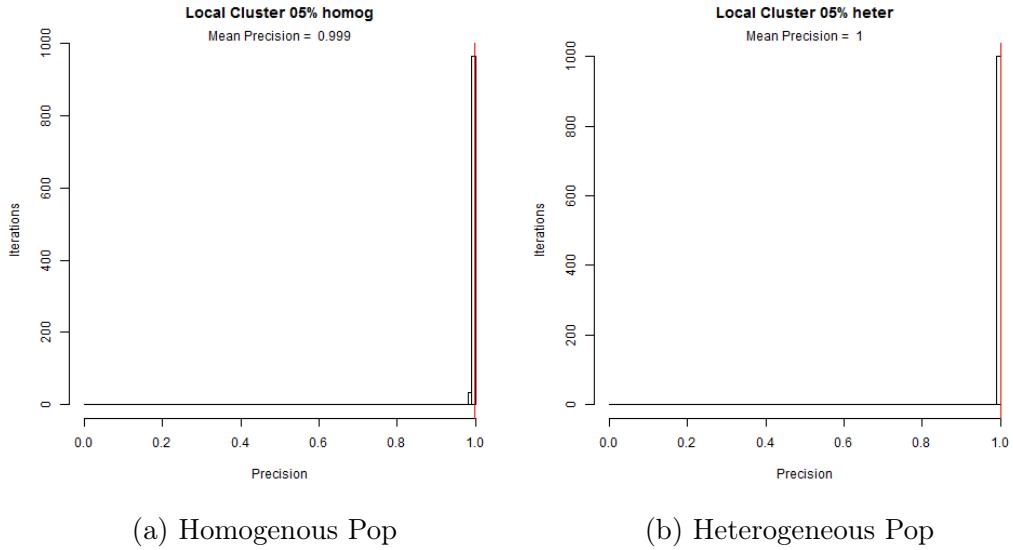


Figure A.35: Precision Results: 5% Cluster Size, SaTScan

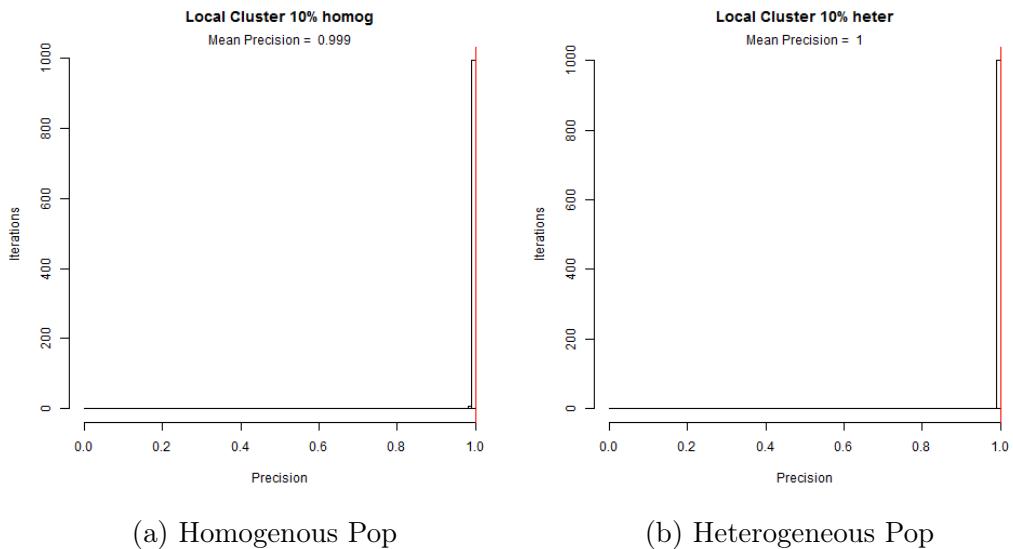


Figure A.36: Precision Results: 10% Cluster Size, SaTScan

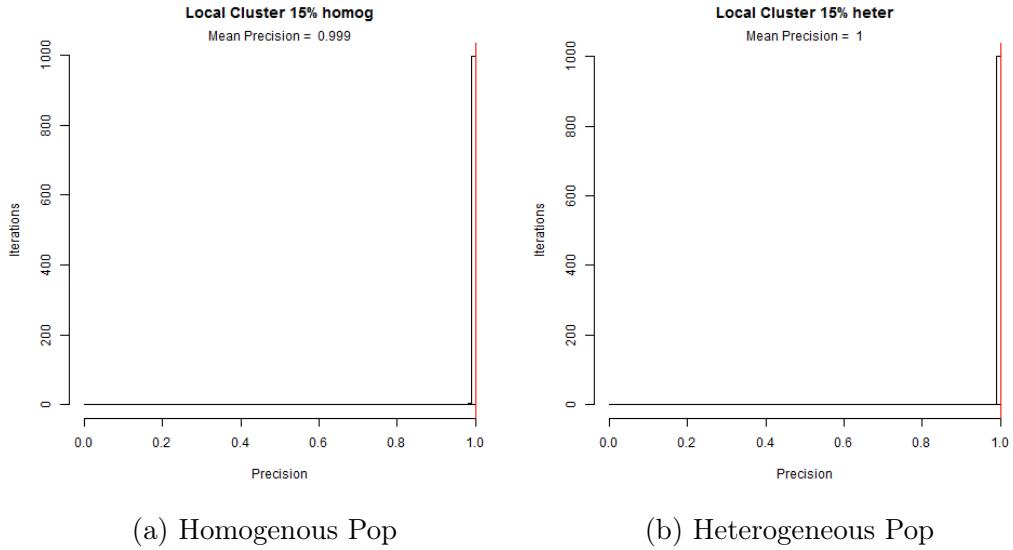


Figure A.37: Precision Results: 15% Cluster Size, SaTScan

