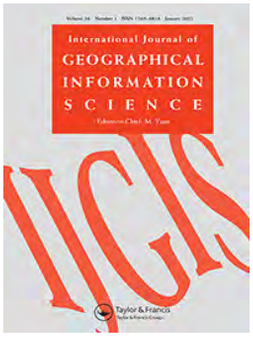

International Journal of Geographical Information Science

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tgis20

A multiscale measure of spatial dependence based on a discrete Fourier transform

Hanchen Yu & A. Stewart Fotheringham

To cite this article: Hanchen Yu & A. Stewart Fotheringham (2021): A multiscale measure of spatial dependence based on a discrete Fourier transform, International Journal of Geographical Information Science, DOI: 10.1080/13658816.2021.2017440

To link to this article: https://doi.org/10.1080/13658816.2021.2017440

Published online: 21 Dec 2021.

INTERNATIONAL JOURNAL OF GEOGRAPHICAL INFORMATION SCIENCE
https://doi.org/10.1080/13658816.2021.2017440

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

A multiscale measure of spatial dependence based on a discrete Fourier transform

Hanchen Yu[a] and A. Stewart Fotheringham[b]

[a]Center for Geographic Analysis, Harvard University, Cambridge, MA, USA; [b]Spatial Analysis Research Center, School of Geographical Sciences and Urban Planning, Arizona State University, Tempe, AZ, USA

ABSTRACT

The measurement of spatial dependence within a set of observations or the residuals from a regression is one of the most common operations within spatial analysis. However, there appears to be a lack of appreciation for the fact that these measurements are generally based on an a priori definition of a spatial weights matrix and hence are limited to detecting spatial dependence at a single spatial scale. This paper highlights the scale-dependence problem with current measures of spatial dependence and defines a new, multi-scale approach to defining a spatial weights matrix based on a discrete Fourier transform. This approach is shown to be able to detect statistically significant spatial dependence which other multi-scale approaches to measuring spatial dependence cannot. The paper thus serves as a warning not to rely on traditional measures of spatial dependence and offers a more comprehensive, and scale-free, approach to measuring such dependence.

ARTICLE HISTORY
Received 17 January 2021
Accepted 8 December 2021

KEYWORDS
Multiscale; spatial dependence; *Moran's I*; Fourier transform; spatial weights matrix

1. Background and introduction to the problem

Spatial dependence is an attribute of almost all spatial data; indeed it is quite difficult to conceive of real data which are distributed randomly over space. However, the degree to which similar values of a variable are located in close proximity to each other can vary from being very weak to very strong so that the measurement of spatial dependence is a common component of much spatial analysis. It is useful to be able to describe quantitatively the degree to which similar values are clustered together in space. This is not only important as a descriptive measure of a distribution but it also has implications for the calculation of the effective sample size in assessing inference and power in many statistics applied to spatial data. The measurement of spatial dependence is also important in assessing the degree to which residuals in a regression exhibit spatial autocorrelation because if the residuals exhibit sufficient dependence, this may invalidate the assumption that they are independent and invalidate the basis on which inference is made (Cliff and Ord 1972, Kelejian and Robinson 1992).

For these reasons, there is a vast literature on various types of statistical measures of spatial dependence although two measures have come to dominate the literature: *Moran's I* and *Geary's C* (Moran 1950, Geary 1954, Getis and Ord 1992, Anselin 1995,

CONTACT A. Stewart Fotheringham sfotheri@asu.edu

2019, Ord and Getis 1995, Kelejian and Prucha 2001, Delgado and Robinson 2015). Arguably, the former has gained pre-eminence in the literature possibly because of its convenient scale between −1 (extreme negative spatial autocorrelation) and +1 (extreme positive autocorrelation) with a mid-point which is almost zero [−1/(N-1), where N is the number of points] indicating no spatial dependence. Consequently, in what follows we use *Moran's I* as an exemplar of the standard type of spatial dependence measure which we show can have severe flaws if the data being examined result from multiple processes operating at different scales.

