



Rectifying I : three point and continuous fit of the spatial autocorrelation metric, Moran's I , to ideal form

T. J. DeWitt · J. I. Fuentes · T. R. Ioeberger · M. P. Bishop

Received: 11 June 2020 / Accepted: 30 April 2021
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract

Context Interpreting spatial autocorrelation is complicated by differences in data type, spatial conformation, and contiguity definitions. Though lacking consistent meaning, Moran's I is commonly reported, compared, and interpreted based on conceptual ideals. To provide consistent, logical, and intuitive meaning and enable broader synthetic work, a new approach to I is needed.

Objectives We sought to standardize I and true it to conceptual ideals and existing intuition regarding regular correlations. We also wished to test performance of transformed metrics over a diversity of designed and empirical datasets.

Methods We developed two means to rectify I . Both fit null distributions from data permutation to a target

frame of $[-1, 0, 1]$, followed by projection of original I into this conformation. One method used three-point registration employing the distribution median and select tail percentiles. The other directly projected all I based on theory or cumulative frequencies reflecting the distribution of regular correlations. Repeatability and sensitivity of results were examined for varied permutation replication and framing parameter choices. Empirical and designed datasets were used to compare rectified to traditional metrics.

Results Both rectification methods improved distributional characteristics of I . Three-point registration produced overly broad distributions with discontinuous peaks. Continuous projection fit the distribution for regular correlations precisely. Diverse case studies demonstrated failings of I and the clarity gained by rectification.

Conclusions Rectified I enabled meaningful comparisons of spatial patterns for diverse data and landscape conditions. Preserving the intuitive value of Moran's I while providing a theoretically sound and consistent approach for standardizing its values should foster sustained use.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10980-021-01256-0>.

T. J. DeWitt (✉) · J. I. Fuentes
Department of Ecology & Conservation Biology, Texas A&M University, College Station, TX 77843-2258, USA
e-mail: tdewitt@tamu.edu

J. I. Fuentes · T. R. Ioeberger
Department of Computer Science & Engineering, Texas A&M University, College Station, TX 77843-3112, USA

M. P. Bishop
Department of Geography, Texas A&M University, College Station, TX 77843-3147, USA

Keywords Geospatial analysis · Distribution statistics · Statistical null distributions · Monte Carlo permutation

Introduction

Measurements made in proximity to each other tend to be similar (Moran 1947; Durbin and Watson 1950; Tobler 1970; Cliff and Ord 1981). A traditional measure of such spatial autocorrelation is Moran's I (Moran 1947, 1950). I is conceptually elegant for its basic and intuitive nature, although it is often idealized beyond its actual character. Several authors have detailed I 's systematic properties, biases, and limitations (Cliff and Ord 1973; Sen 1976; de Jong et al. 1984; Waldhör 1996; Fortin and Dale 2005). Yet presumptions and mistaken perceptions of I 's character persist.

I is often presented in idealized terms wherein it is said to range between -1 and 1 , with perfect

interspersion at -1 , random dispersion at 0 , and perfect clumping or gradient conformation at 1 (Fig. 1). It is often interpreted as analogous or homologous to a regular (Pearson) correlation coefficient r . Values of I considered small for r are regularly treated as respectively weak autocorrelation. Analogy with r suggests an opposite scale ($\pm I$ indicate the same magnitude of pattern). These properties if realized would facilitate interpretation within and comparisons across studies. The intuitive, pedagogical, and potential comparative value of this ideal are important to protect if I can be modified to achieve its reputed or assumed qualities. A list of assumed or desirable characteristics for an autocorrelation metric is given in Table 1.

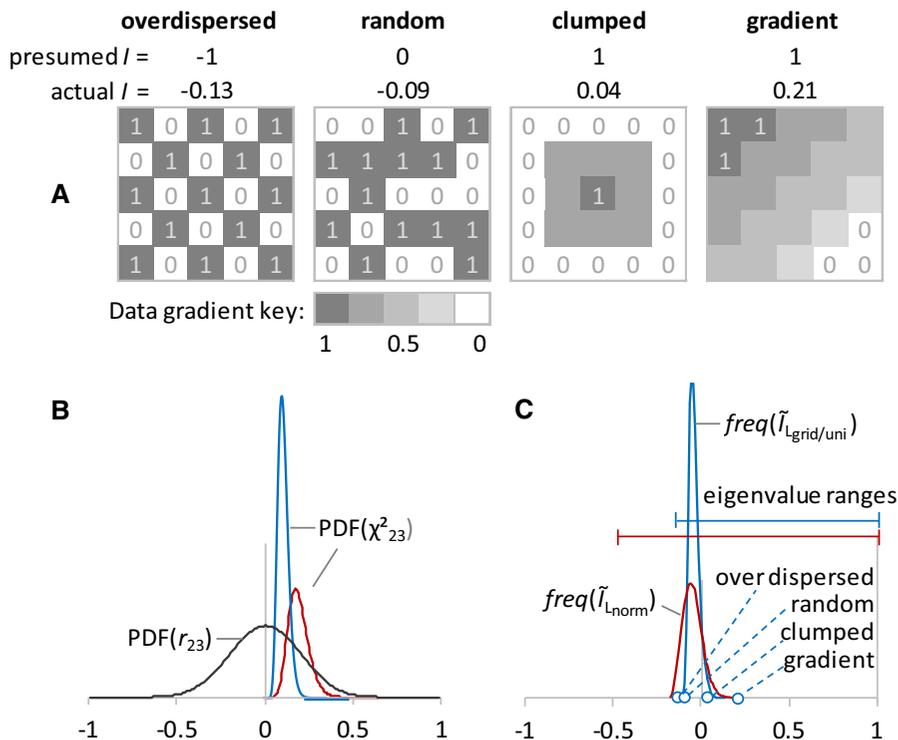


Fig. 1 Spatial patterns and assessment metric values expected and observed. **a** Spatial data arrays illustrating over-dispersed, random, or positively autocorrelated (clumped or gradient) landscape patterns. A common expectation is that I for these caricatured patterns should be at or near -1 , 0 , or 1 . Actual values for the four patterns are given above each array. **b** Probability density functions (PDFs) for Pearson r and χ^2 are given for comparison. The χ^2 PDFs were scaled to the data breadth observed in simulation using uniform grid spatial conformation (blue line) versus random normal coordinate

values (red line). **c** Distributions of \tilde{I} realized through 3×10^4 random permutations of binomial data in grid conformation as in panel A (blue line) and in a random bivariate normal location matrix (red line). Mathematically derived range limits (eigenvalue extremes of the n -sum weights matrix, $n\mathbf{W}_1$) for each distribution are given as horizontal bars. Observed ranges of \tilde{I} are given as distribution stop points on the abscissa. I values for the conformations in panel A are given in the context of their null distribution. Calculations are archived in Supplementary file ESM_1

Table 1 Qualities expected, assumed, or desirable for an autocorrelation metric

Expected or desired property of an autocorrelation metric	/	Would be evinced by:
Responsive to dispersion	+	Metric is low for overdispersion and high for clumped data conformation
Responsive to gradients	+	Metric increases with increasing gradient structure
Distinguishes clumping from gradient effects	-	Metrical values differ systematically for gradient-free clumping versus gradient structure
Strictly bounded between - 1 and 1	-	Metric is mathematically limited to the range from - 1 to 1
Value for perfect interspersion: - 1 or commensurate with expectations for regular correlation	-	Perfect interspersion yields a value of - 1 or the maximum statistic constrained by probability theory
Value for perfect gradient: 1 or commensurate with expectations for regular correlation	-	Perfect gradient patterns yield a metrical value of 1 or the maximum statistic constrained by probability theory
Value for gradient-free clustering: 1 or commensurate with expectations for regular correlation	-	Clumping without gradient structure yields a value of 1 or the maximum statistic constrained by probability theory
Central tendency for random patterns: Mean is 0	-	Mean of metrics from unpatterned landscapes or data randomizations is 0
Median is 0	-	Median of metrics from unpatterned landscapes or data randomizations is 0
Distribution of random pattern metrics is symmetric (opposite)	-	Distribution tail percentages at x and $1-x$ have the same absolute value
Metric has stand-alone interpretability (contextualized by n)	-	Metrical values consistently match existing expectations for effect strengths
Metric can be compared among studies (contextualized by n)	-	Similarity of metric values among studies having similar sample sizes and spatial pattern effect sizes
Metric itself is a standard effect size	-	Metric directly indicates effect size (e.g. % variance explained)
Metric itself is a test statistic	-	Significance of metric value can be looked up in standard tables (e.g. Rohlf and Sokal 1995)

Key: +, -, indicate the metric meets or fails the respective criterion

In reality, I is negatively biased and variably so based on sample size. Its null distribution—the distribution of values from replicated random data permutations—is right skewed and excessively narrow compared to other statistical metrics such as t or r . Its extrema are not limited to -1 and 1 and its scale is not opposite (de Jong et al. 1984). For example, I of 0.2 in one neighborhood may represent stronger or weaker autocorrelation than -0.2 in another (Anselin 1995). Theory suggests the actual bounds of I are the minimum and maximum eigenvalues of the normalized (unit sum) hollow matrix of inverse-distances (de Jong et al. 1984, Griffith 1996, Chen 2013). Yet these eigenvalues provide little indication of actual null distribution position or breadth (Fig. 1c). I varies with the geometric conformation of measurement locations (Tiefelsdorf 1998), the means of

representing sample proximity, and the quantitative nature of the measured variables (extensively reviewed in Cliff and Ord 1973). These problems render I incomparable among or even within studies for different variables.

Despite the well documented reality that I has no consistent meaning beyond singular contexts, the metric is still commonly reported, compared, and often interpreted based on its conceptual ideals. Even when reported in context with its theoretical mathematical limits, which is rarely done by empiricists, values of I are still difficult or even misleading to interpret. To provide consistent, logical, and intuitive meaning beyond individual contexts, and hence to enable conceptual and analytical syntheses such as meta-analyses (Rosenthal and Rubin 1986) or expansion to multi-domain use (Kim et al. 2015; Ritters

Table 2 Nomenclature used in this paper

Conventions	
~	Tilde in superposition indicates values derived from permutations
.	Dot in subposition indicates mean centering
^	Circumflex in subposition indicates median centering
Arrays	
v	Vector of spatially explicit data
z	Standardized (mean-centered, unit-variance) data vector
1	Column vector of n ones
L	Location matrix (e.g. latitude and longitude or grid row and column designations)
D	Matrix of distances among sample locations $\mathbf{D}=[diag(\mathbf{L}\mathbf{L}')\mathbf{1}'-2(\mathbf{L}\mathbf{L}')+1diag(\mathbf{L}\mathbf{L}')]^{1/2}$
C	Contiguity matrix. Binary matrix indicating neighboring sites on a grid
P	Proximity matrix $\mathbf{P}=1/(\mathbf{I}_n+\mathbf{D})-\mathbf{I}_n$ is a hollow ($P_{ii}=0$) inverse-elements distance matrix
W₁	Spatial weights matrix scaled to unit sum as $\mathbf{W}_1=\mathbf{P}/\sum P_{ij}$ or $\mathbf{C}/\sum C_i/n$
Scalars	
I	Moran's index of spatial autocorrelation, $I=\mathbf{z}'\mathbf{W}_1\mathbf{z}$
n	Number of cases in an empirical dataset
CP_I	Cumulative percentage of values $\leq I$ in a null distribution
\tilde{I}_{CP}	\tilde{I} at which the cumulative percentage of values is as given in subscript (e.g. \tilde{I}_{50} is the median)
x	Bounds for null distribution tail percentages (e.g. I_x and I_{1-x})
I_{3P}	I metric rectified to three points ($\tilde{I}_x, \tilde{I}_{50}, \tilde{I}_{1-x}$) in its null distribution
I_r	I metric projected to the theoretical distribution of correlations
k	Number of permutations of n data rows

2019), an improved standard for I and similarly plagued spatial pattern metrics is needed.