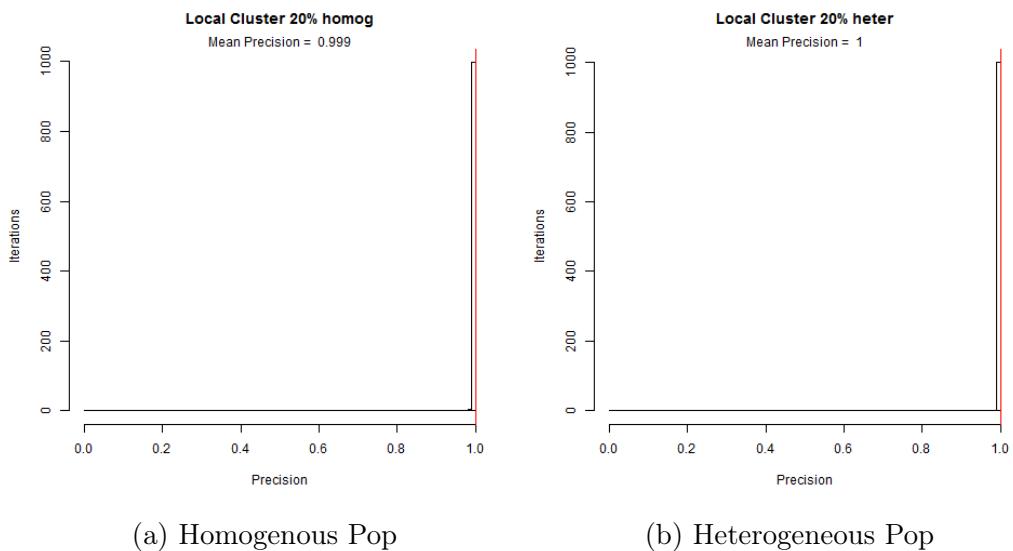


Figure A.38: Precision Results: 20% Cluster Size, SaTScan

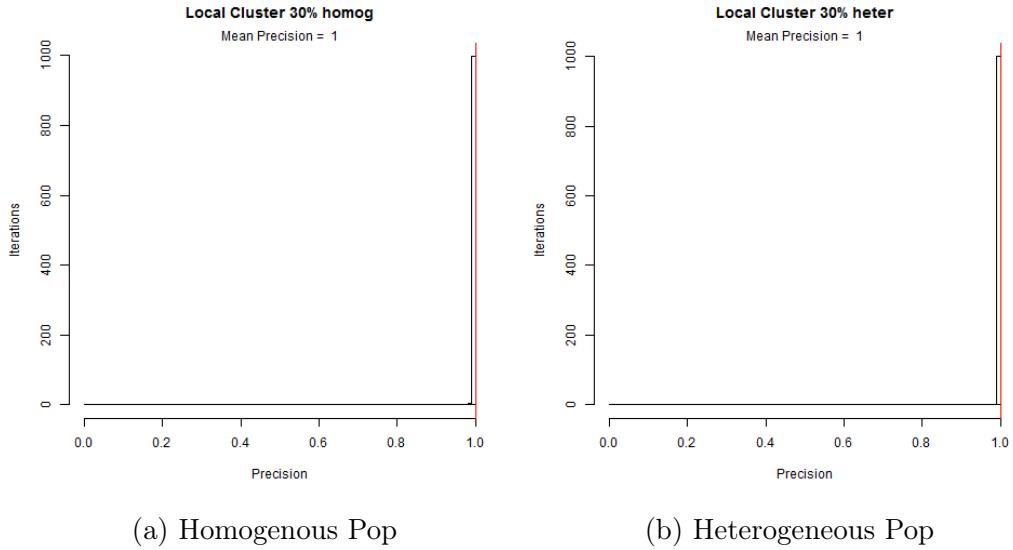


Figure A.39: Precision Results: 30% Cluster Size, SaTScan

Figures A.25 - A.29 plot true positives versus false positives for SaTScan under each cluster scenario.

