The Moran Coefficient and the Geary Ratio: Some mathematical and numerical comparisons

Qing Luo[1,2], Daniel A. Griffith[2], Huayi Wu[1]

[1]Wuhan University, 129 Luoyu Road, Hongshan district, Wuhan, Hubei, 430079

[2]The University of Texas at Dallas, 800 W. Campbell Rd Richardson, Texas 75080-3021
Email: luoqing11@whu.edu.cn, dagriffith@utdallas.edu, wuhuayi@whu.edu.cn

Abstract

This paper discusses relationships between the Moran Coefficient (MC; Moran, 1950) and the Geary Ratio (GR; Geary, 1954) under different distributional assumptions [normal, uniform, beta, and exponential] and selected geographic neighbourhood definitions (linear, square rook, hexagon, square queen, maximum planar, maximum hexagon, and a constant neighbor). It focuses on comparisons of efficiency and power for the MC and the GR. Its results should inform features of spatial data analysis.

Keywords: MC, GR, efficiency, power, geographic configuration.

1. Introduction

The MC and GR are statistics used to quantify the nature and degree of spatial autocorrelation. Cliff and Ord (1973, 1981) established their asymptotic normal sampling distribution properties and the power superiority of the MC versus the GR for only a few types of selected small surface partitionings. Tiefelsdorf and Boots (1995) derived the exact distribution of the MC for small samples, which is seminal work establishing the novel eigenvector spatial filtering spatial statistics methodology (Griffith, 1996). This paper explores relationships between the MC and GR for a wide range of surface partitonings across sizes that expand to infinity, derives their approximate variances under different distributional assumptions, analyzes their statistical efficiency, presents large sample power comparisons for them, and documents some comparative features.

2. The relationship between the MC and GR

Let X be the georeferenced variable of interest distributed over a tessellation. Its observations are $x_1, x_2, \cdots, x_n$. The average of these observations is denoted by

$$\bar{x} = \sum_{i=1}^{n} x_i \Big/ n. \tag{1}$$

For regular surface partitionings, $n = P \times Q$, where P and Q respectively are the number of rows and columns. Let $\boldsymbol{C} = \left(c_{ij}\right)_{n \times n}$ be a surface partitioning's connectivity matrix, where $c_{ij} = 1$ if i and j are adjacent (i.e., neighbors), and 0 otherwise; $\boldsymbol{C}$ is symmetric.

The sample MC and GR of variable X are defined by

$$MC = n \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} \left(x_i - \bar{x}\right)\left(x_j - \bar{x}\right) \Big/ \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} \sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2, \tag{2}$$

$$GR = (n-1)\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}\left(x_i - x_j\right)^2 \Big/ 2\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 . \tag{3}$$

The GR can be rewritten as (Griffith, 1987, p.44)

$$\frac{n-1}{2\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}} \frac{2\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2\left(\sum_{j=1}^{n} c_{ij}\right)}{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2} - \frac{n-1}{n} MC. \tag{4}$$

Proof. Substituting equation (2) into equation (4) yields

$$GR = (n-1)\left[2\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2\left(\sum_{j=1}^{n} c_{ij}\right) - 2\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}\left(x_i - \bar{x}\right)\left(x_j - \bar{x}\right)\right] \Big/ \left[2\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2\right]. \tag{5}$$

Substituting $\left(x_i - x_j\right)^2 = \left[\left(x_i - \bar{x}\right) - \left(x_j - \bar{x}\right)\right]^2$, and utilize the symmetry of matrix $\boldsymbol{C}$, yields GR = equation (4). ■

3. Derivation of the MC and GR asymptotic variances

The exact variances of these two statistics (Cliff, and Ord, 1973), where subscript N denotes normality and R denotes randomization, are

$$Var_N\left(MC\right) = \frac{n^2 S_1 - nS_2 + 3S_0^{\ 2}}{(n-1)(n+1)S_0^{\ 2}} - \frac{1}{(n-1)^2}, \tag{6}$$

$$Var_R\left(MC\right) = \frac{n\left[\left(n^2 - 3n + 3\right)S_1 - nS_2 + 3S_0^{\ 2}\right] - b_2\left[\left(n^2 - n\right)S_1 - 2nS_2 + 6S_0^{\ 2}\right]}{(n-1)(n-2)(n-3)S_0^{\ 2}} - \frac{1}{(n-1)^2}, \tag{7}$$