Given the perceptions, potential, and problems described for Moran's I , taken with the long history in spatial analysis, instead of continuing efforts to reframe its narrative, it may be better to reframe the metric itself to require fewer or no qualifications. Two methods of reframing I were developed and herein are described. The first method used a dichotomous Procrustes approach, anisometric three-point registration, to fit the null distribution based on its median and values at select tail percentages, to a target frame of $[-1, 0, 1]$, followed by projection of the original I into this conformation. The second method projected I in a continuous manner, using its cumulative percentage in the null distribution, to the theoretical distribution of regular (Pearson) correlations. We referred to these projections as 'rectification' because they scaled alternative regions of the distribution of I differently such that the collective distribution fit a designated target

frame. Differential scaling to geometric fit is eponymously termed 'Procrustes' methodology by analogy with the character in Greek mythology who variously contorted victims to fit a specific (bed) frame (Hurley and Cattell 1962). Whereas the first method used two independent scaling operations, each homogeneous on a given side of the median, the second used continuous but inhomogeneous scaling. Our goal in both cases was to register the original metric in a frame that obviates many or all the problems of the original (Table 1), so it has consistent and intuitive meaning and sustainable impact in the field of spatial analysis.

Background

Nomenclature in this article was summarized in Table 2. The original metric, I , was developed by Moran (1947, 1950). The expected mean, variance, and z and χ^2 test statistics to assess significance of

I are well known (e.g., Cliff and Ord 1973; Goodchild 1988; Rogerson 1999; Getis 2010). I is χ^2 distributed commensurate with its derivation from inverse distances (Thirey and Hickman 2015; e.g., Fig. 1c).

We use Chen's (2013) formulation of Moran's I as

$$I = \mathbf{z}'\mathbf{W}_1\mathbf{z} \quad (1)$$

where \mathbf{z} is a standardized data vector for n measurements from known locations and \mathbf{W}_1 is the spatial weights matrix scaled to unit sum. In this formulation, local indicators of spatial autocorrelation, 'LISA' (sensu Anselin 1995), are diagonal elements of $\mathbf{z}'\mathbf{W}_1$ (Chen 2013). Because local I are additive, global I can also be defined as the trace of $\mathbf{z}'\mathbf{W}_1$. Range limits of I are expected to be the signed minimum and maximum eigenvalues of $n \cdot \mathbf{W}_1$ (de Jong et al. 1984, Griffith 1996). These limits are highly sensitive to proximity definitions and their values give no practical guidance to actual distributions within them (Fig. 1c). These limits and even empirical values of I often exceed 1 (de Jong et al. 1984; Tiefelsdorf 1998). The nature of I in practice, therefore, bears little resemblance to what is widely presumed and taught regarding the metric.

\mathbf{W} can be rendered from variously composed physical distance or contiguity mappings among sample sites (well presented in Getis 2010). Distance matrices, \mathbf{D} , may be defined linearly (without exponentiation) or nonlinearly (exponentiated by >1). For example, squared distances are often used to enhance down-weighting of increasingly distant neighboring samples to reduce their influence on the magnitude of I . Contiguity matrices, \mathbf{C} , are generally binary definitions of sampling plot edge adjacencies in regular or irregular polygonal grids and may include or exclude vertex adjacency. \mathbf{C} is row-standardized to compensate differences in neighbor numbers by site as in cases where edge sites have fewer neighbors. If appropriate and desired, contiguity and distance matrices may be combined (Cliff and Ord 1969). Distance and contiguity elements are converted to inverses and can be regarded as proximities composing a matrix \mathbf{P} . Perfect proximity, the proximity of a datum's location to its own location, \mathbf{P}_{ii} , and distances beyond a selected limit if desired, are set to zero so those cases drop out of subsequent calculations. \mathbf{P} can be calculated to ensure diagonal elements are 0 as $\mathbf{P} = 1/(\mathbf{I}_n + \mathbf{D}) - \mathbf{I}_n$, where \mathbf{I}_n is an n

$\times n$ unit (identity) matrix where $\text{diag}\mathbf{I}_n = \mathbf{1}$. \mathbf{W} is a scaled version of \mathbf{P} , such as the unit-sum $\mathbf{W}_1 = \mathbf{P}/\sum \mathbf{P}_{ij}$. The type of distance and contiguity representations used to derive \mathbf{W} should be based on the logic of the process being studied (Anselin 1988). Exploratory investigation of weighting schemes based on model performance and validation may be a useful complement to the logical and functional approach.

Expectations and interpretations of I may reflect those of regular correlations (Fig. 1b) due to similar nomenclature: 'index of autocorrelation' and 'correlation coefficient' or the widely repeated notion that I was developed from Pearson's r (e.g., Getis 2010) which implies mathematical homology. Yet Pearson's r is a cross product between two similarly ($n \times 1$) dimensioned, similarly (unit-variance) scaled, and similarly (normally) distributed variables. In contrast, I is a cross product between a normally distributed $n \times 1$ data vector and and inverse elements of a variably dimensioned, χ^2 -distributed $n \times n$ matrix scaled to unit sum. For example, the distance matrix for the example in Fig. 1 has one dimension but as a proximity or weights matrix it has six (see eigenvalue counts in Supplementary file ESM_1). Thus, compared to r , I is dimensionally, volumetrically, and distributionally convoluted (Tiefelsdorf 1998, in part).

Expected I under a null hypothesis of no spatial pattern, $E(I)$, is $-1/(n-1)$ with expected variance and a z test statistic as given in Cliff and Ord (1973) or a χ^2 statistic as given in Rogerson (1999). Significance of I metrics is better assessed by permutation analysis than by inference from a test statistic because the distribution of I can be unpredictably skewed based on spatial arrangement of the sample sites, varied distributional properties of measured variables, and other characteristics that violate parametric assumptions (Upton and Fingleton 1985; Goodchild 1988; Tiefelsdorf and Boots 1997; Li et al. 2007). \bar{P} values from Monte Carlo simulations are determined as the proportion of permutation results more extreme than the observed I as appropriate to the null hypothesis. For example, the probability of unbiased processes underlying an I equal or smaller than that observed is the cumulative percentage of data up to I . For extremely patterned data, I will often exceed the maximum or minimum of values obtained in even large numbers of permutations. As probability

estimates are limited by permutation number, the logical limit must be imposed that it be not less than $1/(k+1)$ for k permutations (Griffith 1987) or half that for two-tailed tests. Permutations also produce an explicit null distribution that can be scaled, as demonstrated herein, to fit a specific frame or desired density function.

Methods

3-point registration

This method to rectify I used dichotomous scaling of null distributions of \tilde{I} obtained by Monte Carlo simulation using replicated random data permutations. Based on the ideology described in the Sect. “Introduction” and Table 1, null distributions were median centered, and each half of the distributions was separately scaled, homogeneously on a given side, so that the bounds of chosen distribution tail percentages occurred at -1 and 1 . Use of the median to center null distributions guaranteed that half of the values obtained by randomly permuting the data were positive and half were negative. This criterion for centrality did not rely on specific expectations for the geometry of central tendency. Obvious choices for extremes to serve as the ± 1 limits for the null distribution were the minimum and maximum values, \tilde{I}_0 and \tilde{I}_{100} . If all possible $n!$ permutations or an infinite number of random permutations were conducted, this choice of distributional limits would ensure that no conformation of data could yield $|I_{3P}| > 1$, making an absolutely bounded distribution frame. In practice, we attempted to approach the absolute frame with stringent definitions of tail percentages and high permutation effort. In theory, this could also be achieved by using the $n\mathbf{W}_1$ eigenvalue limits. But in practice these extrema are far and variably removed from the bulk of the realizable distribution, so that frame would result in greatly contorted transformed distributions. Because the repeatability and proximity to a reasonable frame could have depended on permutation number, choice of tail percentages for the reference frame, and possibly their interaction, as well as unique system-specific properties, sensitivity analyses as described below Sect. (“Survey of published datasets”) were performed. Our methods were intended to identify

combinations of frame definitions and number of permutations that balanced accuracy and effort to provide an acceptably repeatable standard for diverse datasets.

We calculated I following Chen (2013; Eq. 1). Then, either the standardized data vector \mathbf{z} or the location matrix \mathbf{L} was permuted (shuffled case wise) k times and \tilde{I} was recalculated each time to yield a ‘null’ distribution of frequencies. Null distributions were therefore those for values of \tilde{I} calculated from datasets with no covariation, on average, between \mathbf{z} and \mathbf{D} , which is the matrix of pairwise distances calculated from \mathbf{L} . \tilde{I} values at designated cumulative percentages of the null distribution were then used to enact median centering and the separate scaling of the positive and negative sides of the centered distribution. I values obtained by permutation were denoted with a tilde in super-position, \tilde{I} , median-centered values with a circumflex in sub-position, $\tilde{\tilde{I}}$, and fully rectified (centered and scaled) values were denoted for the 3-point registration method with subscript, as in \tilde{I}_{3P} . These conventions apply hereafter for any parameter obtained through permutation. The cumulative percentage of values $\leq \tilde{I}$ was designated with numerical subscript. For example, the null distribution minimum was denoted \tilde{I}_0 , the maximum as \tilde{I}_{100} , and the median as \tilde{I}_{50} . Tail percentages defining frame boundaries were mirrored as in \tilde{I}_x and \tilde{I}_{1-x} (Table 2).

After median-centering, each half of the null distributions was scaled by the distance from the median to each respective frame boundary:

$$\tilde{I}_{3P}(\tilde{I}) = \begin{cases} \frac{\tilde{I}-\tilde{I}_{50}}{\tilde{I}_{50}-\tilde{I}_x} & \text{if } \tilde{I} < \tilde{I}_{50} \\ 0 & \text{if } \tilde{I} = \tilde{I}_{50} \\ \frac{\tilde{I}-\tilde{I}_{50}}{\tilde{I}_{1-x}-\tilde{I}_{50}} & \text{if } \tilde{I} > \tilde{I}_{50} \end{cases} \quad (2)$$

Scaling was applied similarly for the observed I . This scaling method set all values on a given side of the centered null distribution to their proportional position between the median and the respective tail boundary. For example, I of $+0.2$ in a null distribution having median 0 and \tilde{I}_{1-x} of 0.4 , yields an I_{3P} of $+0.5$ ($=^{0.2/0.4}$). In this manner, any original null distribution frame $[\tilde{I}_x, \tilde{I}_{50}, \tilde{I}_{1-x}]$ was anisometrically scaled about the median to $[-1, 0, 1]$.