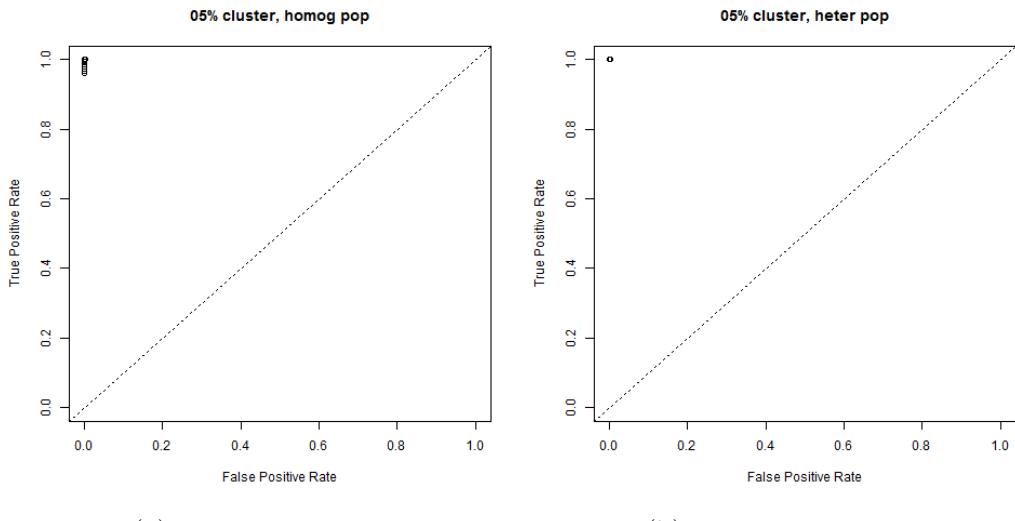


Figure A.40: ROC Curves: 5% Cluster Size, SaTScan

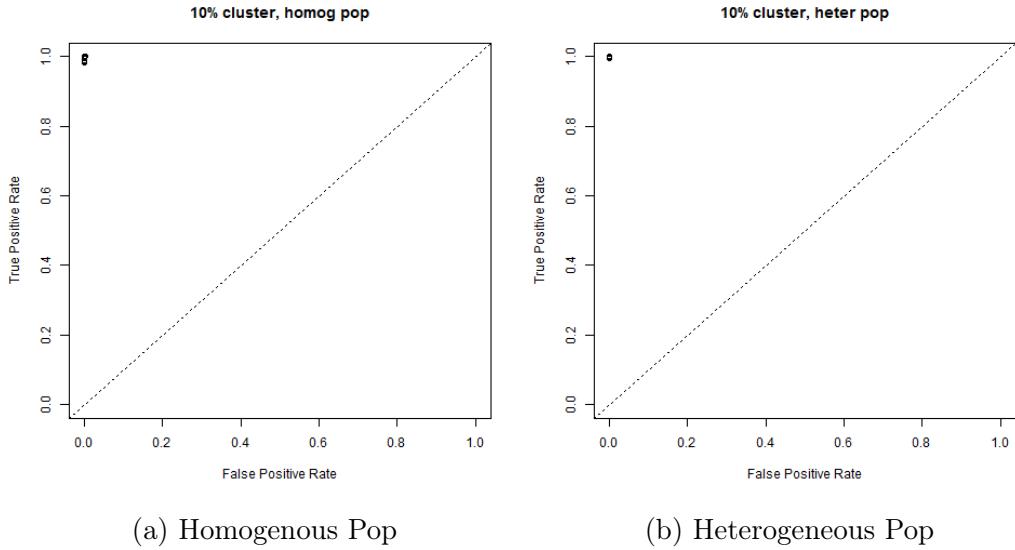


Figure A.41: ROC Curves: 10% Cluster Size, SaTScan

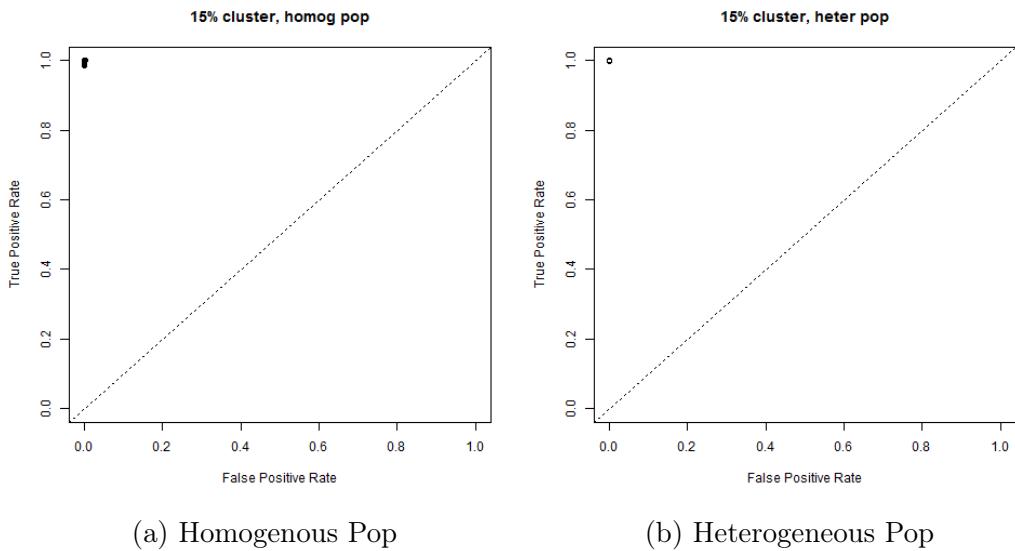


Figure A.42: ROC Curves: 15% Cluster Size, SaTScan