$$Var_N\left(GR\right) = \left[\left(2S_1 + S_2\right)(n-1) - 4S_0^{\ 2}\right] \Big/ \left[2(n+1)S_0^{\ 2}\right], \tag{8}$$

$$Var_R\left(GR\right) = \frac{(n-1)S_1\left[n^2 - 3n + 3 - (n-1)b_2\right] - \frac{1}{4}(n-1)S_2\left[n^2 + 3n - 6 - \left(n^2 - n + 2\right)b_2\right]}{n(n-2)(n-3)S_0^{\ 2}} + \frac{S_0^{\ 2}\left[n^2 - 3 - (n-1)^2 b_2\right]}{n(n-2)(n-3)S_0^{\ 2}} \tag{9}$$

where $S_0 = \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}, S_1 = \sum_{i=1}^{n}\sum_{j=1}^{n}\left(c_{ij} + c_{ji}\right), S_2 = \sum_{i=1}^{n}\left(\sum_{j=1}^{n}\left(c_{ij} + c_{ji}\right)\right)^2$, and for $z_i = x_i - \bar{x}$, $b_2 = \frac{1}{n}\sum_{i=1}^{n} z_i^{\ 4} \Big/ \left(\frac{1}{n}\sum_{i=1}^{n} z_i^{\ 2}\right)^2$ is kurtosis. Because $\boldsymbol{C}$ is symmetric, $S_1 = 2\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij} = 2S_0$, and $S_2 = 4\sum_{i=1}^{n}\left(\sum_{j=1}^{n} c_{ij}\right)^2$.

Griffith (2010) proposes simplifying equations (6)-(9) through asymptotics assuming a normal distribution, producing

$$Var_A\left(MC\right) = 2 \Big/ \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij} = 2/S_0, \tag{10}$$

$$Var_A(GR) = 2\Big/\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij} + 2\sum_{i=1}^{n}\left(\sum_{j=1}^{n} c_{ij}\right)^2 \Big/ \left(\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}\right)^2 = \frac{2}{S_0} + \frac{S_2}{2S_0^{\ 2}} = \frac{2S_0 + S_2}{2S_0^{\ 2}}, \qquad (11)$$

where the subscript A denotes asymptotic.

The asymptotic variance for the MC is insensitive to the normality and randomization assumptions.

Theorem 1 $\lim_{n\to\infty} Var_N(MC) = Var_A(MC)$.

Proof. $\lim_{n\to\infty} Var_N(MC) = \lim_{n\to\infty}\left[equation(6)\right] = 2/S_0 = Var_A(MC)$. ■

Theorem 2 $\lim_{n\to\infty} Var_R(MC) = Var_A(MC)$.

Proof. $\lim_{n\to\infty} Var_R(MC) = \lim_{n\to\infty}\left[equation(7)\right] = 2/S_0 = Var_A(MC)$. ■

In contrast, the asymptotic variance of the GR is sensitive to the normality and randomization assumptions.

Theorem 3 $\lim_{n\to\infty} Var_N(GR) = Var_A(GR)$.

Proof. $\lim_{n\to\infty} Var_N(GR) = \lim_{n\to\infty}\left[equation(8)\right] = (2/S_0) + (S_2/2S_0^{\ 2}) = Var_A(GR)$. ■

Theorem 4 $\lim_{n\to\infty} Var_R(GR)$ depends on b_2, the kurtosis of a distribution.

Proof. $\lim_{n\to\infty} Var_R(GR) = \lim_{n\to\infty}\left[equation(9)\right] = (2/S_0) + \left[(b_2 - 1)S_2/4S_0^{\ 2}\right]$. ■

For the normal, uniform, beta ($\alpha = \beta = 0.5$), and exponential distributions, the b_2 values are 3, 9/5, 3/2, and 9, respectively, yielding asymptotic variances for GR of

$$Var_{AN}(GR) = 2/S_0 + S_2/2S_0^{\ 2}, \qquad (12)$$

$$Var_{AU}(GR) = 2/S_0 + S_2/5S_0^{\ 2}, \qquad (13)$$

$$Var_{AB}(GR) = 2/S_0 + S_2/8S_0^{\ 2}, (\alpha = \beta = 0.5), \qquad (14)$$

$$Var_{AE}(GR) = 2/S_0 + 2S_2/S_0^{\ 2}, \qquad (15)$$

where the subscripts AN, AU, AB, and AE respectively denote the selected distributions. Equation (12) coincidences with Griffith's (2010) result.