That measures of spatial dependence are scale-dependent is not new: the issue has been studied for decades and various solutions have been proposed involving the construction of local measures of spatial dependence (Getis and Ord 1992, Anselin 1995, Ord and Getis 1995), using measures at different scales (Meisel and Turner 1998, Fauchald *et al*. 2000, Wu *et al*. 2000, 2020, Hay *et al*. 2001, Zhang and Zhang 2011, Westerholt *et al*. 2015, 2018, Oman and Mateu 2019), and optimizing the spatial weights matrix (Getis and Aldstadt 2004, Rogerson 2011, Rogerson and Kedron 2012, Bauman *et al*. 2018). Here, we propose a new measure of spatial dependence based on a Fourier transform of a spatial distribution (Fuentes 2007, Bandyopadhyay *et al*. 2015, Rao 2018, Guinness 2019). We first describe the problem that typical measures of dependence are usually arbitrary being sensitive to how spatial scale is defined. We then examine two methods that might provide a solution, a multiscale measure of *Moran's I* and a spatial variogram, but find them lacking. We therefore propose a new multiscale measure of spatial dependence based on a Fourier transform and show how this can be used to define a scale-free spatial weights matrix for generic use in spatial analysis.

2. Single scale and multiscale measures of spatial dependence

Consider a variable x which is distributed over a set of N locations, each labeled i. *Moran's I* is then defined as:

$$I = \frac{N \sum_{i=1}^{N} \sum_{j=1}^{N} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{N} \sum_{j=1}^{N} w_{ij} \sum_{i=1}^{N} (x_i - \bar{x})} \quad (1)$$

where N is the number of locations indexed by i and j; $\underset{0.3em\$-\$}{x}$ is the mean of x; and w_{ij} is a matrix of spatial weights. One immediate problem with the use of any measure of spatial dependence is the definition of the spatial scale over which dependence is defined as this can have a major impact on the determination of whether significant spatial dependence exists.[1] In the formula above, for example, this scale definition problem manifests itself through the definition of the spatial weights matrix which can have many forms such as:

(i) $w_{ij} = 1$ if i and j share a common edge (if i and j represent polygons); 0 otherwise;

(ii) $w_{ij} = 1$ if i and j are the 1-kth nearest neighbors; 0 otherwise;

(iii) $w_{ij} = 1$ if i and j share a common point of contact (if i and j represent polygons); 0 otherwise;

(iv) w_{ij}=1 if the distance between i and j is smaller than a threshold d; 0 otherwise;

(v) $w_{ij} = 1/d_{ij}^{\beta}$ where d_{ij} is the distance between i and j.

Each of these definitions will lead to a different value of *I* being calculated. Even where the weights exhibit a continuous decline as distance increases, as in (v), the scale problem exists because the weights will depend on β with the weights decreasing more rapidly as distance increases the larger is β.

The scale problem in measuring spatial dependence is shown clearly in the situation described in Figure 1. Figure 1(a) represents a distribution $\lambda = \xi + \eta$ with a checkerboard pattern with maximum and minimum values of 1.5 and −1.5 $\eta = \pm 1.5$ overlaying a mean-stationary Gaussian random field $\xi \sim N(0, \Sigma)$, where Σ denotes the covariance matrix arising from a squared exponential spatial covariance function as follows:

$$\Sigma(k,\ h) = exp\left(-0.5 * \frac{{d_{ij}}^2}{b^2}\right) \tag{2}$$

In this notation, *b* is the parameter controlling the surface's dependence and is set to 4 to derive the surface in Figure 1(a), and d_{ij} is the distance between locations *i* and *j*. The distribution is one such that at a regional scale high values tend to cluster in the south and southeast with low values clustering in the north but at a local scale there is negative dependence with relatively high values being adjacent to relatively low values. Figure 1(b) represents the values of *Moran's I* and associated β values for this distribution using a continuous distance-based weights matrix with different distance powers (effectively measuring spatial dependence at different scales) such that:

$$w_{ij} = \frac{1}{{d_{ij}}^{\beta}} \tag{3}$$

where β is a scaling factor. As β tends to 0, the weights tend to 1 and as β increases, spatial dependence is measured at an increasingly local scale. When β is low (≤ 1.5), spatial dependence is measured regionally and *Moran's I* is significantly positive. As β increases, dependence is measured increasingly locally and decreases to zero becoming insignificant when $1.5 < \beta < 2$. When $\beta > 2$, significant negative spatial dependence is detected. This demonstrates the problem with standard measures of spatial dependence – they are