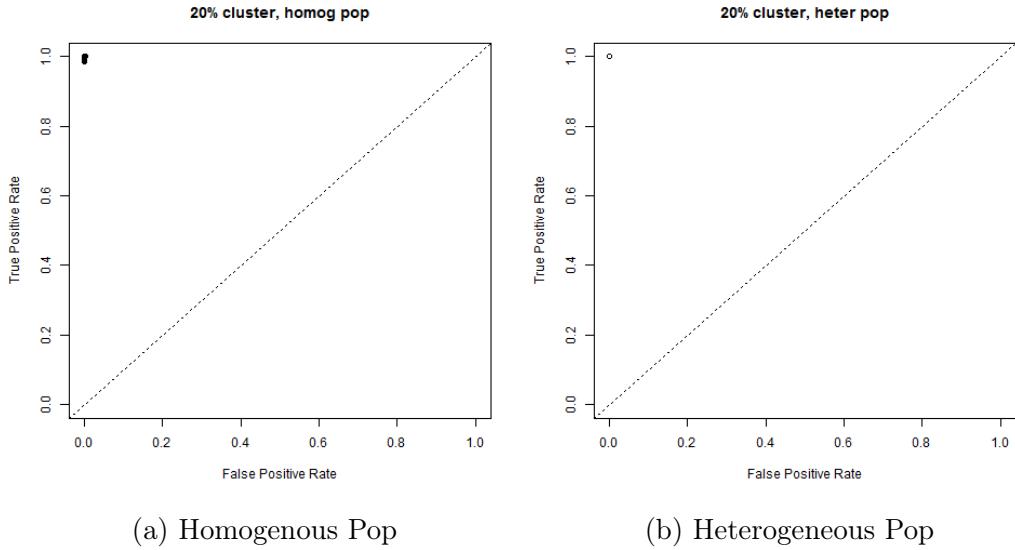


Figure A.43: ROC Curves: 20% Cluster Size, SaTScan

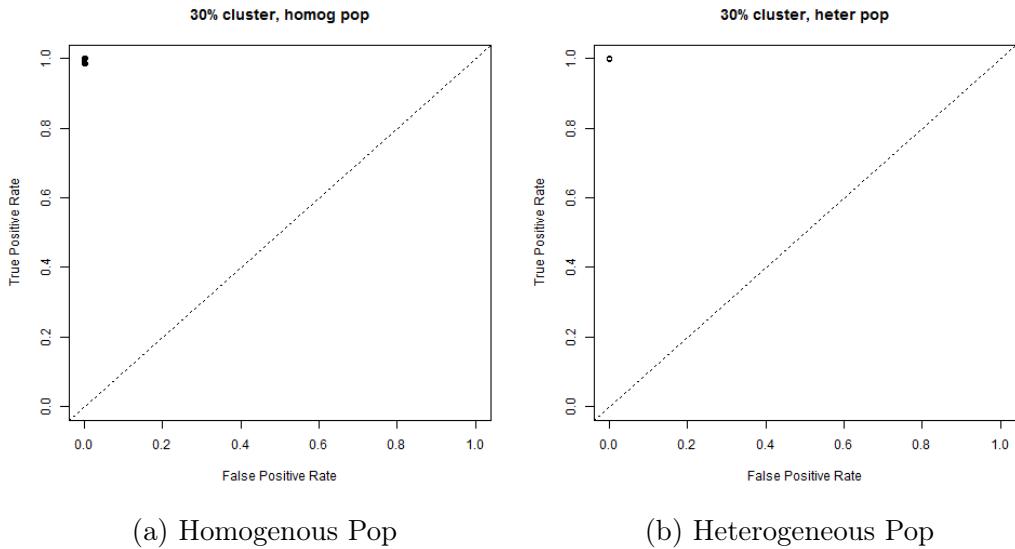


Figure A.44: ROC Curves: 30% Cluster Size, SaTScan

Figures A.45 - A.49 shows the chance of each county being detected as clusters.