4. Efficiency analysis

A statistic with a smaller variance is more efficient. The variance ratio of the MC and the GR may be defined by

$$r_{exact} = Var_{exact}(MC)/Var_{exact}(GR), \qquad (16)$$

where subscript *exact* denotes the exact MC and GR variances, given by equations (6) and (8), or (7) and (9).

The following asymptotic variances also are of interest:

$$r = Var_A(MC)/Var_{A*}(GR) = 2/S_0/S, \qquad (17)$$

where $A*$ denotes AN, AU, AB or AE, and S denotes equation (12), (13), (14), or (15). If $r < 1$, then the MC is more efficient than the GR.

Five different geographic configurations rendering different connectivity matrices illuminate the range of possible geographic situations (Table 1). Analysis here focuses on their respective variance ratios, which converge to 1 in the limit for all but maximum planar connectivity.

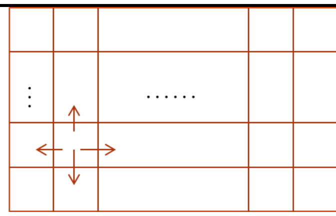
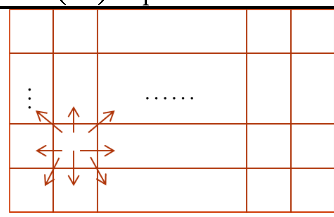
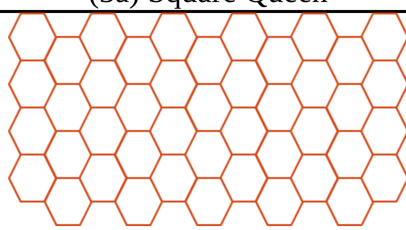
Geographic Configurations	Neighbour sums
 (1a) Linear	$\sum_{i=1}^{n}\sum_{j=1}^{n}c_{ij}=2(n-1)$, $\sum_{i=1}^{n}\left(\sum_{j=1}^{n}c_{ij}\right)^2=4n-6$ (1b)
 (2a) Square Rook	$\sum_{i=1}^{n}\sum_{j=1}^{n}c_{ij}=2\left[P(Q-1)+Q(P-1)\right]$ $\sum_{i=1}^{n}\left(\sum_{j=1}^{n}c_{ij}\right)^2=2(8PQ-7P-7Q+4)$ $(n=P\times Q)$ (2b)
 (3a) Square Queen	$\sum_{i=1}^{n}\sum_{j=1}^{n}c_{ij}=2(4PQ-3P-3Q+2)$ $\sum_{i=1}^{n}\left(\sum_{j=1}^{n}c_{ij}\right)^2=2(32PQ-39P-39Q+46)$ $(n=P\times Q)$ (3b)
 (4a) Hexagon	$\sum_{i=1}^{n}\sum_{j=1}^{n}c_{ij}=2(3PQ-2P-2Q+1)$ $\sum_{i=1}^{n}\left(\sum_{j=1}^{n}c_{ij}\right)^2=2(18PQ-20P-19Q+19)$ $(n=P\times Q)$ (4b)
 (5a) Maximum Planar Connectivity	$\sum_{i=1}^{n}\sum_{j=1}^{n}c_{ij}=6(n-2)$ $(n=Q+2)$ $\sum_{i=1}^{n}\left(\sum_{j=1}^{n}c_{ij}\right)^2=2(n^2+6n-22)$ (5b)

Table 1. Areal unit configuration cases and their neighbor sums

In Table 1, (5a) portrays an unlikely connectivity scheme, which differs from other cases. This configuration has Q+2 units (P-by-Q+2 becomes 1-by-Q+2), with two units adjacent to all Q+1 other units. Table 2 also introduces maximum hexagon connectivity (1a-2a). The number of units in both configurations can be expressed as P-by-Q+2. In addition, some of the configurations in Table 1 can be generalized to three dimensions (3-D). For example, two ends of a linear landscape (Table 1, 1a) can be connected so that it becomes a closed circle; two pairs of the opposite ends of a square partitioning can be connected so that it becomes a torus. Table 2 portrays these 3-D cases (3a-4a); their neighbor sums are $\sum_{i=1}^{n}\sum_{j=1}^{n}c_{ij}=kn$, $\sum_{i=1}^{n}\left(\sum_{j=1}^{n}c_{ij}\right)^2=k^2n$, where k is a constant (2, 4 or 8 here).