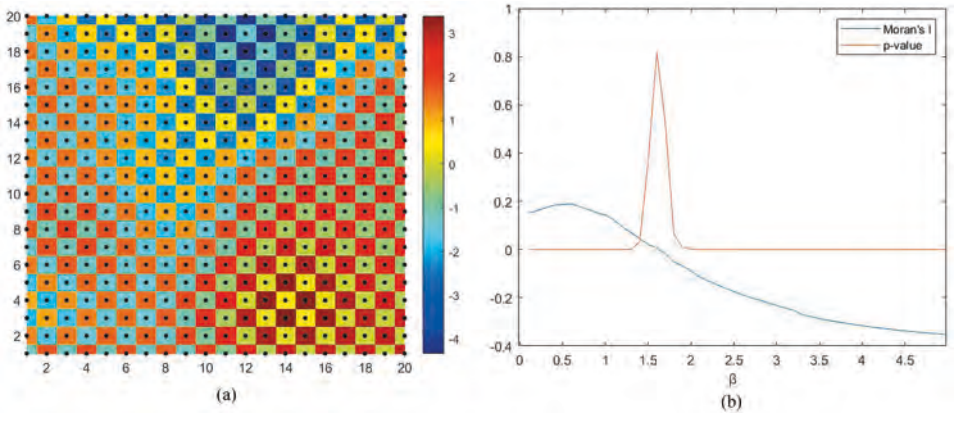

Figure 1. (a) A spatial distribution; (b) *Moran's I* and associated β values calculated at different spatial scales.

scale dependent and any particular measurement of spatial dependence will only be valid at one particular spatial scale. If a measure such as *Moran's I* or *Geary's C* is to be used, values should be reported at all spatial scales such as is done in Figure 1(b), unless of course a specific scale of analysis has meaning and in this case a single-scale measure of dependence is appropriate. Generally though the measurement of spatial dependence is undertaken at a purely arbitrary scale as defined by the definition of the spatial weights matrix.

However, a multi-scale description of *Moran's I*, as shown in Figure 1(b), still has limitations when a spatial distribution results from a combination of processes operating at different spatial scales. Consider, for example, the three surfaces in Figure 2. The distributions shown in Figure 2(a) and 2(b) are derived from a Gaussian Random Field in which the parameter b in equation (2) equals 6 and 1 respectively. The distribution in Figure 2(a) has a high degree of spatial dependence; the distribution in Figure 2(b) has a low degree of spatial dependence. The distribution in Figure 2(c) is the cell-wise summation of the two distributions in Figure 2(a) and 2b. The multi-scale plots of *Moran's I*, equivalent to that in Figure 1(b), for each of these three surfaces are shown in Figure 3.

In Figure 3(a) the value of *Moran's I* is always significantly positive and rises rapidly as the value of I is calculated at increasingly local scales. In Figure 3(b) the value of *Moran's I* is insignificant at a broad regional scale (when $\beta < 0.25$) and becomes significantly positive at more local scales of analysis (when $\beta > 0.25$). The increase in *Moran's I* as β

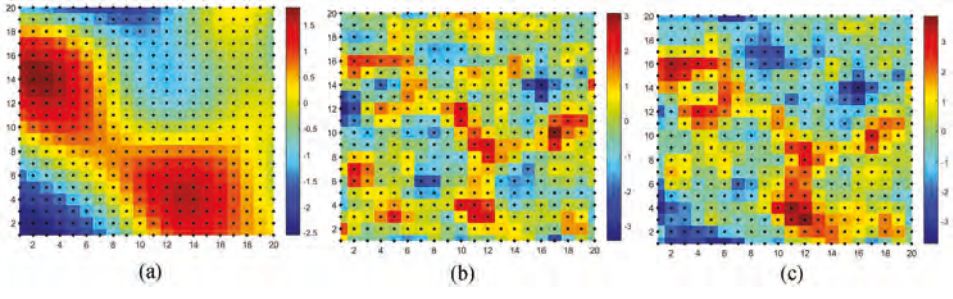

Figure 2. (a) Distribution with high spatial dependence; (b) Distribution with low spatial dependence; (c) A cell-wise summation of the two distributions.