Figure A.45: 5% Cluster Size, Heterogeneous Population

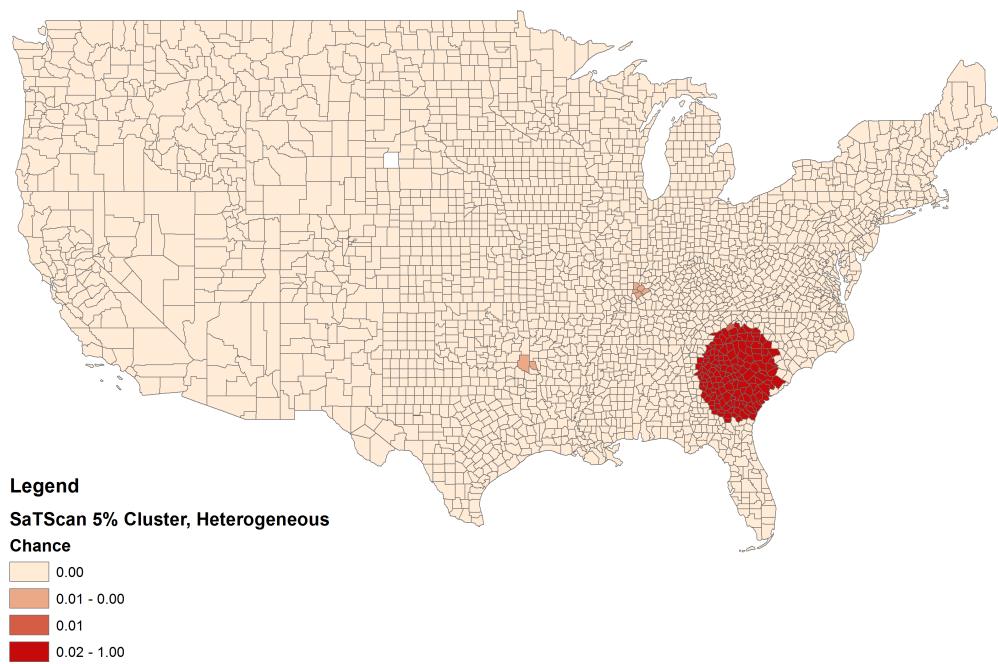


Figure A.46: 10% Cluster Size, Heterogeneous Population