Many remote sensing images can be divided into square lattices, and various connectivity criteria can be employed to establish their spatial adjacency.

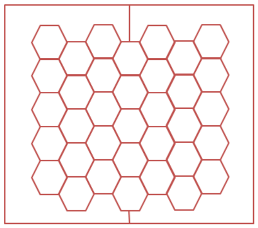
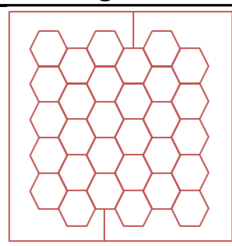
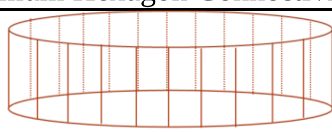
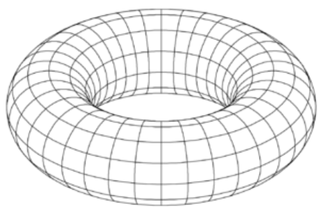
Geographic Configurations	Neighbour sums
 (1a) Maximum Hexagon Connectivity, odd Q	$\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij} = 6PQ \quad (n = P\times Q+2)$ $\sum_{i=1}^{n}\left(\sum_{j=1}^{n} c_{ij}\right)^2 = 2(P^2+Q^2+20PQ$ $-11P-10Q+6)$ (1b)
 (2a) Maximum Hexagon Connectivity, even Q	$\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij} = 6PQ \quad (n = P\times Q+2)$ $\sum_{i=1}^{n}\left(\sum_{j=1}^{n} c_{ij}\right)^2 = 2(P^2+Q^2+20PQ$ $-11P-10Q+8)$ (2b)
 (3a) Circle connectivity	$\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij} = 2n$, $\sum_{i=1}^{n}\left(\sum_{j=1}^{n} c_{ij}\right)^2 = 4n$ $(n = P\times Q)$ (3b)
 (4a) Torus connectivity	$\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij} = 4n$, $\sum_{i=1}^{n}\left(\sum_{j=1}^{n} c_{ij}\right)^2 = 16n$ (rook) $\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij} = 8n$, $\sum_{i=1}^{n}\left(\sum_{j=1}^{n} c_{ij}\right)^2 = 64n$ (queen) $(n = P\times Q)$ (4b)

Table 2. Other connectivity cases and their neighbor sums

4.1 Normal variance ratios

Substituting equations (10) and (12) into equation (17) yields $\left[Var_A(MC)/Var_{AN}(GR)\right]_* = S_0/(S_0+S_2/4)$, where * denotes linear (L), square rook (SR), square queen (SQ), hexagon (H), maximum planar (MP), maximum hexagon (MH), and constant neighbors (CN) connectivity. Substituting corresponding S_0 and S_2 values (see Table 1 (1b-5b), Table 2 (1b-4b)), and letting $n\to\infty$, yields the asymptotic variance ratios between the MC and GR.

4.2 Uniform variance ratios

Repeating the steps in §4.1, but with equation (13) rather than (12), yields $\left[Var_A(MC)/Var_{AU}(GR)\right]_* = S_0/(S_0+S_2/10)$, and then the asymptotic variance ratios.

4.3 Beta variance ratios

Repeating the steps in §4.1, but with equation (14) rather than (12), yields $\left[Var_A(MC)/Var_{AB}(GR)\right]_* = S_0/(S_0+S_2/16)$, and then the asymptotic variance ratios.

4.4 Exponential variance ratios

Repeating the steps in §4.1, but with equation (15) rather than (12), yields $\left[Var_A(MC)/Var_{AE}(GR)\right]_* = S_0/(S_0+S_2)$, and then the asymptotic variance ratios.

Table 3 summarizes these asymptotic variance ratios, and their related exact counterparts and adjustment factors for asymptotic variances. Except for the maximum planar case, all exact variance ratios are 1; hence, asymptotic ratios need to be adjusted for GRs. This furnishes quantitative evidence that the GR, unlike the MC, is far more sensitive to the underlying frequency distribution of a variable.