Values of $|I_{3P}| > 1$ were expected to occur at frequency $2xn$ (frequency xn on each side of the null distribution). As well, empirical datasets with strong

autocorrelation would often yield $|I_{3P}| \gg 1$. Values of $|I_{3P}|$ and $|\tilde{I}_{3P}|$ in excess of 1 were set based on their signs to ± 1 :

$$|I_{3P}| > 1 = \begin{cases} -1 & \text{if } \tilde{I}_{3P} < -1 \\ 1 & \text{if } \tilde{I}_{3P} > 1. \end{cases} \tag{3}$$

P-values for hypothesis tests and analysis were obtained directly from cumulative proportions at *I* or I_{3P} in their respective null distributions. For example, a directional hypothesis of overdispersion is supported when the cumulative proportion of I_{3P} in its null distribution is less than or equal to α , where α is the established minimum probability of making a statistical type-I inferential error. A hypothesis of any systematic spatial structure is supported if the cumulative proportion of negative $I \leq \alpha/2$ or that for positive $I \geq 1-\alpha/2$. Calculation of \tilde{P} values for null hypothesis tests by frequencies are:

$$\tilde{P}(H_0) = \begin{cases} A = \frac{\sum_{-1}^I \text{freq.}(I)}{k} & H_0 : \text{no over dispersion} \\ B = \frac{\sum_I^1 \text{freq.}(I)}{k} & H_0 : \text{no clumping or gradients} \\ \frac{2 \cdot \min(A,B)}{k} & H_0 : \text{no autocorrelation.} \end{cases} \tag{4}$$

Parametric *P* values for some specific demonstrations were also calculated using *z* statistics per Cliff and Ord (1973).

Fit to Pearson's *r*

Our second method of rectification also made use of permutations as described above. However, since our goal was to register the null distribution to a theoretically known distribution, there was no need to define target frame boundaries. *I* values were fit to *r* using inverse distribution functions based on their cumulative proportions in null distributions. The probability distribution of correlation coefficients from random bivariate normal or uniform data is

$$\text{prob.}(r) = \frac{(1 - r^2)^{\frac{n-4}{2}}}{\sqrt{\pi}} \cdot \frac{\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n-2}{2})} \tag{5}$$

(Hotelling 1953). This function can be integrated from -1 to that *r* having the same cumulative proportion as *I* in its null distribution, which *r* is then deemed I_r .

$$I_r := r \left| \int_{-1}^r \text{prob.}(r) dr = CP_1 \tag{6}$$

In practice it was simpler to invoke the inverse *t* distribution as it is widely available in software packages such as R, MATLAB, Stata, and Excel. The expected *t*-distribution of Pearson correlations for bivariate normal or uniform random variables is given by

$$t(r) = r \sqrt{(n-2)/(1-r^2)} \tag{7}$$

(Rahman 1968), where $n-2$ is the degrees of freedom. To map *I* values to this distribution, their cumulative proportion in the null distribution was used to calculate a t_{n-2} statistic using the inverse *t* distribution as

$$t(CP_1) = \begin{cases} t.\text{inv}\left(\frac{1}{k+1}/2, n-2\right) & CP_1 = 0 \\ t.\text{inv}(I_r, n-2) & 0 < CP_1 < 1 \\ t.\text{inv}\left((1 - \frac{1}{k+1})/2, n-2\right) & CP_1 = 1 \end{cases} \tag{8}$$

then *t* was converted to *r* as

$$r(t) = t / \sqrt{n-2+t^2} := I_r \tag{9}$$

with I_r given to equal *r*. Where *t* was undefined at the distribution limits, $CP=0$ and 1 , I_r was set to the respective limits of *r* given the constraint of $P \geq \pm 1/(k+1)$ (Griffith 1987). Thus Eqs. 5, 8 and 9 functioned to prevent I_r from being set to a value with greater significance than the limit imposed by permutation number.

Mapping from cumulative proportions to *r* was used both to rectify empirical *I* values but also to convert permutation results to I_r . I_r was graphically illustrated by denoting its position within either the theoretical distribution function (Eqs. 5, 9) or the empirical histogram of all \tilde{I} obtained through permutation. The latter was used for intuitive value and contrast with the original null distribution histogram.

Both rectification methods are available as options in the R package Irescale (Fuentes et al. 2020).

Focal case: autocorrelation of major Chinese city population sizes

Both rectification procedures were applied to data from Chen (2013; Fig. 2a) on population sizes of and rail distances among 29 major Chinese cities. This dataset is referred to hereafter as *Cities*. I was calculated following Chen (2013; Eq. 1), whose result we wished to replicate. The rail distance matrix, \mathbf{D} , was non-positive definite due to rail paths being variously indirect (Dokmanic et al. 2015). Therefore, \mathbf{D} could not have been used to derive an \mathbf{L} matrix to be permuted and still replicate Chen's results. Therefore, we permuted only the standardized data vector \mathbf{z} . The *Cities* example was examined using the original data, a natural logarithm (\log_e) transformed version, and a designed dataset composed to have extremely high autocorrelation (Fig. 2b). The latter two data vectors were included to assess the responsiveness of I and its rectified metrics to variable transformation and inform on the positive metric limit. Each of the three data vectors were used to calculate I for each of 10^4 permutations. From these results we calculated \tilde{I} (i.e., $\bar{I} - \bar{I}_{50}$), I_{3P} using $x=0$ and 0.1 tail percentages as frame boundaries, and I_r . As a control measure to test for permutation bias, a Pearson correlation was calculated between the original data vector and each permuted version. The null distribution of these

correlations was examined to ensure $r_{27}^* \approx 0$. Calculations for this case analysis are provided in an Excel file (Supplementary file ESM_2). All calculations cited, except the random r , can be performed with the R package Irescale (Fuentes et al. 2020).

Sensitivity analysis

Sensitivity and repeatability of null distribution parameters were explored for both rectification methods using replicated sets of permutations. All datasets were used for this purpose. We focused on the *Cities* dataset for presentation of detailed dynamics ("Sensitivity analysis" sect.). Summary statistics and exceptional behaviors noted for the broader survey described in "Survey of published datasets" sect. were also assessed and reported in "Survey of published datasets" sect. For the *Cities* analysis, we used the rail \mathbf{D} matrix where possible to align with the analysis by Chen (2013) but could not do so for spatial regressions, in which case we used geographic \mathbf{D} . The geographic and rail distance matrices were highly correlated (matrix $r_{404}=0.93$). Simulations were carried out for 10 replicates of each combination of six permutation numbers ($k=10^1, 10^2, \dots, 10^6$). From simulation results, average and standard error, which in this paper is considered to be the obverse of repeatability, were calculated for all I , \tilde{I} , rectified metrics, and 3-point registration frame boundaries (I_x ,

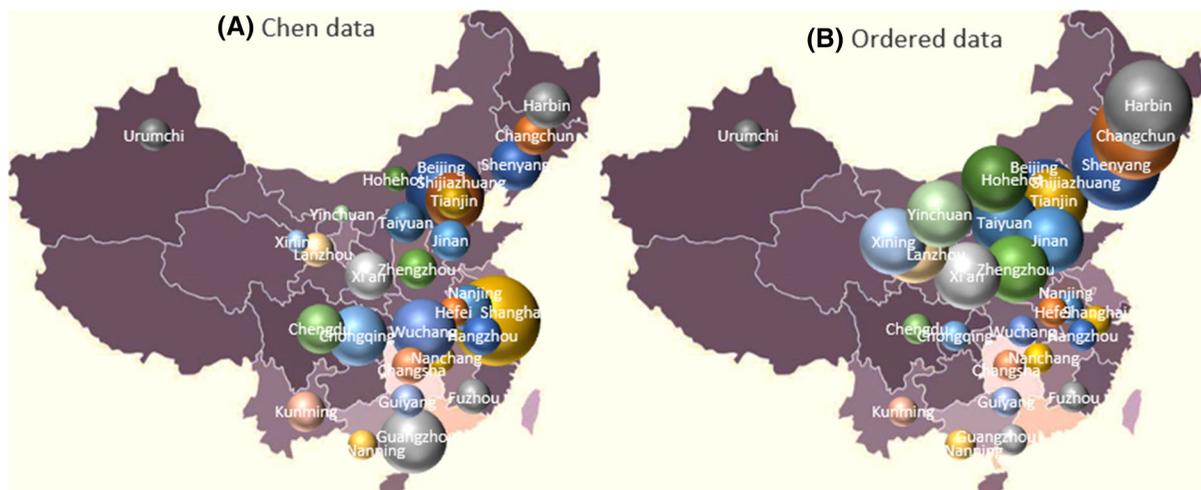


Fig. 2 Focal data to illustrate rectification methods. **a** Original data on populations of major Chinese cities (from Chen 2013). **b** Designed data created to have high spatial autocorrelation

due to data gradients and clumping. Marker diameter corresponds to population size

I_{1-x}) defined by tail percentiles $x=0, 10^{-3}, 10^{-2}, 10^{-1}$, and 10^0 .

To resolve how I tracked with other autocorrelation metrics, we calculated the matrix correlation between \mathbf{D} and the matrix of Euclidian distances among elements of \mathbf{z} . This ‘matrix’ correlation is a Pearson correlation of the corresponding elements taken in pairs from the two distance matrices (Mantel 1967). We also calculated a multiple correlation from a spatial multiple regression of the form

$$\hat{\mathbf{z}} = b_0 + b_1 \cdot \text{latitude} + b_2 \cdot \text{longitude} + b_3 \cdot \text{latitude} \cdot \text{longitude} \quad (10)$$

The square of multiple r is the proportion of variance in \mathbf{z} explained by gradient effects (Sokal and Rohlf 1995). This measure of gradient effect size was used with I_r^2 , a dimensionally comparable metric of non-dispersion, to create complementary indices that differentially capture aspects of spatial pattern due to clumping versus gradient effects as

$$I_C^2 = I_r^2 / (I_r^2 + \text{multiple } r^2) = 1 - I_G^2 \quad (11)$$

To enable the last three calculations (Eqs. 9, 10 and 11) for *Cities*, we used an \mathbf{L} matrix composed of GPS coordinates obtained from Google Earth. We also applied the gradient and clumping metrics to the ‘number fledged’ data vector of Marrot et al. (2015), three designed versions of that vector, and an array of patterns from the 25-plot grid illustrated in Fig. 1. Designed data vectors from *Cities*, Marrot et al. (2015), and targeted simulations of gradient, random, clumped, and ‘clumpy-gradient’ data were used to assess sensitivity of I_C^2 and I_G^2 to cases with known magnitude patterns.

Sensitivity results were compared within and between studies with emphasis on repeatability and magnitude of I_r metrics and correlations between I , I_{3P} , I_r , matrix r , and multiple r from regression. The sensitivity analysis for *Cities* as implemented in Excel is available in Supplemental file ESM_3. Irescale (v. 2.3.0; Fuentes et al. 2020) was used to confirm the *Cities* result and calculate sensitivity analyses for the case studies described in the following section.