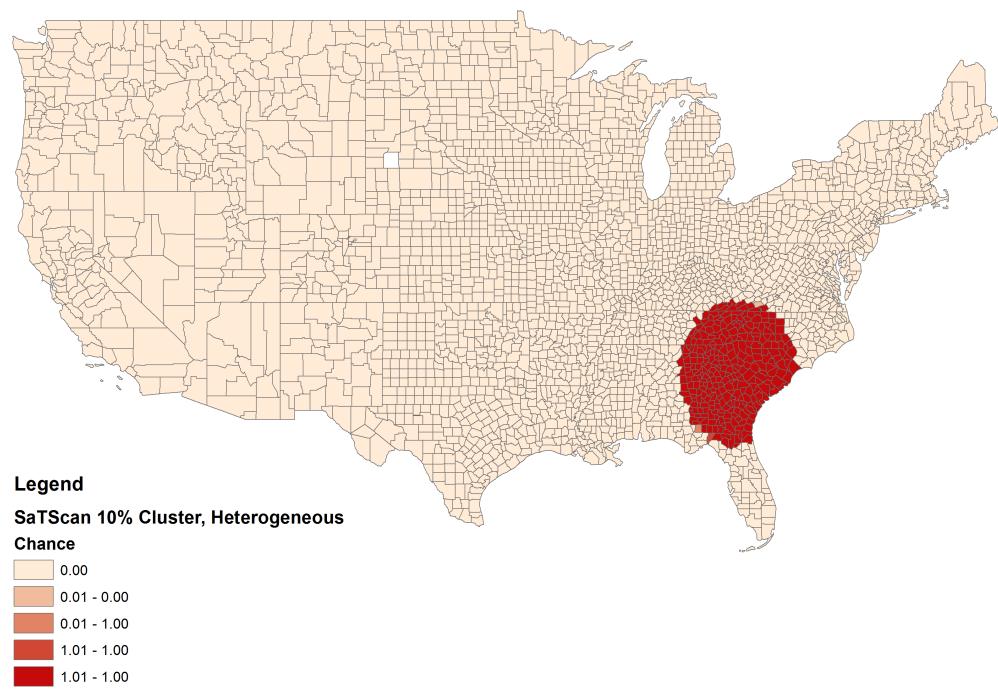


Figure A.47: 15% Cluster Size, Heterogeneous Population

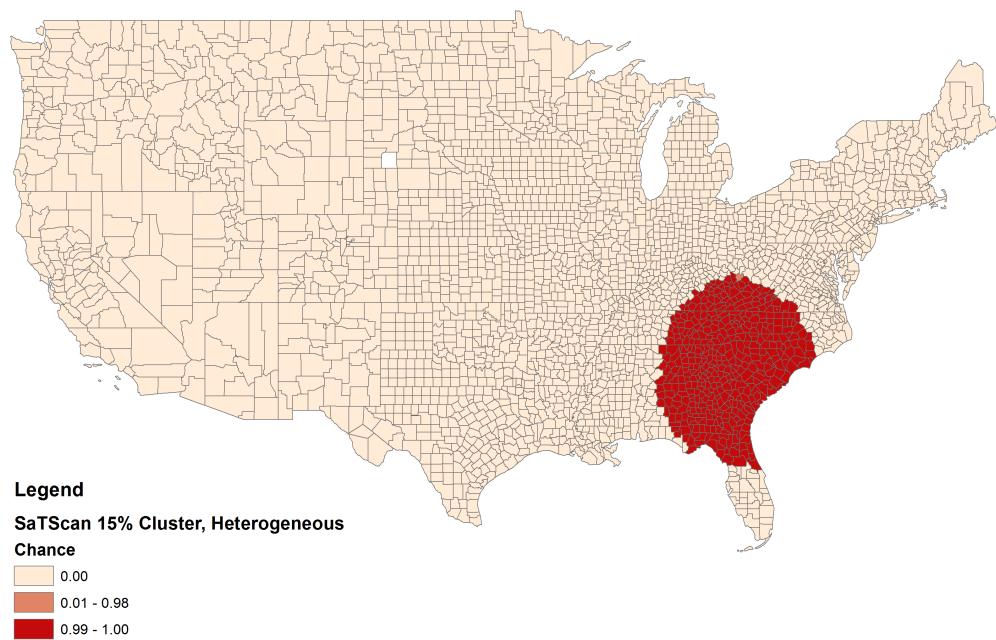