Distribution \ Landscape		Linear	Square Rook	Hexagon	Square Queen	Maximum Planar	Maximum Hexagon	Constant Neighbors
Normal	**AVR**	1/3	1/5	1/7	1/9	0	3/25	1/(k+1)
	EVR	1	1	1	1	0	3/7	1
	AVMC	1	1	1	1	1/3	1	1
	AVGR	1/3	1/5	1/7	1/9	1	7/25	1/(k+1)
Uniform	**AVR**	5/9	5/13	5/17	5/21	0	15/59	5/(5+2k)
	EVR	1	1	1	1	0	0.6522	1
	AVMC	1	1	1	1	1/3	1	1
	AVGR	5/9	5/13	5/17	5/21	1	1/2.565	5/(5+2k)
Beta (a=b=0.5)	**AVR**	2/3	1/2	2/5	1/3	0	6/17	4/(4+k)
	EVR	1	1	1	1	0	3/4	1
	AVMC	1	1	1	1	1/3	1	1
	AVGR	2/3	1/2	2/5	1/3	1	8/17	4/(4+k)
Exponential	**AVR**	1/9	1/17	1/25	1/33	0	3/91	1/(1+4k)
	EVR	1	1	1	1	0	0.1579	1
	AVMC	1	1	1	1	1/3	1	1
	AVGR	1/9	1/17	1/25	1/33	1	1/4.79	1/(1+4k)

Table 3. Asymptotic-to-exact variance ratios and adjustment factors (= *VMC/*VGR. AVR or EVR: asymptotic or exact variance ratio. AVMC or GR: AVR of MC or GR).

Landscape	***Abs(Asy/act-1)<=0.025***		***Abs(Act-asy)<=0.01***	
	MC	**GR**	**MC**	**GR**
Linear	42	56	23	27
Square Rook	88	12	36	11
Square Queen	161	7	37	93
Hexagon	121	7	34	72
Maximum Planar	15	403	10	333
Maximum Hexagon	157	14	34	5438
Circle	43	83	23	35
Torus Rook	84	124	29	36
Torus Queen	167	207	37	41

Table 4. Minimum sample sizes for the MC and the GR.

Table 4 presents the minimum sample size for which the asymptotic variances approximate their exact variance counterparts, and indicates that the MC asymptotic variance furnishes a useful result for relatively small sample sizes.

4.5 Variance ratio convergence

Figure 1 portrays exact variance ratio curves as well as values for 184 specimen irregular surface partitions. Figures 1a-4a are ideal increasing sample size ratio trajectories for normal, uniform, beta, and exponential random variables, respectively, which respectively depict convergence in the interval [13, 100], [10, 100], [8, 100], [23, 100]. Figures 1b-4b are the same curves extended to n = 7,250, with the 184 specimen surface values superimposed (black dots).