Survey of published datasets

To achieve a broad view of how rectified I compared to traditional I and related metrics, we identified and

examined four published datasets in addition to *Cities*. The additional data were identified by searching the phrase ‘Moran’s I ’ in Google Scholar (in mid-2019) and examining recent publications returned to find a subset of 12 in which the authors openly archived data. The 12 datasets were further examined and four were selected by qualitative review for having diverse variable types within studies and diverse sample size and spatial structures among studies. The four datasets identified were Cozzarolo et al. (2018), Goldsmith et al. (2019), L’Herpinier et al. (2019), and Marrot et al. (2015). These datasets collectively included 24 variables as accounted in detail in the ‘‘Results’’ sect. For each dataset, we conducted rectification and companion procedures, including the sensitivity analysis described above (‘‘Sensitivity analysis’’ sect.). Results for sensitivity dynamics and correlations of spatial pattern metrics within and among studies were reported.

Summation regarding metric ideals

Results of the varied methods described above were scrutinized for congruence or dissonance with the characteristics given in Table 1 as regards I and comparator metrics. These inferences were drawn primarily from mathematical definitions and null distribution shapes as presented below and confirmed with broader explorations archived in the supplementary files.

Results

Three-point registration

General results are recounted here and specific case study detail and procedural assessments are reported in sections that follow. Null distributions rectified by 3-point registration resulted in the frame $[\tilde{I}_x, \tilde{I}_{50}, \tilde{I}_{1-x}] = [-1, 0, 1]$. The mean of 3P-rectified distributions was generally mismatched with the median, most often in a direction opposite that for the unrectified null distribution. Empirical I from strongly patterned datasets when projected into their rectified null distributions in several cases exceeded 1 and were subjected to capping (Eq. 3). Selecting smaller distribution tails for target frame references reduced

I_{3P} and the frequency of times limits were necessary to invoke. Even when using maximum and minimum values of \tilde{I} as the target frame, the caps were still necessary because some empirical patterns represented a degree of patterning not discovered with limited permutation. For I observed to be significant by permutation, rectification by 3-point registration generally resulted in $|I_{3P}| \gg |I|$ and \tilde{I} .

Rectification by 3-point registration greatly increased similarity of null distribution tail shapes relative to the unrectified distribution (Fig. 3a, b). However, separate scaling of the upper and lower distribution halves resulted in strong leverage in the skewed side of distributions, such that the long tail contracted toward the bound at the median. Tail contraction resulted in pushing up the area (distribution density) near the median in that half of the distribution. In opposite manner, the short-tailed half of the distribution drew out the tail and density near the median decreased, producing a slumping effect. The result of the opposing shifts in density created a shape discontinuity at the median (Fig. 3b). Despite creating an odd distribution shape around the median, skewness was most often considerably reduced and sometimes was reversed (changed in sign). \tilde{P} values for I , I_{3P} , and I_r were equivalent when calculated from frequencies. P values calculated from z tests per Cliff and Ord (1973) differed among the metrics and frequently fell below the practical limit of \tilde{P} from frequencies.

Fit to distribution of Pearson's r

The distribution-fitting method produced null distributions fully contained in the interval $[-1, 1]$ without applying caps (Fig. 3c), although distributional extrema $[I_0, I_{100}]$ where t is undefined had to be limited to that having $\tilde{P}=1/(k+1)$ per Eq. 8. The mean and median were coincident at 0. By definition, the rectified null distribution, being fit to a t -distribution, was neither appreciably skewed nor asymmetric about the median, although it was subject to moderate platykurtism for sample sizes below 10. Left and right tail percentiles were opposite; that is, tail areas were of the same magnitude at $\pm \tilde{I}_r$. \tilde{P} -values were functionally equivalent when calculated from simulation frequencies (Eq. 4) or by t statistics derived from I_r (Eq. 7).

Focal case: autocorrelation of major Chinese city population sizes

The *Cities* population and rail distance data yielded I of -0.031 in concurrence with Chen (2013; $z_{27}=0.11$, one tailed $P=0.46$; $\tilde{P}=0.42$). Null distributions for the original, log and designed data from simulations with 10^4 permutations yielded mean I within a few ten-thousandths of the expected mean. Medians were approximately 9% more negatively biased than the means for all three data vectors and all were right skewed (Table 3). Null distributions from untransformed data were most skewed and log-transformed data were least skewed. Simulation results in this analysis never produced median-centered I in excess

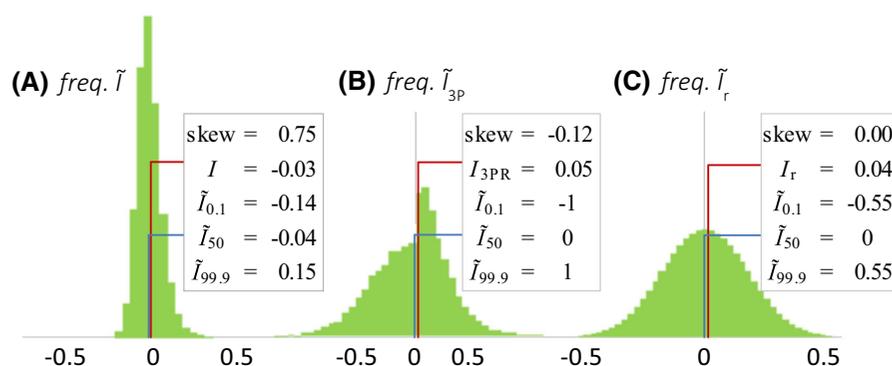


Fig. 3 Null distribution of I from 10^4 permutations of the *Cities* data. **a** Moran's \tilde{I} . **b** \tilde{I} rectified using 3-point registration. **c** \tilde{I} fit to the distribution of Pearson's r . The position of unrectified and rectified I on the abscissa are indicated in red

with the respective medians in blue. Distribution tail boundaries for $\alpha=0.1$ and skewness is given in each panel. Calculations were archived in Supplementary file ESM_4

of 0.27, suggesting a limit. However, based on I for the designed dataset ($=0.28$) and sensitivity analyses, the limit appeared to be ≈ 0.33 (“Survey of published datasets” sect.). This limit from simulation contradicts the theoretical limit of the maximum eigenvalue of $n\mathbf{W}_1$ which was 1.09. \tilde{P} was similar to that from parametric z -tests except where $P < 1/(k+1)$. The limit imposed by Eq. 4 prevents assigning I_r any higher value than that which can be resolved as significant given any fixed k . (Table 3). These calculations are archived in Supplementary file ESM_2.

I statistics differed markedly in this analysis, most strongly among data vectors but also among modes of rectification (Table 3). For the original data vector and its log version, \hat{I} was 0.01 and 0.06. The latter value was marginally significant by one-tailed test

against a null hypothesis of no positive autocorrelation ($\tilde{P}=0.09$). The rectified metrics produced similarly low values for the untransformed data (0.03–0.04) but rose for the log data to values that would be moderate as regular correlations (0.26–0.39). Finally, the designed positive autocorrelative dataset yielded $I=0.28$, though the values for the rectified metric were much greater ($I_r=0.64$, both $I_{3P}=1$). Both I_{3P} in this case were constrained by the cap set at 1 (Eq. 3). I_r was limited due to permutation replication because t_{n-2} , part of the conversion procedure from I to I_r , was limited to that having probability $1/(k+1)$ —in this case $t_{27}=4.3$, which translates to $r_{27}=I_r=0.64$.

Overall pattern strength was much greater for the designed dataset at $I_r = 0.64$ compared to 0.26 for the log data. However, the autocorrelation partitioning

Table 3 SpaSpatial autocorrelation metrics and right tailed P -values for the *Cities* dataset

Parameter	Metrics per data series		
	n	$\ln(n)$	Designed
Skewness	0.72	0.33	0.49
$\lambda_{\min} n\mathbf{W}_1$		-0.35	
$\lambda_{\max} n\mathbf{W}_1$		1.09	
Moran's I	-0.031	0.024	0.28
$I - I_{50}$	0.008	0.063	0.31
$I_{3P,0.1}$	0.043	0.388	1
$I_{3P,0}$	0.030	0.316	1
I_r	0.041	0.257	0.64
P	0.42	0.088	10^{-4}
Multiple r	0.34	0.38	0.92
P	0.36	0.26	$<10^{-12}$
Matrix r	-0.085	-0.021	0.43
P	0.74	0.54	<0.01
I_c^2	-	0.31	0.32
I_G^2	-	0.69	0.68

indices I_G^2 and I_C^2 demonstrated similar patterns between the log data and the designed dataset. For both datasets, the cluster metrics I_C^2 were approximately 0.3; thus the gradient metrics I_G^2 were approximately 0.7. The joint presence of clumping per se and gradient effects for the designed data was visually apparent from Fig. 2b as datum magnitudes were clumped at one scale but the clumps were arrayed as a north–south gradient. The pattern partitioning indices were not calculated for untransformed data due to lack of evidence for any spatial pattern to be characterized ($\bar{P} > 0.4$ for both I and multiple r , Table 3).

Figure 3a illustrates a focal null distribution of \tilde{I} from 10^4 random permutations of the *Cities* data. The distribution was negatively biased to the expected level and in this case was particularly right skewed. Over 83% of \tilde{I} values from permutation were negative and the skew was highly significant based on D'agostino et al.'s (1990) test for excess skewness ($z_{27} = 27.6$, $P < 10^{-20}$). The range of median-centered values extended 1.3-fold as far to the right as they did on the left, but this range was still far less than was suggested given the 3.2-fold difference in theoretical limits imposed by eigenvalues of $n\mathbf{W}_1$. The null distribution mean of Pearson correlations between the original data vector and its permuted version was near 0 ($r_{27} = 0.002$) with no noteworthy skewness ($= 0.01$). Matrix correlations of geographical distance and Euclidean distance among permuted measurements were not notably biased for any *Cities* datasets (matrix $r_{\bar{r}} \leq 0.0021$), but these distributions were strongly skewed ($= 0.97, 0.46$, and 0.80 for actual, log transformed, and designed datasets). Null distributions of multiple r from the spatial regressions averaged 0.30 – 0.31 and were moderately skewed ($= 0.55, 0.27, 0.25$).

Figure 4 gives the null distributions of I , its transforms, matrix r , and Pearson's r for the untransformed *Cities* data vector. The only null distribution that was opposite (symmetric about the mean/median) was that for I_r , which was fitted to and therefore represented the theoretical function for Pearson's r . Null distributions of I_{3P} at some frame definitions were found to have similar properties to those of Pearson's r but presented the characteristically strong shape discontinuity about the median (Fig. 4b). Plotting null distributions in log space enhanced the apperency of null distribution tail

shapes (Fig. 4c). Extremal probability decays for I_{3P} were similar in nature to, but broader than the theoretical expectation for Pearson's r .

Sensitivity analysis

Permutation with different numbers of iterations demonstrated the repeatability of rectified metrics and in the case of I_{3P} in particular the ultimate values obtained. Since the natures of the two rectification methods were different, they were considered separately.