Figure A.48: 20% Cluster Size, Heterogeneous Population

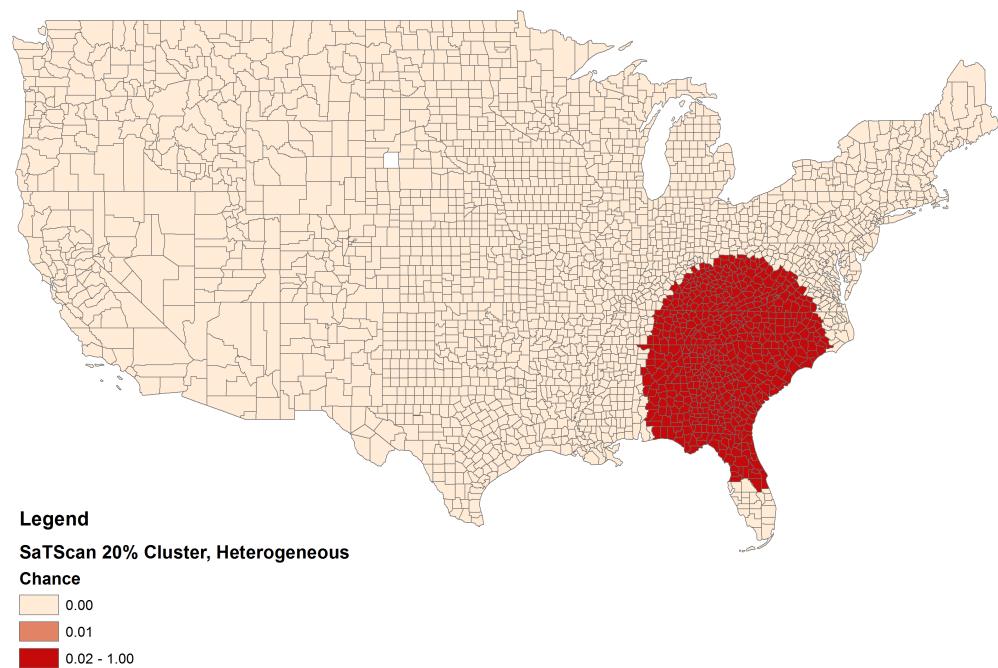
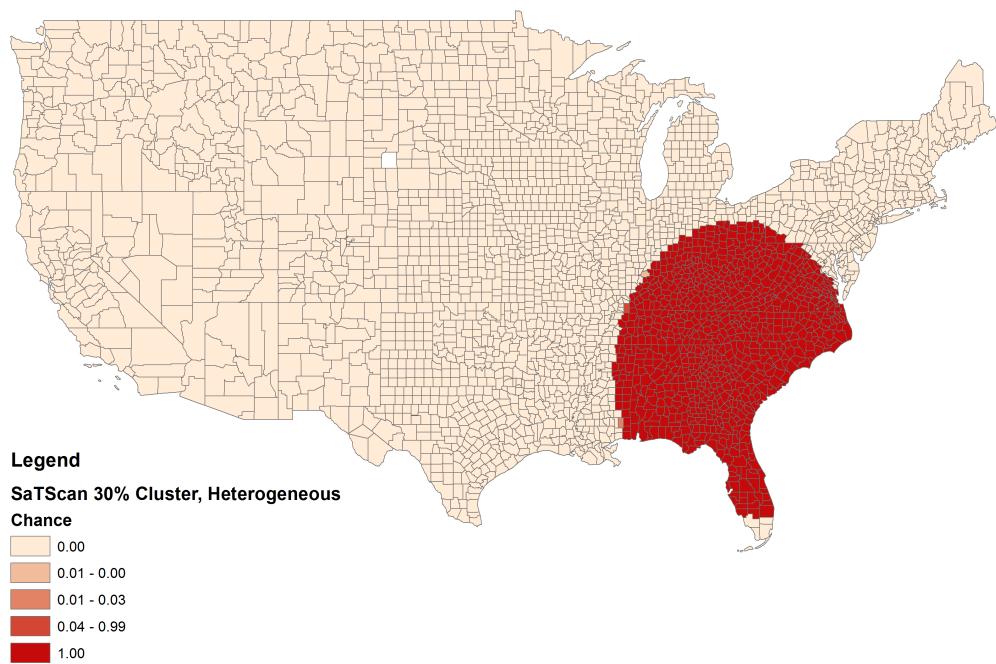