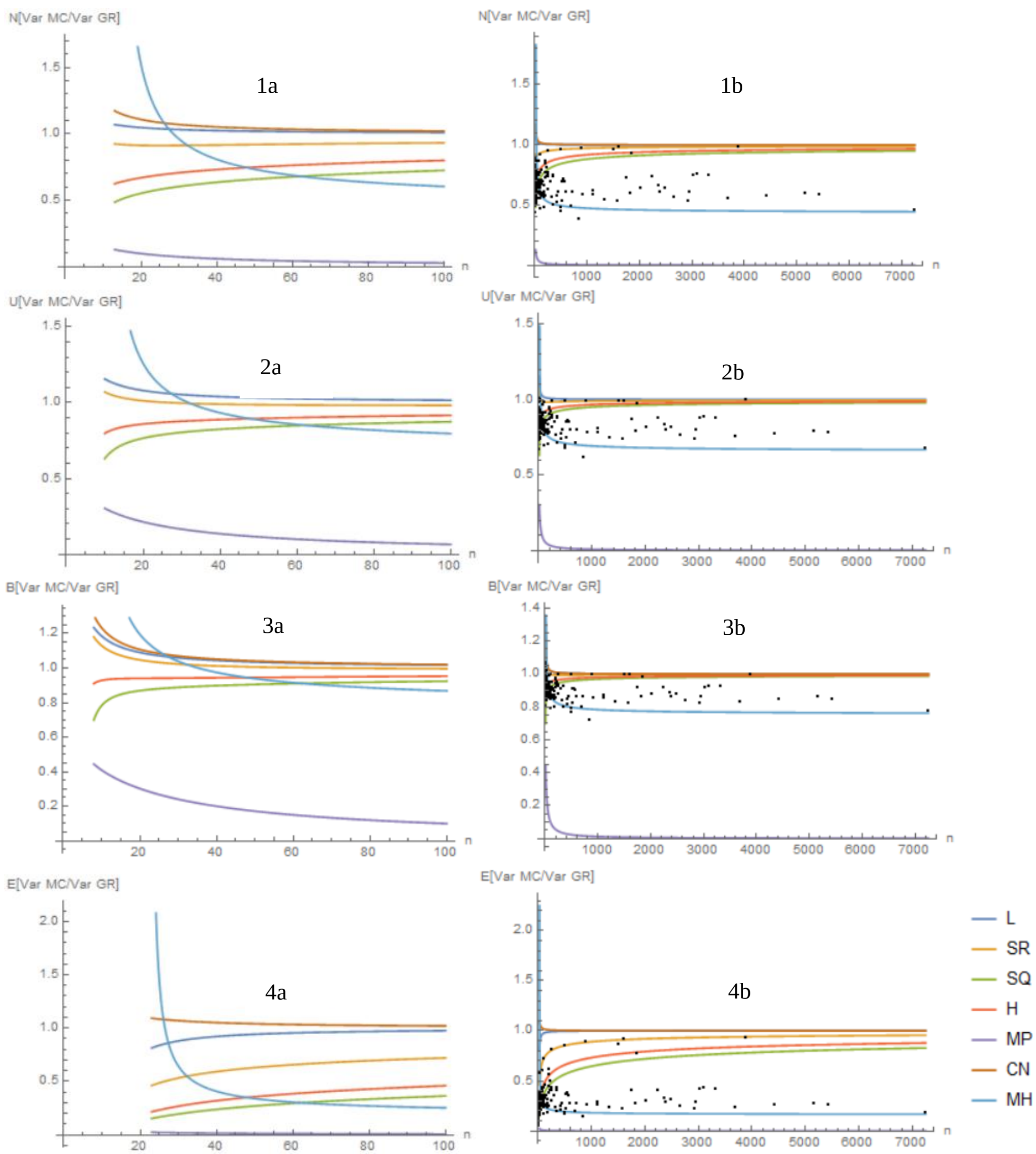


Figure 1. Exact variance ratio curves.

5. A power comparison

Cliff and Ord (1973) conducted limited simulation experiments comparing the power of the MC and GR, concluding that the MC is more powerful. One remaining unanswered question asks what happens across a wider range of sample sizes and attribute variable types.

5.1 Establishing statistical power

The power of a test is $1-\beta$, where β is the probability of committing a Type II error (i.e., failing to reject the null hypothesis when it is false) for a given significance level α, which is the probability of committing a Type I error (i.e., rejecting the null hypothesis when it is true). If $1-\beta_{MC} > 1-\beta_{GR}$ (i.e., the Type II error probability for the MC is less than the Type II error probability for the GR), then the MC is more powerful than the GR.

Step 1. Let the null and alternative hypotheses be

H_0 : *No spatial autocorrelation*; H_1 : *Nonzero spatial autocorrelation.*

Suppose $\alpha = 0.05$.

Step 2a. Consider the Moran test assuming normality defined by equation (18), where $Var_N(MC)$ is equation (6). The critical value (CV) is $z_{MC} = \left[MC + 1/(n-1)\right] / \sqrt{Var_N(MC)}$.

$$N_{MC}\left(-\frac{1}{n-1}, Var_N(MC)\right) \tag{18}$$

If $|z_{MC}| < 1.96$, then the statistical decision is to fail to reject *H₀*.

Step 2b. Parallel results for the GR include

$$N_{GR}\left(1, Var_N(GR)\right) \tag{19}$$

where $Var_N(GR)$ is equation (8), and $z_{GR} = (GR-1)/\sqrt{Var_N(GR)}$. If $|z_{GR}| < 1.96$, then the statistical decision is fail to reject H_0.

Step 3a. If the statistical decision is fail to reject H_0 when it is not true, a Type II error occurs. The CV under the true sampling distribution is given by

$$N\left(a, Var_N(MC)\right)\left(a \in [-1,1], a \neq -1/(n-1) \right). \tag{20}$$