3-point registration

Regardless of the tail percentiles selected as a frame of reference to scale I , increased permutation replication increased the range of extreme values discovered. This was especially so in the skewed end of the distribution. The greater extrema revealed through higher permutation replication had the effect of widening reference frames (Fig. 5a). Wider frames entailed frame boundaries more distant from the median. Since I was scaled by these distances (Eq. 2), greater permutation replication resulted in smaller $|I_{3P}|$. This relationship between more permutation and reduced $|I_{3P}|$ was asymptotic at accessible levels of permutation (Fig. 5b). For example, the maximum

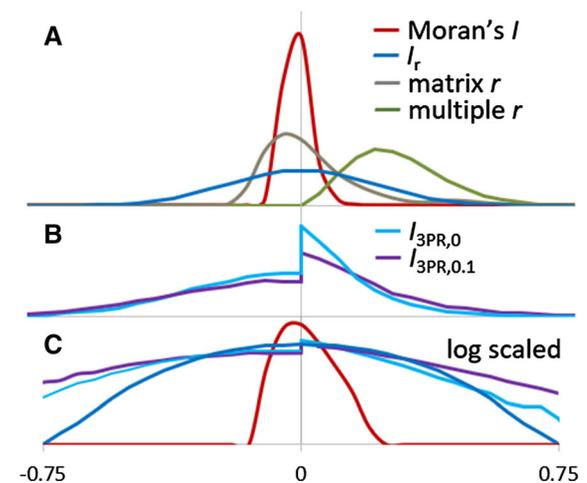


Fig. 4 Null distribution shapes for 3×10^4 permutations of the untransformed *Cities* data. **a** Moran's I , I_r , and matrix and multiple r . **b** I_{3P} for two frame definitions. **c** I metric null distributions in log space to emphasize tail conformations relevant for hypothesis testing (Supplementary file ESM_5)

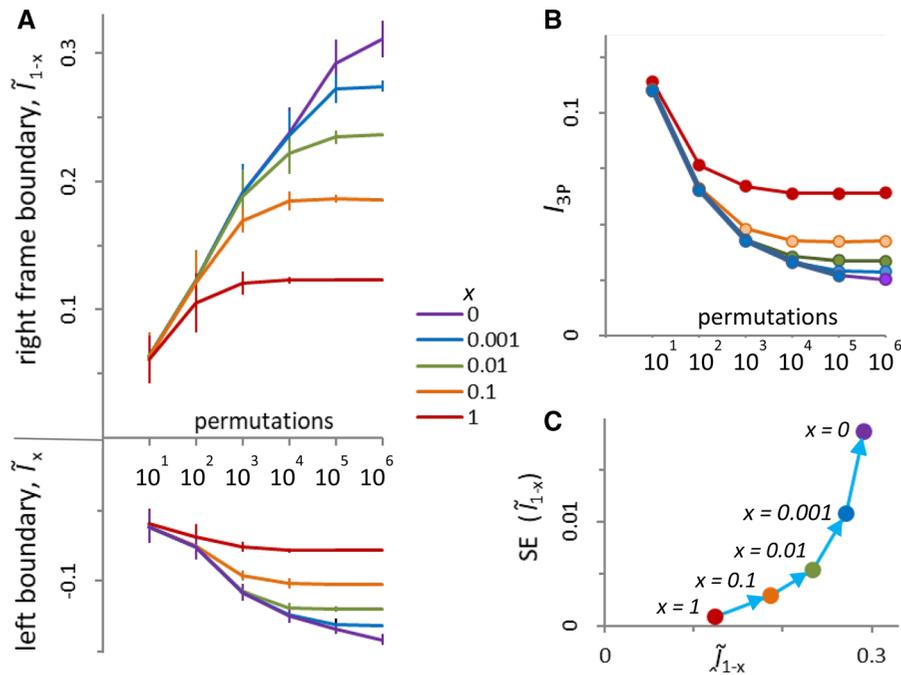


Fig. 5 Sensitivity analysis for 3-point rectification of \tilde{I} from the *Cities* dataset. Values are from 10 replicate simulations at various permutation numbers and target frame percentiles. **a** Frame boundaries expanded with increased permutation replication and reduced target frame percentiles. Repeatability ($\sim SE^{-1}$) of frame boundaries increased with permutation replication and more stringent tail percentiles. **b** Dependency of I_{3P} on permutation replication and frame criteria. I_{3P}

frame width using 1% tails as frame boundaries was evident and highly repeatable with a total of 10^4 permutations (10 replicates of 10^3 permutations). In general, as asymptotes became evident with sufficient replication, error became trivial. Thus, asymptotic values being so identified by sensitivity analysis demonstrated the levels of sufficient replication needed to attain I_{3P} nearly invariant to increased permutation effort.

Despite the repeatability and stable limit for frame boundaries, and hence stability of I_{3P} for a given frame definition, alternative frame definitions had different asymptotes and different errors. For example, while the right frame boundary using 1% tails to define the frame resolved an average frame bound of 0.12 in 10^4 permutations (red symbol, Fig. 5c), the 0.01% tail at that replication averaged twice that (0.24; green symbol, Fig. 5c), with tenfold increased error. Error was mitigable with more permutation, yet higher permutation widened the reference frame

decreased with increasing permutation replication to asymptotes unique for each frame designation. Asymptotic values for alternative frames were reached at different permutation replications. **c** Summary of tradeoffs among illustrated parameters. Increasingly diminished tail area (arrow train from red to purple symbols) led to a stable limit for the frame definition (0.29 on the abscissa) but also increased error (ordinate; Supplementary file ESM_3)

further and reduced I_{3P} . Therefore, a basic finding of this sensitivity analysis involves complex tradeoffs mitigable only by joint optimization of frame definitions and replication, the dynamics and appreciable balance of which could only be discerned through explicit analysis.

Fit to Pearson's r

This method of rectification was highly repeatable even at modest permutation replication (Fig. 6). For example, I_r calculated from the *Cities* data using 10 replicate null distributions of 10^4 – 10^6 permutations, all reached the asymptotic value (0.04) with no noteworthy error. Thus, the original I of -0.03 would be solidly $+0.04$ on the expected scale for regular correlation. Even with lower replication by an order of magnitude, the asymptotic value was met but with concern-worthy (11.7% of the mean) error. The two lowest replication levels resulted in

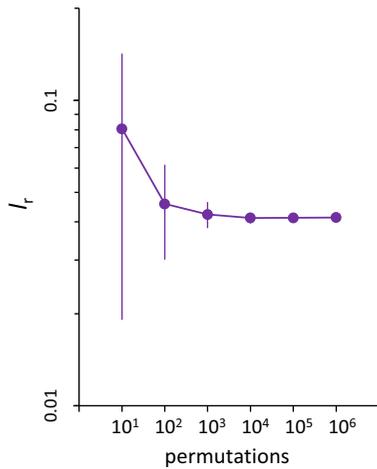


Fig. 6 I_r average and standard error of 10 replicate simulations plotted in log space (ordinate) as a function of 6 magnitudes of permutation (abscissa) of the *Cities* data

exponentially greater error and inflation of I_r . This sensitivity analysis and that for the 3-point registration method are archived in Supplementary file ESM_3.

Table 4 gives correlations of the three spatial pattern metrics (\tilde{I}_r , matrix \tilde{r} , and multiple \tilde{r}) over 2×10^4 permutations both with each other (for a given data vector) and across data vectors (for a given metric). These correlations reflect the co-sensitive nature of the metrics to spatial data patterns in the Monte Carlo simulation. In general, there were strong relationships for values of a given metric calculated for the untransformed and log data vectors ($0.79 \leq \tilde{r} \leq 0.85$; blue shading in Table 4). The strongest relationships among metrics for a given data vector were between matrix \tilde{r} and multiple \tilde{r} ($0.54 \leq \tilde{r} < 0.66$; yellow shading in Table 3). Of the two \tilde{r} metrics, multiple \tilde{r} was more closely correlated with \tilde{I} ($0.42 \leq \tilde{r} < 0.43$; green shading in Table 3) than matrix r ($0.24 \leq \tilde{r} \leq 0.29$; orange shading in Table 4). There was no reason to expect correlation between the designed dataset and the empirical data vectors so metric correlations for these are not given in the table. These calculations are archived in Supplementary file ESM_2.

Table 4 Correlations among spatial pattern metrics from permutations of *Cities* data vectors

		I	Matrix r			Multiple r		
		$\tilde{z}_{\ln(n)}$	\tilde{z}_n	$\tilde{z}_{\ln(n)}$	$\tilde{z}_{\text{designed}}$	\tilde{z}_n	$\tilde{z}_{\ln(n)}$	$\tilde{z}_{\text{designed}}$
I	\tilde{z}_n	0.851	0.259	–	–	0.428	0.372	–
	$\tilde{z}_{\ln(n)}$	–	–	0.235	–	0.332	0.429	–
	$\tilde{z}_{\text{designed}}$	–	–	–	0.2931	–	–	0.4157
Matrix r	\tilde{z}_n	–	–	0.789	–	0.659	–	–
	$\tilde{z}_{\ln(n)}$	–	–	–	–	–	0.539	–
	$\tilde{z}_{\text{designed}}$	–	–	–	–	–	–	0.5635
Multiple r	\tilde{z}_n	–	–	–	–	0.796	–	–

Table 5 Spatial pattern metrics from five published studies

Source	Variable	<i>n</i>	<i>I</i>	$I_{3P}^{a,b}$	I_r^a	\tilde{p}	Skew
Chen	Pop	29	− 0.031	0.030	0.041	0.83	0.73
	ln(pop)	29	0.024	0.316	0.257	0.18	0.34
	Designed	29	0.275	<i>l</i>	0.771	10^{−6}	0.56
Cozzarolo et al. (2018) ^c	Sex	617	0.032	0.23	0.053	0.19	0.79
	wt	575	0.023	0.05	0.053	0.20	3.53
	tars.len	609	0.066	0.52	0.091	0.03	0.11
	hyb.ind	576	0.510	<i>l</i>	0.202	10^{−6}	0.16
	inf.plas	617	0.182	<i>l</i>	0.195	10^{−6}	0.11
	inf.leuc	617	0.084	0.78	0.116	0.004	0.10
	mixed.p	273	− 0.046	− 0.19	− 0.033	0.58	1.48
	mixed.l	311	− 0.004	0.04	0.011	0.85	0.43
	ct.host.a	151	0.013	0.02	0.045	0.59	1.88
	ct.host.b	146	0.037	0.03	0.078	0.35	1.22
	ct.para.a	144	0.015	0.04	0.040	0.64	0.96
ct.para.b	140	0.054	0.08	0.080	0.35	0.92	
Goldsmith et al. (2019)	br.wtr.ox	81	− 0.043	− 0.36	− 0.133	0.24	0.35
	br.wtr.h	81	0.019	0.5597	0.270	0.0149	0.34
	tree.dbh	74	− 0.003	0.28	0.168	0.15	0.41
	leaf.watr.ox	81	− 0.016	0.13	0.076	0.50	0.38
	leaf.watr.h	81	0.026	0.60	0.300	0.007	0.39
L'Herpinierie et al. (2019)	Parasite	270	1.17	<i>l</i>	0.293	10^{−6}	0.20
	tmax	270	0.452	0.70	0.216	<10^{−3}	0.13
	LAI	270	0.656	<i>l</i>	0.293	10^{−6}	0.04
	calc	270	0.399	0.61	0.195	10^{−3}	0.26
	Humid	270	0.834	<i>l</i>	0.293	10^{−6}	0.05
	Vision.pcl	270	− 0.001	0.01	0.008	0.90	0.05
Marrot et al. (2015)	lay.date	229	− 0.004	0.1099	0.053	0.42	0.53
	clutch.size	229	0.009	0.40	0.150	0.02	0.52
	incub.duration	229	− 0.008	0.03	0.017	0.80	0.58
	num.fledged	229	0.027	0.51	0.231	<10^{−3}	0.55
	perf.grad	229	0.315	<i>l</i>	0.317	10^{−6}	1.43
	perf.clump	229	0.298	<i>l</i>	0.317	10^{−6}	1.29