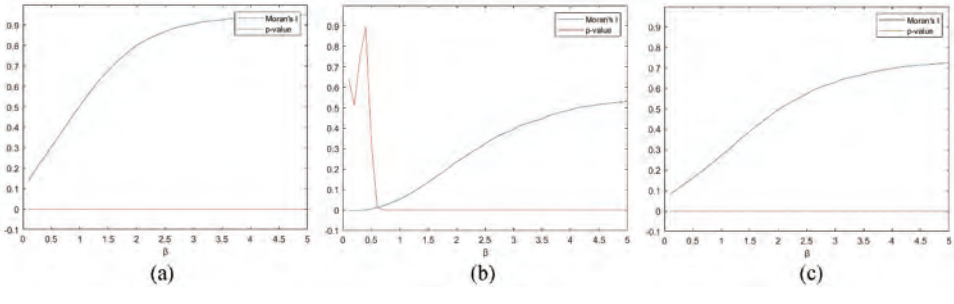

Figure 3. Multi-scale *Moran's I* plots with associated β values for the distributions shown in Figure 2 (a-c).

increases is less steep than for the distribution in Figure 2(a). The plot in Figure 2(c) is a compromise of the plots in Figure 2(a) and 2(b), which is to be expected given the distribution in Figure 2(c) is a combination of the two distributions in Figure 2(a) and 2(b). The problem is that if faced with Figure 2(c) alone, we have no way of identifying that the surface results from two different processes operating at different spatial scales. What is needed is a multi-scale measure of spatial dependence that can identify if a distribution results from two or more processes operating at different spatial scales. To this end, the spatial variogram was developed based on a measure of semivariance defined as:

$$\gamma_{ij} = 1/2 * var(Z_i - Z_j) \tag{4}$$

where Z_i is the value of an attribute at location i, Z_j is the value at location j and *var* is a variance operator so that semivariance is a measure of the degree to which values of a distribution a certain distance apart are dissimilar. If a surface were constant, for example, the semivariance would be 0 for all distances. For a stationary isotropic process, the semivariance only depends on the distance h between locations (Cressie 1993) so that semivariance can be represented by $\gamma(h)$ and a semivariogram can be constructed by plotting the value of $\gamma(h)$ against h.

Figure 4(a-c) contains the variograms for the three distributions shown in Figure 2 with semivariance on the vertical axis plotted against h. In distributions with positive spatial dependence, as h tends to 0, $\gamma(h)$ tends to 0 and as h increases, the semivariance increases. The slope of the variogram indicates the degree of dependence in the distribution; the steeper the slope, the lower is the spatial dependence in the distribution. Comparing the results for the two distributions shown in Figure 2(a) and 2(b), the semivariogram corresponding to Figure 2(a) is less steep and reaches a plateau (the sill) at a value of h =9. In Figure 4(b) the slope is steeper indicating less spatial dependence and reaches its sill at around h =3. The semivariogram of the composite distribution in Figure 2(c) indicates the possibility that the distribution is comprised of two processes operating at different scales with a steeper slope (less dependence) from h = 0 to 3 and a gentler slope (greater dependence) between h =3 and h =9, at which point the sill is reached and there is no further dependence at this distance.

Although semivariograms can identify multi-scale processes and therefore appear to be superior to a multi-scale *Moran's I*, they are not always reliable. For instance, consider the three distributions in Figure 5 which are similar to those in Figure 2 except that

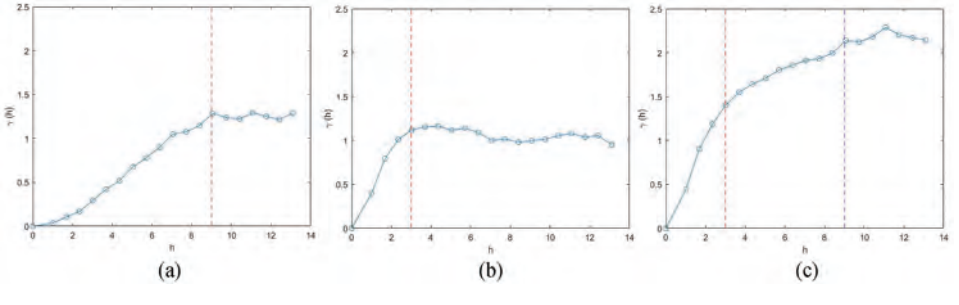

Figure 4. Semivariograms for the three distributions shown in (a) Figure 2(a); (b) Figure 2(b); (c) Figure 2(c).