The CV under the null distribution [equation (18)] is

$$z_{CV} = \pm 1.96\sqrt{Var_N(MC)} - 1/(n-1), \tag{21}$$

The CV under the true distribution [equation (19)] is $z_{ts} = (z_{CV} - a)/\sqrt{Var_N(MC)}$. Substituting equation (21) into this equation yields $z_{ts} = \pm 1.96 - \frac{1}{\sqrt{Var_N(MC)}}\left(\frac{1}{n-1} + a\right)$, where the subscript *ts* denotes the standard CV under the true distribution. Given $z_{\alpha/2} = 1.96 - \frac{1}{\sqrt{Var_N(MC)}}\left(\frac{1}{n-1} + a\right)$ and $-z_{\alpha/2} = -1.96 - \frac{1}{\sqrt{Var_N(MC)}}\left(\frac{1}{n-1} + a\right)$, the power of the MC is

$$1-\beta_{MC} = 1 - \mathrm{P}\left(x \le z_{\alpha/2}\right) + \mathrm{P}\left(x \le -z_{\alpha/2}\right). \tag{22}$$

Step3b. For the GR, the CV under the true distribution is equation (23), and the CV under the null distribution [equation (20)] is equation (24). The CV under the true distribution is $z'_{ts} = \left(z'_{CV} - \mu\right) / \sqrt{Var_N(GR)}$, where μ is the mean in equation (23). Substituting equation (24) into this formula yields equation (25).

$$N\left(\frac{n-1}{\sum_{i=1}^{n}\sum_{j=1}^{n}c_{ij}}\frac{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2\left(\sum_{j=1}^{n}c_{ij}\right)}{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2}-\frac{n-1}{n}a, Var_N(GR)\right)\left(a\in[-1,1], a\neq -\frac{1}{n-1}\right). \quad (23)$$

$$z'_{cv} = \pm 1.96\sqrt{Var_N(GR)}+1, \quad (24)$$

$$z'_{ts} = \pm 1.96 + \frac{1}{\sqrt{Var_N(GR)}} - \frac{(n-1)\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2\left(\sum_{j=1}^{n}c_{ij}\right)}{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2\sqrt{Var_N(GR)}} + \frac{a(n-1)S_0}{n\sqrt{Var_N(GR)}}, \quad (25)$$

Defining $z'_{\alpha/2}$ and $-z'_{\alpha/2}$ in the same way as $z_{\alpha/2}$ and $-z_{\alpha/2}$, the power of the GR is

$$1-\beta_{GR} = 1 - \text{P}\left(x \le z'_{\alpha/2}\right) + \text{P}\left(x \le -z'_{\alpha/2}\right). \quad (26)$$

Figures 2 and 3 present selected power curves plotted with their respective minimum sample sizes. The MC is more powerful than the GR for the rook configuration, while the GR is more powerful than the MC for the torus rook connectivity. For the hexagon and maximum hexagon cases, the MC is more powerful than the GR for positive spatial autocorrelation, but not always for negative spatial autocorrelation.

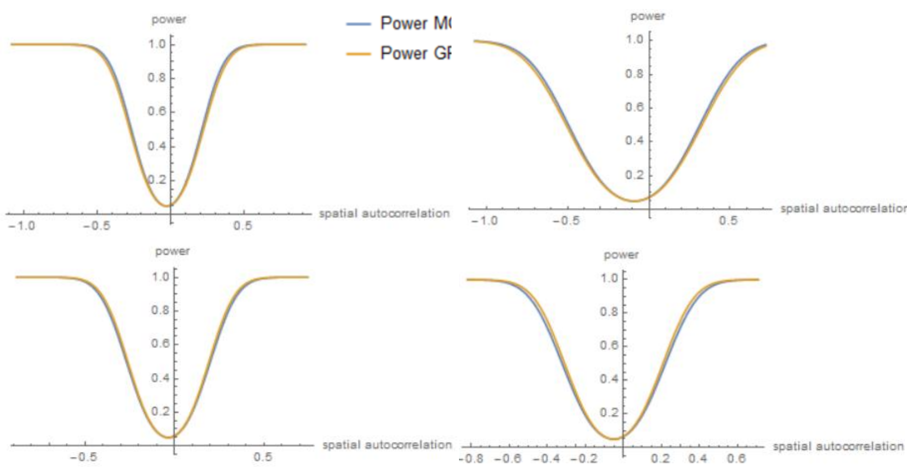


Figure 2. Planar and torus rook square surface partitionings. (a) Top left: square rook, P = 6, Q = 6; (b) Top right: square rook, P = 3, Q = 4; (c) Bottom left: torus rook, P = 5, Q = 6; (d) Bottom right: torus rook, P = 6, Q = 6.