^a $k=10^6$; ^b $x=0.01$; ^c Data subset (2005–2007 only)

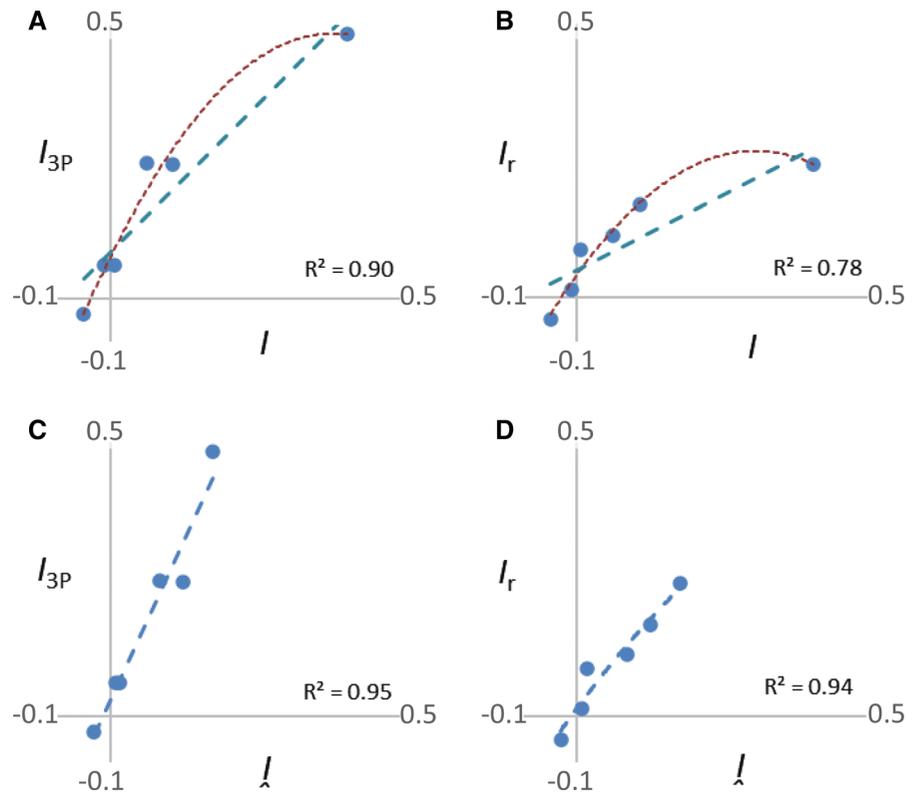
Italics indicates values limited by capping (for I_{3P} per Eq. 3) or logical maximum based on k (for I_r). Bold— P values indicated significant autocorrelation using $\alpha=0.05$

Survey of published datasets

I varied considerably among variables and studies (Table 5). Few (4 of 29) empirical I in the survey were below the null distribution median and none were less than -0.03 . Of the 25 I values above the median, 13 were significant. The range of median-centered values was 0.034–1.17. Rectified values

ranged from -0.36 to 1. The positive cap (Eq. 3) was invoked in five cases exclusive of designed data vectors. Summary metrics for the 29 variables from five studies plus three designed and one log transformed data vector were collected in Table 5. Both rectified metrics were highly correlated with original I scores. In general, I_r was a more conservative metric than I_{3P} ; larger magnitude values occurred especially

Fig. 7 Relationship between I values. **a** I_{3P} calculated with null distribution extrema and **b** I_r plotted relative to raw I . **c** $I_{3P,0}$ and **d** I_r plotted with median-centered I . Shown are the data for six variables from the 2006 data series of Cozzarolo et al. (2018)



(A) gradient

(B) clumped

(C) clumpy gradient

(D) autocorrelation metrics

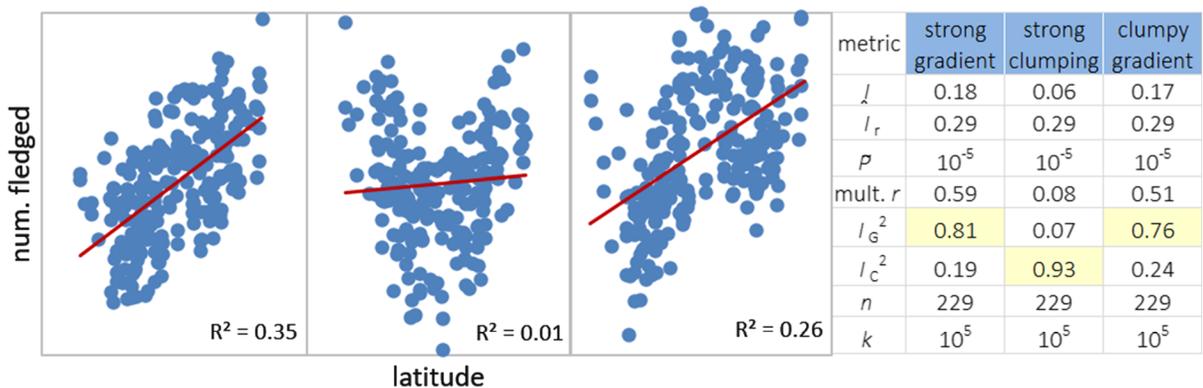


Fig. 8 **a–c** Designed data vectors (ordinate) as a function of transect position (abscissa). **d** Autocorrelation metrics for the designed data including the pattern-partitioning metrics I_G^2 and I_C^2 (Supplementary file ESM 6)

less frequently for I_r relative to I_{3P} . Scaling of both rectified metrics attenuated for larger uncentered I ; however, centered I and the rectified metrics exhibited a strongly linear relationship. The average within-study correlation between I and the rectified metrics was ≈ 0.9 . Example correlations are illustrated in Fig. 7.

The pattern partitioning metrics, I_C^2 and I_G^2 , for the data vector based on the replication and spatial structure of Marrot et al. (2015), but designed for strong gradient, clumping, or mixed patterns, acted in accord with our intent (Fig. 8). The gradient, clumped data, and clumpy gradient had equivalent I_r of 0.29, which is the maximum correlation that can be

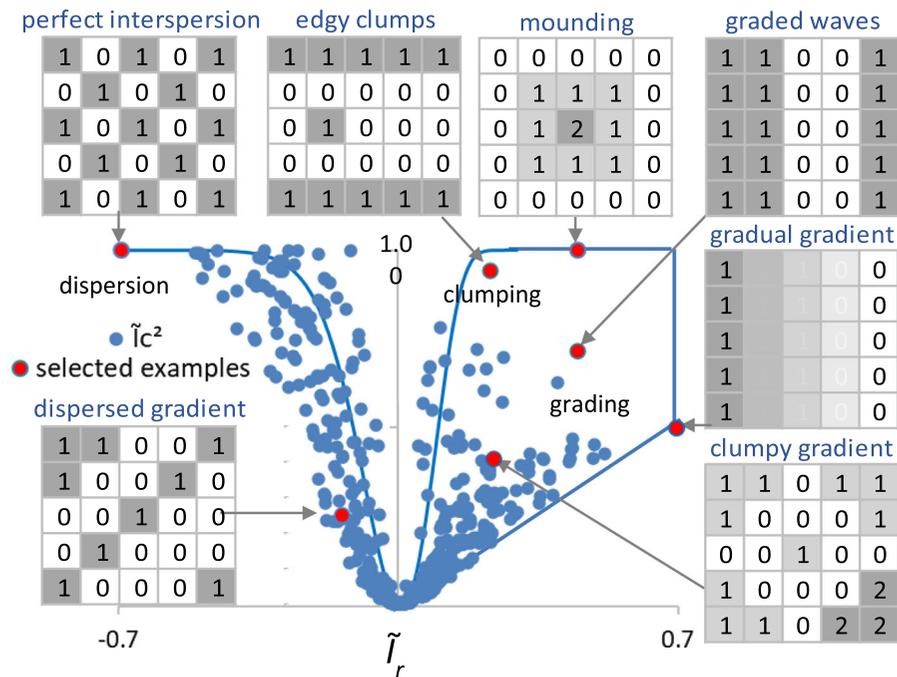


Fig. 9 Covariation of the spatial pattern partitioning metric I_C^2 with the nondispersion metric I_r . Overdispersion was gaged in this space by increasingly negative I_r which also compelled concomitant increase in I_C^2 . The gradient term (multiple r) in the denominator for I_C^2 attenuated the rise of I_C^2 . Nondispersion (clumping and gradients) was indexed by increasingly positive

I_r and either an accelerated or attenuated increase in I_C^2 for patterning that was due to gradients. Thus, I_C^2 rose faster as a function of nongradient clumping over the space of I_r . The differential response of I_C^2 to the two data structures therefore created a means to partition influences of the two structuring mechanisms on overall pattern

achieved with 229 cases and 10^5 permutations with this spatial array. That is, \tilde{I} from none of 10^5 permutations exceeded the observed I for any of the designed arrays (all $\tilde{P}=10^{-5}$). Median-centered I , however, suggested a weaker signal for clumping (Fig. 8d, row 1). Spatial regression demonstrated strong signal for the two cases with gradient patterns with no appreciable signal for the clumped-only data conformations.

The partitioning metric I_G^2 strongly signaled both pure and ‘clumpy gradient’ effects (Fig. 8a, c). I_C^2 strongly signaled only the pure clumping effect (Fig. 8b) with weak to modest indications of pattern in the other two conformations (Fig. 8d, row 6). A broader view of the behavior of I_C^2 emerged from permutations of the data illustrated in Fig. 1a. Figure 9 demonstrates the response of I_C^2 over the range of dispersion observed in random permutations and selected designed data in 5×5 grid conformations (Supplementary file ESM 7). For negative I_r , greater magnitude (neglecting sign) indicated greater dispersion. The negative extreme represented perfect

interspersion. Conversely, the positive extreme of I_r indicated pure gradient conformation which by definition entails clumping of data with progressive but collective monotonic increase in data values over a spatial vector. However, the metric I_C^2 for increasingly positive I_r was particularly sensitive to nongradient clumping (Fig. 9). I_C^2 's complement, I_G^2 , was concomitantly more sensitive to gradients.