Figure A.49: 30% Cluster Size, Heterogeneous Population



References

- [1] M. C. Jackson, L. Huang, Q. Xie, R. C. Tiwari, A modified version of Moran's I, *International Journal of Health Geographics* 9 (1) 33, doi:10.1186/1476-072X-9-33, URL <http://dx.doi.org/10.1186/1476-072X-9-33>.
- [2] T. Tango, A test for spatial disease clustering adjusted for multiple testing, *Statistics in Medicine* 19 (2) (2000) 191–204.
- [3] L. Anselin, Local Indicators of Spatial Association - LISA, *Geographical Analysis* 27 (2) (1995) 93–115, ISSN 1538-4632, doi:10.1111/j.1538-4632.1995.tb00338.x, URL <http://dx.doi.org/10.1111/j.1538-4632.1995.tb00338.x>.

- [4] W. R. Tobler, A computer movie simulating urban growth in the Detroit region, *Economic Geography* 46 (1970) 234–240.
- [5] L. A. Waller, G. M. Jacquez, Disease Models Implicit in Statistical Tests of Disease Clustering, *Epidemiology* 6 (6) (1995) 584–590, ISSN 10443983, URL <http://www.jstor.org/stable/3703132>.
- [6] N. R. Seabrook, The Obama effect: Patterns of geographic clustering in the 2004 and 2008 presidential elections, *The Forum* 7 (2).
- [7] S. J. Rey, B. D. Montouri, US Regional Income Convergence: A Spatial Econometric Perspective, *Regional Studies* 33 (2) (1999) 143–156, doi:10.1080/00343409950122945, URL <http://dx.doi.org/10.1080/00343409950122945>.
- [8] J. Dockerty, K. Sharples, B. Borman, An assessment of spatial clustering of leukaemias and lymphomas among young people in New Zealand, *Journal of Epidemiology and Community Health* 53 (1999) 154–158.
- [9] B. W. Turnbull, E. J. Iwano, W. S. Burnett, H. L. Howe, L. C. Clark, Monitoring for clusters of disease: application to leukemia incidence in upstate New York, *American Journal of Epidemiology* 132 (supp1) (1990) 136–143.
- [10] D. C. Wheeler, L. A. Waller, W. Cozen, M. H. Ward, Spatial-temporal analysis of non-Hodgkin lymphoma risk using multiple residential locations, *Spatial and Spatio-temporal Epidemiology* 3 (2) (2012) 163–171.
- [11] L. Waller, C. Gotway, *Applied Spatial Statistics for Public Health Data*, Wiley Series in Probability and Statistics, Wiley, ISBN 9780471662679, URL <https://books.google.com/books?id=0uQwgShUdGAC>, 2004.
- [12] M. Kulldorff, A spatial scan statistic, *Communications in Statistics - Theory and Methods* 26 (6) (1997) 1481–1496.
- [13] P. A. Moran, Notes on continuous stochastic phenomena, *Biometrika* (1950) 17–23.
- [14] Environmental Systems Research Institute, Cluster and Outlier Analysis: Anselin Local Moran's I (Spatial Statistics), URL http://resources.esri.com/help/9.3/arcgisengine/java/gp_toolref/

`spatial_statistics_tools/cluster_and_outlier_analysis_colon_anselin_local_moran_s_i_spatial_statistics_.htm`, 2016.

- [15] L. A. Waller, E. G. Hill, R. A. Rudd, The geography of power: Statistical performance of tests of clusters and clustering in heterogeneous populations, *Statistics in Medicine* 25 (5) (2006) 853–865, ISSN 1097-0258, doi:10.1002/sim.2418.
- [16] M. C. Jackson, L. Huang, J. Luo, M. Hachey, E. Feuer, Comparison of tests for spatial heterogeneity on data with global clustering patterns and outliers, *International Journal of Health Geographics* 8 (1) (2009) 55.