5.2 Theoretical evaluation

Equation (2) can be rewritten using matrix notation such that $\left(\boldsymbol{I} - \frac{\boldsymbol{1}\boldsymbol{1}^T}{n}\right)\boldsymbol{C}\left(\boldsymbol{I} - \frac{\boldsymbol{1}\boldsymbol{1}^T}{n}\right)$ is in the numerator of the MC, where **1** is an n-by-1 vector of ones, and T denotes the matrix transpose operation. The eigenvalues of this matrix times $\frac{n}{\boldsymbol{1}^T\boldsymbol{C}\boldsymbol{1}}$ furnish the complete set of distinct MC values for a geographic landscape, with the extreme values establishing the minimum and maximum possible MC values. Corresponding GR values can be calculated with the eigenvectors of this matrix. The relationship between MC and GR in this context is given by $GR = a + b\left(MC - MC_{\min}\right)^{c}$. For linear (circle) adjacency and a square tessellation with rook (torus rook) adjacency, a = 2, b= -1, and c = 1. For a queen (torus queen) adjacency, a = 1.5, b= -1, and c = 1. For a P-by-Q hexagonal tessellation, a=1.5,

b = -0.99063 - 0.78935$(1/P+1/Q)^{0.87125}$ + 0.00205P/Q, and

c = 1.05828 - 0.87248*$(1/P+1/Q)^{0.57358}$ + 0.00385P/Q.

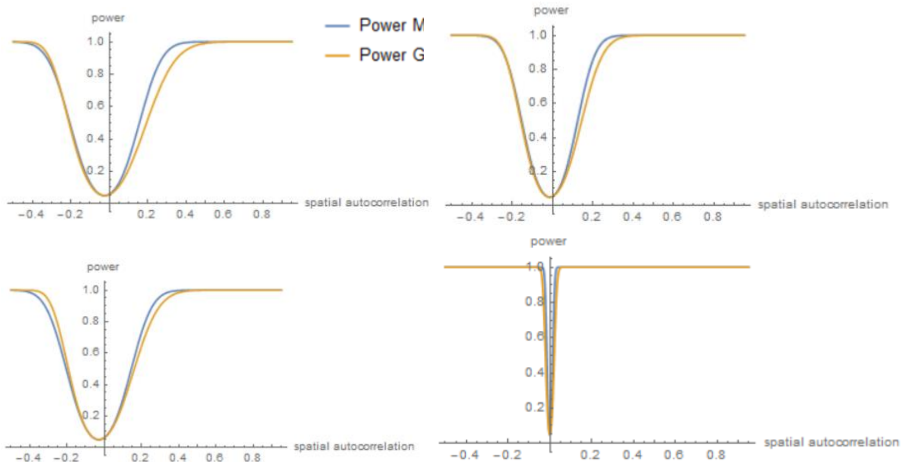


Figure 3. Hexagon and maximum hexagon surface partitionings. (a) Top left: hexagon, P = 5, Q = 7; (b) Top right: hexagon, P = 8, Q = 9; (c) Bottom left: maximum hexagon, P = 5, Q = 7 ;(d) Bottom right: maximum hexagon, P = 73, Q = 74 .

6. Conclusions

The MC may not be uniformly more powerful than the GR for all sample sizes and geographic configurations. The MC asymptotic variances are reliable for most modern day sample sizes, which are greater than 100. The GR asymptotic variance varies with geographic configuration as well as attribute variable type, both of which determine efficiency when geographic sample size goes to infinity. Finally, the relationship between the MC and the GR appears to vary across areal unit configuration types.

7. References

Cliff A and Ord J, 1973, *Spatial Autocorrelation*. Pion, London.
Cliff A and Ord J, 1981, *Spatial Process*. Pion, London
Geary R, 1954, The contiguity ratio and statistical mapping. *The incorporated statistician* , 5(3): 115-145.
Griffith D, 1987, *Spatial Autocorrelation: A Primer*. AAG, Pennsylvania.
Griffith D, 1996, Spatial Autocorrelation and Eigenfunctions of the Geographic Weights Matrix Accompanying Geo-Referenced Data. *The Canadian Geographer*, 40(4): 351-367.
Griffith D, 2010, The Moran coefficient for non-normal data. *Journal of Statistical Planning and Inference*, 140: 2980-2990.
Moran P, 1950, Notes on continuous stochastic phenomena. *Biometrika*, 37: 17-23.
Tiefelsdorf M and Boots B,1995, The exact distribution of Moran's I. *Environment and Planning A*, 27:985-999.