Recalling that $I_C^2 = I_r^2 / (I_r^2 + \text{multiple } r^2)$, increasing multiple r as it increases with gradient conformation could only increase the denominator and constrain I_C^2 to smaller values. These effects were evident in Fig. 9 as positive autocorrelation due to pure clumping drove I_C^2 upward to the limit of 1 while wholly graded conformations constrained I_C^2 to its lower limit. These effects were visually fit with an inverted χ^2 function of appropriate dimensionality for the clumping effect and with linear demarcation of the covariance limit fit to its apparency in the plot (Fig. 9; Supplementary file ESM_7).

Sensitivity analyses for the 29 empirical data vectors most often demonstrated the same pattern

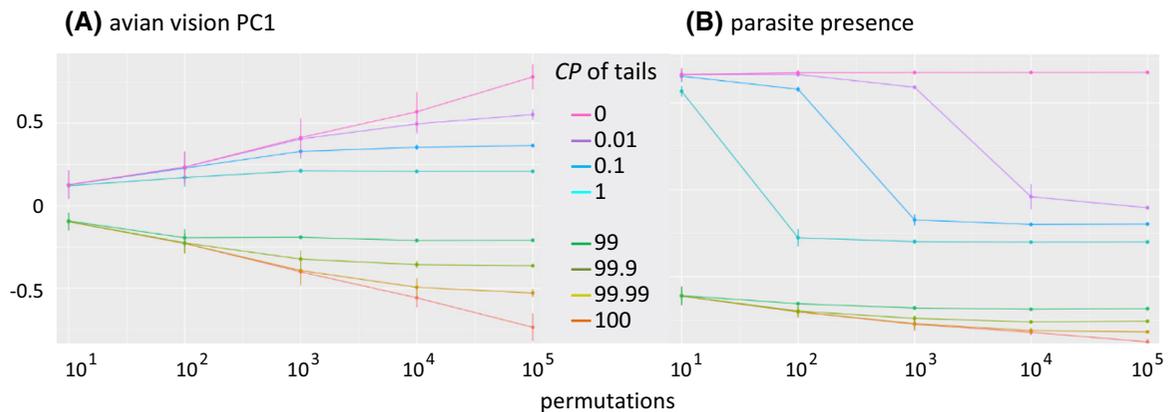


Fig. 10 Sensitivity of frame boundary definitions by permutation number for two data vectors from L’Herpinere et al. (2019). Three point rectification demonstrated two patterns in which greater permutation replication resulted in either wider

(a) or constricted (b) reference frames. Widening of frame boundaries with greater permutation effort was more commonly observed within and across studies

found for the *Cities* data (Fig. 5). Wider frames, hence lesser I_{3P} , generally arose from increasing replication which resulted in differential *increase* in absolute values of frame boundaries, especially for more stringent tail frame designations (Fig. 10a). However, in several instances in the broader survey, increased replication resulted in a *decrease* in frame boundaries for more liberal tail frame designations (Fig. 10b).

Summation regarding metric ideals

Table 6 summarizes performance of each metric of autocorrelation with respect to ideals framed in the Introduction and Background. The cardinal results of these evaluations were: (1) Moran’s I met only 2 of 14 ideal (desired) criteria, failing even those it is reputed to fit. (2) No metric alone met the ability to distinguish pure clumping and gradient data conformations. (3) The ratio $I_C^2 = I_r^2 / (I_r^2 + \text{multiple } r^2)$ was differentially sensitive to nongradient clumping and gradient effects.

Discussion

Through diverse case studies and simulations, we found that Moran’s I varied with spatial pattern but also with several aspects of data structure that distort metrical representations of pattern. I , its transforms, and the other spatial pattern metrics examined,

including matrix correlation and spatial regression, demonstrated a diverse mixture of qualities populating the ‘ideals’ table (Table 1). A reprise of these qualities and performance of the metrics for each is given in Table 6.

The original metric, I , failed to provide a consistent standard for comparing autocorrelation within or among studies (see also Sen 1976; de Jong et al. 1984; Waldhör 1996; Tiefelsdorf and Boots 1997). Our goal was to ‘unwarp’ I —to morph it into an intuitive metric that fit conceptual expectations for both I and for correlations in general. We believe that both rectification methods presented, 3-point registration and fitting to the distribution of r , did this. However, analysis revealed that the 3-point method had a diversity of caveats that made it advisable to conduct sensitivity analyses in each application. Although we provided software for doing the analyses (Fuentes et al. 2020), the sensitivities we felt made 3-point rectification, though preferable to no rectification, less desirable than fitting to the distribution for r . The latter method appeared apt for all the goals described in the “Introduction” sect. and Table 1, except for distinguishing non-gradient clumping from gradient effects. No single metric would be able to distinguish three unique states (overdispersion and two types of exception) which is a two-dimensional issue. But a ratio of I_r and multiple r provided a measure of distinction by indexing contributions of the two non-exclusive patterns (Figs. 8d, 9).

Table 6 Qualities expected, assumed, or desirable for autocorrelation metrics

Expected or desired property of autocorrelation metrics	I	Matrix r	Multiple r	I_{3P}	I_r	I_C^2	Evidence regarding I_r
Responsive to dispersion	+	~	~	+	+	+	Figures 1, 9, ESM_6, 7
Responsive to gradients	+	+	+	+	+	+	Figures 1, 8, 9, ESM_6, 7
Distinguishes clumping from gradient effects	-	-	-	-	-	~	Figures 1,9, ESM_6, 7
Strictly bounded between - 1 and 1	-	+	-	+ ¹	+	+	Equations 5, 6, 7, 8 and 9
Value for perfect interspersions:							
- 1 or commensurate with expectations for regular correlation	-	-	-	+	+	+	Figure 9, ESM_7
Value for perfect gradient:							
1 or commensurate with expectations for regular correlation	-	~	+	+	+	-	Figures 9, ESM_7
Value for gradient-free clustering:							
1 or commensurate with expectations for regular correlation	-	-	-	~	~	-	Figures 8,9, ESM_6, 7
Central tendency for random patterns:							
Mean is 0	-	+	-	-	+	-	Figures 3,4, 11
Median is 0	-	-	-	+	+	-	Figures 3,4,11
Distribution of random pattern metrics is symmetric (opposite)	-	-	-	-	+	-	Figures 3,4, ESM_7
Metric has stand-alone interpretability (contextualized by n)	-	~	+	~	+	-	I_r distributed as r ; Eq. 9, Figs. 4, 11
Metric can be compared among studies (contextualized by n)	-	+	+	+ ³	+	-	I_r distributed as r ; Eq. 9, Figs. 4, 11
Metric itself is a standard effect size	-	-	+	-	+	-	Equations 6, 9
Metric itself is a test statistic	-	-	+	-	+	-	Equations 6,9 Rohlf and Sokal (1995)
Derived null distributions retain original type I error rate				+	+	-	ESM_4, 6

Key: +, -, and ~ indicate that the metric meets, fails, or approximately (or weakly) fits the criterion

I_{3P} may be capped such that values in excess of |1| are redefined based on their sign to ± 1 (Eq. 3)

² The median is a preferable measure of central tendency for asymmetric distributions where 0 is intended to indicate lack of pattern

³ Comparability among studies requires the same frame definitions and contextualization by n

All of the I metrics rendered identical \tilde{P} -values from Monte Carlo simulation, evincing preservation of type I error despite centering and the two scaling methods. Scaling I was not intended to change inferential outcomes regarding autocorrelation. Our intention was to project the well-recognized traditional metric to a quantitative and conceptual space concordant with the inertia of existing ideals and intuition. Such inertias have been termed ‘apperception’, meaning mental anticipation based on the sum of prior experience (Greenwood 2015). Apperception regarding I mirrors that for regular correlation per Pearson (1924). Correlations are expected to be t -distributed and to span the opposite frame $[-1,0,1]$

with 0 indicating absence of pattern (Sokal and Rohlf 1995). A given value of correlation implies the same effect strength for variable relationships from similarly replicated studies, though qualitative interpretation of the effect strength varies by discipline (e.g., Möller and Jennions 2002; Mukaka 2012). Rectification by 3-point registration eliminated bias (i.e. the median null value was 0), greatly reduced skewness, and generally expanded the scale of the null distribution of I to a semblance of that for correlation (Fig. 4). However, the scale and null distribution of I_r was fit by definition to that for regular correlation. Thus, I_r registered precisely with apperception for correlation.

Moran's I is often cited as derivative of Pearson's correlation coefficient, implying a homology and hence theoretical basis for I that is entirely false. An I metric that is dimensionally, volumetrically, and distributionally homologous to r may not be possible to derive theoretically given the multifarious disparities that emerge due to factors such as alternative data types and patterns of site proximities. Suspending doubt, if such a direct, mathematically homologous version of I were possible, with dimensionality, scale, and variable distributions homologous to r , then it follows that the distribution of said metric would match that for r . In substitution for that derivation, we can be true to the Pearsonesque theoretical basis for a new I derived directly from probability theory.

The approach of converting other statistics to r is increasingly used in meta-analysis to enable the sorts of comparisons we have described (Rosenthal and Rubin 1986; Möller and Jennions 2002). Although the conversion to r in the approach illustrated herein is conceptually related, it is free from parametric assumptions. Instead of calculating r from test statistics, effect sizes, or P -values based on them, fitting uses null distributions created directly from empirical data. Therefore, this fitting method makes no assumptions about the specific geometry of residual or null distribution shapes, or the scale of data matrices. While not being wedded to assumptions of parametric test statistics, I_r relied on permutation and its improbable null extrema were limited by replication effort. Minimum \tilde{P} depended on permutation number and hence maximum I_r was limited through Eqs. 8 and 9. For example, maximum I_r from 20-case datasets with a true I_r of 0.9, assuming that is known, by using 10^4 , 10^5 , or 10^6 replicates yields I_r estimates of 0.76, 0.82, and 0.86 respectively. In our frequentist approach, \tilde{P} did not change if I_r exceeded the null distribution maximum by 1 unit or 100. In parametric statistics, it is common for P to short \tilde{P} by many orders of magnitude where \tilde{P} is limited by permutation replication. This limit to I_r only occurred when an observed dataset was extremely (very improbably if randomly) patterned. So for most empirical cases the issue may not be relevant, but at least two solutions may obviate the limit. First, for cases in which \tilde{P} equals the lower limit of $1/(n+1)$, one could increase replication effort until a threshold number of \tilde{I} in excess of I are

returned. If replication had to be incremented enough to overtax computing resources, I_r from the simulations conducted to that point could be plotted as in our sensitivity analyses and an asymptotic value would likely be evident. Second, rather than the try-and-try-again procedure just outlined, one could use a parametric test to get an initial estimate for k ($=1/P-1$) sufficient to yield an unconstrained I_r or the next order of magnitude higher to ensure adequacy. This assures that I_r is not capped and therefore is comparable to an empirical parametric r .