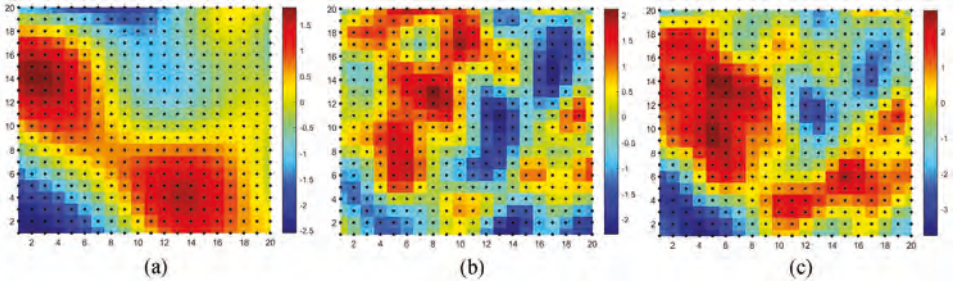

Figure 5. Three spatial distributions (a) strong spatial dependence; (b) medium spatial dependence; (c) a combination of (a) and (b).

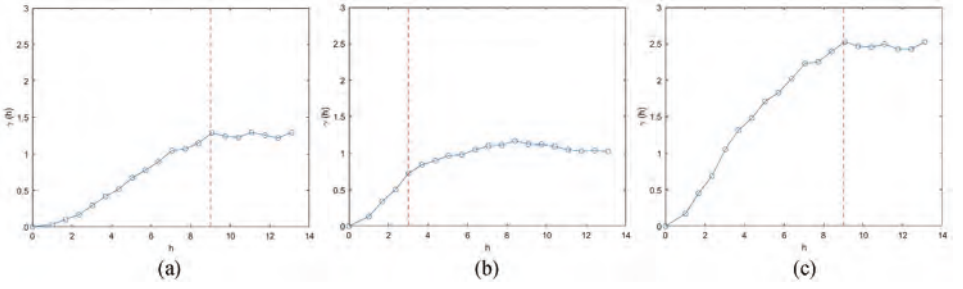

Figure 6. Semivariograms for the three distributions in Figure 5.

the second distribution now has a greater degree of spatial dependence. Figure 6 shows the semivariograms for the three distributions with Figure 6(c) highlighting the problem. In Figure 6(a) the semivariogram has a relatively gentle slope indicating strong dependence with a sill at around h =9. In Figure 6(b) the semivariogram has a slightly steeper slope indicating slightly weaker dependence with a sill around h =3. In Figure 6(c), however, there is no evidence of multi-scale processes generating the distribution with a single gradient indicating relatively weak dependence up to the sill which is reached around h =9. Hence, there is a need for a more sensitive multi-scale measure of spatial dependence and for this we turn to discrete Fourier transforms.

3. The discrete Fourier transform in 1-D and 2-D

There are two perspectives on a distribution. One is to regard it as a set of points in space, which is the traditional view and the one used exclusively so far in this paper. Another is to view a distribution as the combination of a set of waves, each having a frequency and an amplitude (Matsuda and Yajima 2009, Perraudin and Vandergheynst 2017, Deb and Wu 2017). High-frequency waves represent rapid changes in space; low-frequency waves represent more gradual changes over space. The magnitudes of the changes across space are represented by the amplitudes of the waves; the direction of the wave determines the direction of these changes; and the phase of the wave describes the relative position of the wave within its cycle. A discrete Fourier transform is a method of expressing a spatial distribution as a set of waves.