Although the 'exact distribution of I ' has been derived with assumptions standard for parametric statistical inference (Tiefelsdorf 1998), the multifarious nature of real-world applications may limit its utility. For example, dimensionality of null distributions varies based on geometry of sample locations, measurement distributions, and procedural aspects of calculating I such as contiguity definitions. Although positive definite distance matrices have singular dimensionality, inversion of its values produces proximities with unpredictable dimensionality. The 5×5 uniform grid pattern of Fig. 1 had 6 dimensions (positive eigenvalues) yet that for 25 sampling sites in unconstrained (normally distributed) conformation had 7–9 dimensions for alternative permutations (Supplementary file ESM_1). Similarly problematic is the z parameter frequently used as a test statistic for I (Cliff and Ord 1969). Its values are χ^2 -distributed with various dimensionality, so it should perhaps always be paired with significance tests by simulations. Rectifying I to an alternative distribution would seem suspect for subverting 'the exact distribution' and its probability architecture. Though the values of I are rescaled in our correlation-fitting method, the probability structure among values (differences in CP among any pair of I) is unaltered by rectification.

Projecting I based on its null distribution to the theoretical distribution for r also eliminates the need to scale data inputs. Calculation of I_r requires neither scaling the proximity matrix to unit sum nor scaling of the centered data vector to unit variance per Moran (1950). It is only necessary to create a cross product by pre- and post-multiplying the unscaled proximity matrix by the centered data vector, $\mathbf{v}'\mathbf{P}\mathbf{v}$ and then converting this to I_r using its cumulative probability within its null distribution. Supplementary file ESM_8 demonstrates the equivalence of using scaled or unscaled data arrays.

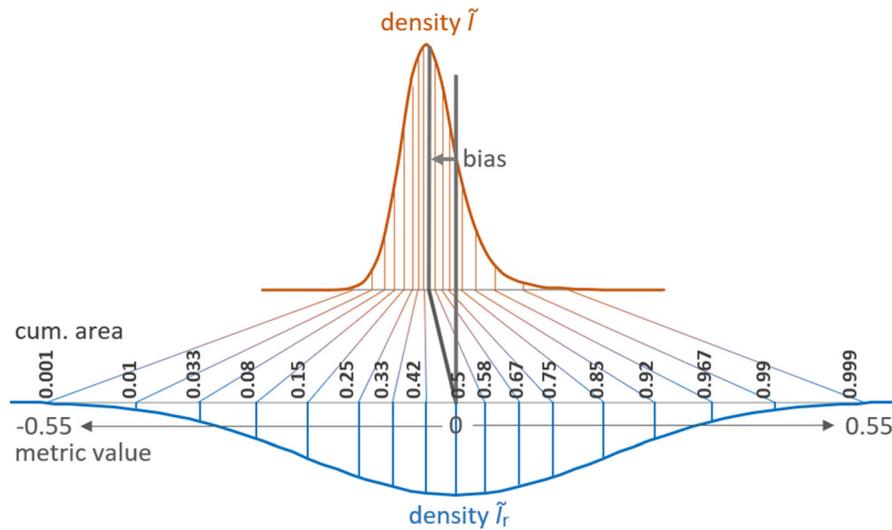


Fig. 11 Mapping of a null distribution of I to that of I_r

The only caveats we conceived for rectifying I to the ideals frequently assumed or desired relate to the ease it grants to pass over consideration of the potentially informative if messy reality of raw null distributions. The tidied view offered by I_r may prevent insights that could flow from observing the original null distribution shape, such as the utility of data transformation. Worse, a ‘normalized’ frame of view may prevent recognition of functional inferences regarding the system or discovery of higher order phenomena (see DeWitt 2016, for ‘bell curve’ biases in evolutionary biology). To prevent such a disconnect between the original and rectified distributions, it is advised to visualize both and perhaps the geometric mapping between them as in Fig. 11 (Supplementary file ESM_9).

The ability to compare I among variables and studies is essential if the metric is to be useful beyond solitary instances and to facilitate its expansion to multiple domains (Kim et al. 2015). Presently, empirical values and null distributions typical of I appear as varied contortions that fit neither each other nor a theoretical or idealized standard. Examples of such contortions can be readily observed among our empirical survey results (Table 5). Two data vectors from Goldsmith et al. (2019) yield $I=0.02$ and -0.04 , yet the former is highly significant and the latter is not ($\tilde{P}=0.007$ and 0.12) in dissonance with the expectation of oppositivity. In terms of scale, an I from L’Herpinier et al. (2019) was inordinate at

1.2 while an I from Marrot et al. (2015) was miniscule at 0.01 and both were significant. Results so misfit to each other and to the scale of regular correlations are difficult to reconcile with intuition. The inscrutable nature of I thereby also precludes comparisons among studies. Rectification trues I to a standard that meets the apperceptive mass of ideals thought, taught, and wrought in the statistical and geospatial literature. We can continue to qualify every I statistic for each of its idiosyncrasies or we can present it in a manner that does not require those labors.

Funding This project was funded by Texas A&M University T3 (Triads for Transformation) grant to TJD, MPB, and TRI.

References

- Anselin L (1988) Spatial econometrics: methods and models. Kluwer, Dordrecht
- Anselin L (1995) Local indicators of spatial association—LISA. *Geogr Anal* 27:93–115
- Chen Y (2013) New approaches for calculating Moran’s index of spatial autocorrelation. *PLoS ONE* 8:e68336
- Cliff AD, Ord JK (1969) The problem of spatial autocorrelation. In: Scott AJ (ed) *Studies in regional science*, vol 1. Pion, London, pp 25–55
- Cliff AD, Ord JK (1973) *Spatial autocorrelation*. Pion, London
- Cliff AD, Ord JK (1981) *Spatial processes: models & applications*. Pion, London

- Cozzarolo CS, Jenkins T, Toews DPL, Brelsford A, Christie P (2018) Prevalence and diversity of haemosporidian parasites in the yellow-rumped warbler hybrid zone. *Ecol Evol* 8:9834–9847
- D’agostino RB, Belanger A, D’agostino RB Jr (1990) A suggestion for using powerful and informative tests of normality. *Am Stat* 44(4):316–321
- de Jong P, Sprenger C, van Veen F (1984) On extreme values of Moran’s *I* and Geary’s *c*. *Geogr Anal* 16:17–24
- DeWitt TJ (2016) Expanding the phenotypic plasticity paradigm to broader views of trait space and ecological function. *Curr Zool* 62:463–473
- Dokmanic I, Parhizkar R, Ranieri J, Vetterli M (2015) Euclidean distance matrices: essential theory, algorithms, and applications. *IEEE Signal Process Mag* 32:12–30
- Durbin J, Watson GS (1950) Testing for serial correlation in least squares regression. *Biometrika* 37:409–428
- Fortin M, Dale M (2005) *Spatial analysis: a guide for ecologists*. Cambridge University Press, New York
- Fuentes JJ, DeWitt TJ, Ioerger TR, Bishop MB (2020) Irescale: calculate and rectify Moran’s *I*. R package v. 2.3.0. <https://cran.r-project.org/package=Irescale>. Accessed 10 Dec 2020
- Getis A (2010) Spatial autocorrelation. In: Fischer MM, Getis A (eds) *Handbook of applied spatial analysis: software tools, methods and applications*. Springer, Heidelberg
- Goldsmith GR, Allen ST, Braun S, Engbersen N, González-Quijano CR, Kirchner JW, Siegwolf RTW (2019) Spatial variation in throughfall, soil, and plant water isotopes in a temperate forest. *Ecohydrology* 12:e2059
- Goodchild MF (1988) Spatial autocorrelation. Concepts and techniques in modern geography. Geobooks, Norwich
- Greenwood JD (2015) *A conceptual history of psychology: exploring the tangled web*. Cambridge University Press, New York
- Griffith DA (1987) *Spatial autocorrelation: a primer*. Association of American Geographers, Resource Publications in Geography, Washington DC
- Hotelling H (1953) New light on the correlation coefficient and its transforms. *J R Stat Soc B* 15:193–232
- Hurley JR, Cattell RB (1962) The procrustes program: producing direct rotation to test a hypothesized factor structure. *Behav Sci* 7:258–262
- L’Herpinier KL, O’Neill LG, Russell AF, Duursma DE, Griffith SC (2019) Unscrambling variation in avian eggshell colour and patterning in a continent-wide study. *R Soc Open Sci* 6:181269
- Kim D, DeWitt TJ, Costa CSB, Kupfer JA, McEwan RW, Stallins JA (2015) Beyond bivariate correlations: three-block partial least squares illustrated with vegetation, soil, and topography. *Ecosphere* 6:135
- Li H, Calder CA, Cressie N (2007) Beyond Moran’s *I*: Testing for spatial dependence based on the spatial autoregressive model. *Geogr Anal* 39:357–375
- Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Res* 27:209–220
- Marrot P, Garant D, Charmantier A (2015) Spatial autocorrelation in fitness affects the estimation of natural selection in the wild. *Methods Ecol Evol* 6:1474–1483
- Möller AP, Jennions MD (2002) How much variance can be explained by ecologists and evolutionary biologist? *Oecologia* 132:492–500
- Moran PAP (1947) Some theorems on time series. *Biometrika* 34:281–291
- Moran PAP (1950) Notes on continuous stochastic phenomena. *Biometrika* 37:17–23
- Mukaka MM (2012) Statistics corner: a guide to appropriate use of correlation coefficient in medical research. *Malawi Med J* 24:69–71
- Pearson K (1924) Historical note on the origin of the normal curve of errors. *Biometrika* 16:402–404
- Rahman NA (1968) *A course in theoretical statistics*. Charles Griffin and Company, Glasgow
- Ritters K (2019) Pattern metrics for a transdisciplinary landscape ecology. *Landsc Ecol* 34:2057–2063
- Rogerson PA (1999) The detection of clusters using a spatial version of the chi-square goodness-of-fit statistic. *Geogr Anal* 31:130–147
- Rohlf FJ, Sokal RR (1995) *Statistical tables*. Macmillan, New York
- Rosenthal R, Rubin DB (1986) Meta-analytic procedures for combining studies with multiple effect sizes. *Psychol Bull* 99:400–406
- Sen A (1976) Large sample-size distribution of statistics used in testing for spatial correlation. *Geogr Anal* 8:175–184
- Sokal RR, Rohlf FJ (1995) *Biometry*, 3rd edn. WH Freeman, San Francisco
- Thirey B, Hickman R (2015) Distribution of Euclidean distances between randomly distributed Gaussian points in *n*-space. arXiv preprint arXiv:1508.02238. <https://arxiv.org/pdf/1508.02238.pdf>. Accessed 20 Dec 2020
- Tiefelsdorf M (1998) Some practical applications of Moran’s *I*’s exact conditional distribution. *Pap Reg Sci* 77:101–129
- Tiefelsdorf M, Boots B (1997) A note on the extremities of local Moran’s *I*_s and their impact on global Moran’s *I**. *Geogr Anal* 29:248–257
- Tobler W (1970) A computer movie simulating urban growth in the Detroit region. *Econ Geogr* 46(Supplement):234–240
- Upton GJ, Fingleton B (1985) *Spatial data analysis by example, volume 1: Point pattern and quantitative data*. Wiley, New York
- Waldhör T (1996) The spatial autocorrelation coefficient Moran’s *I* under heteroscedasticity. *Stat Med* 15:887–892

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.