To begin the discussion of the application of discrete Fourier transforms to measure multi-scale spatial dependence in spatial distributions, we begin by demonstrating the application of a discrete Fourier transform to a 1-dimensional distribution. The discrete Fourier transform, $F(X)$, of a vector X of length m is defined as follows:

$$F(X) = Y_{p+1} = \sum_{j=1}^{m} \omega_m^{jp} X_j \tag{5}$$

where Y_{p+1} is the frequency spectrum, p is the frequency from 0 to m-1, ω_m is the complex root of unity and $\omega_m = e^{-2\pi i/m}$, where i is the imaginary unit. As an example, consider the 1-dimensional distributions in Figure 7(a) and 7(b) which represent two sine waves of different frequencies with random noise having zero mean. Figure 7(c) is a composite of these two waves. Figure 8(a-c) represent the discrete Fourier transforms of these three distributions with Figure 8(a) and 8(b) clearly identifying the two sine waves with different frequencies and Figure 8(c) identifying that the distribution in Figure 7(c) results from a mixture of two processes.

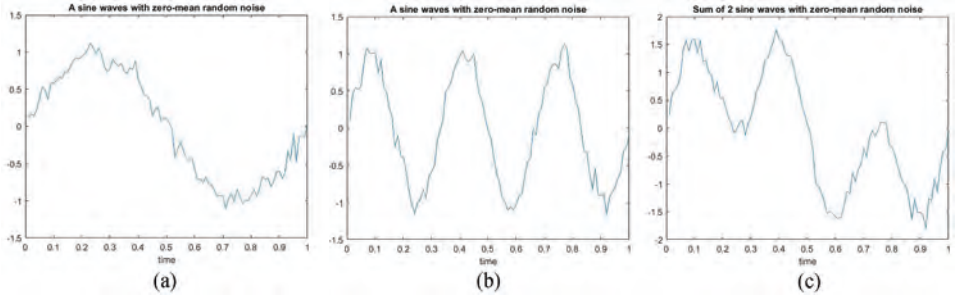

Figure 7. Sine waves (a) frequency XX; (b) frequency YY; (c) a composite of (a) and (b).

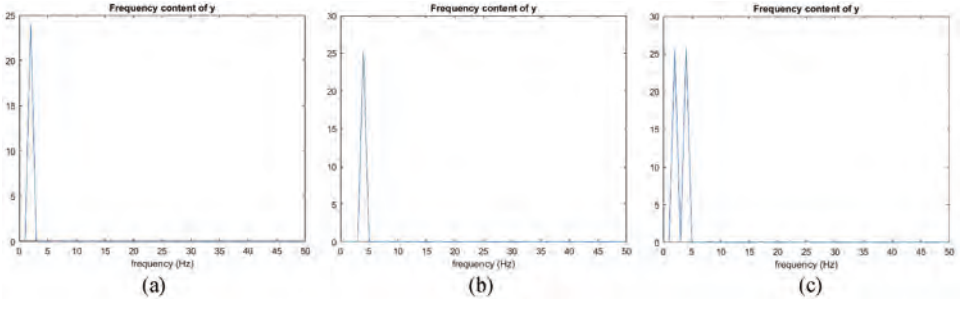

Figure 8. Discrete Fourier transforms of the distributions in Figure 7(a-c).

In two dimensions, the discrete Fourier transform projects a process from the space domain, *(xc, yc)*, where *xc* and *yc* represent the geographic coordinates of a location on a spatial plane, to the frequency domain, *(p, q)*, where *p* and *q* range from 0 to *m-1* and represent the direction and frequency of a 2-D wave. Consider an *m* by *n* spatial process, *X*, whose autocorrelation function is defined as[2]:

$$R_X(i,j) = E\left(X_i * X_j\right) \tag{6}$$

where $R_X(i,j)$ is the autocorrelation function between location *i* and *j* and $E\left(X_i * X_j\right)$ is the expectation of the product of process *X* at locations *i* and j, $X_i * X_j$. The power spectral density of this process, $S_X(p,q)$, is the discrete Fourier transform of the corresponding autocorrelation function and is an expression of dependence in the frequency domain. It indicates at which frequencies variations in autocorrelation are either strong or weak. The power spectral density is defined as the square modulus of the discrete Fourier transform of *X* as follows:

$$S_X(p,\ q) = |F(X)|^2/(m*n) = F(X).\bar{F}(X)/(m*n) \tag{7}$$

where *p* and *q* represent a location in the frequency domain which determine the direction and frequency of a 2-D wave. The period of each wave is $(p^2+q^2)^{-\frac{1}{2}}$. In order to calculate $S_X(p,q)$ we therefore need to calculate the discrete Fourier transform of the process *X*, F(X), and the conjugate transform, $\bar{F}(X)$. The former is given by:

$$F(X) = Y_{p+1,\ q+1} = \sum_{i=1}^{m}\sum_{j=1}^{n} \omega_m^{(i-1)p}\omega_n^{(j-1)q}X_{i,\ j} \tag{8}$$

where $Y_{p+1,q+1}$ is the frequency spectrum, ω_m and ω_n are the complex roots of unity, $\omega_m = e^{-2\pi i/m}$,$\omega_n = e^{-2\pi i/n}$ and *i* is the imaginary unit. The latter is given by:

$$\bar{F}(X) = \bar{Y}_{p+1,\ q+1} = \sum_{i=1}^{m}\sum_{j=1}^{n} \omega_m^{-(i-1)p}\omega_n^{-(j-1)q}X_{i,\ j} \tag{9}$$

The Wiener-Khinchin theorem states that the power spectral density, $S_X(p,q)$, of a wide-sense stationary random process is the Fourier transform of the corresponding autocorrelation function (Wiener 1930, Yao and Journel 1998, Ver *et al*. 2004, Perraudin and Vandergheynst 2017):

$$S_X(p,\ q) = F(R_X(i,\ j)) \tag{10}$$

For a realization of a process, the expectation $R_X(i,j)$ can be estimated by the convolution of data:

$$C_X(i,\ j) = \frac{1}{mn}\sum_{k=1}^{m}\sum_{l=1}^{n} X_{k,\ l}X_{k+i,\ l+j} \tag{11}$$

which is a local measure of spatial dependence and which generally decreases with distance. For example, in Figure 9 we display the values of $C_X(i,j)$ relating to the three surfaces in Figure 2. In Figure 10 we show the corresponding power spectral densities, and

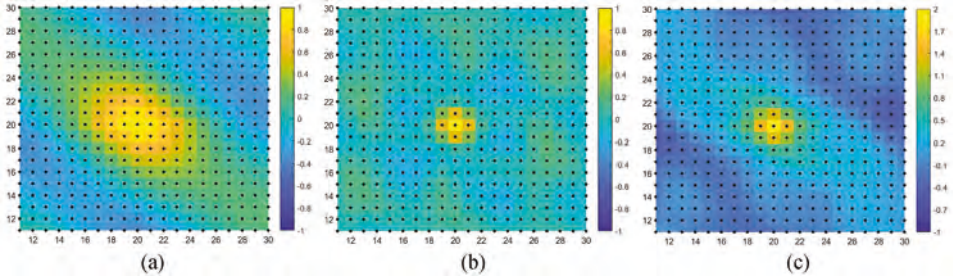

Figure 9. The values of $C_X(i,j)$ corresponding to the three surfaces in Figure 2.

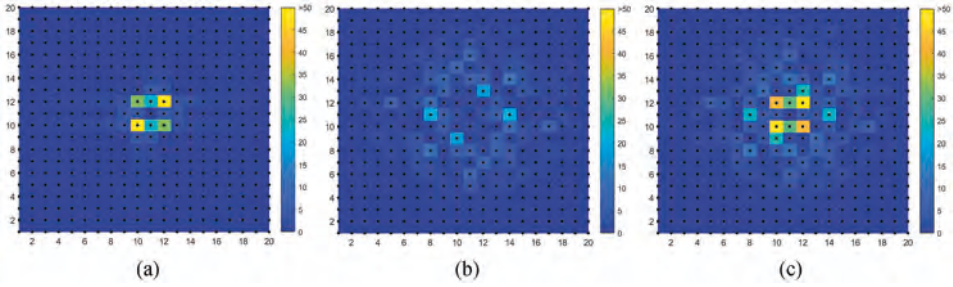

Figure 10. The power spectral densities corresponding to the three surfaces in Figure 2. The colors represent the average power (square of amplitude/m*n) of each wave. High values indicate strong spatial dependence.

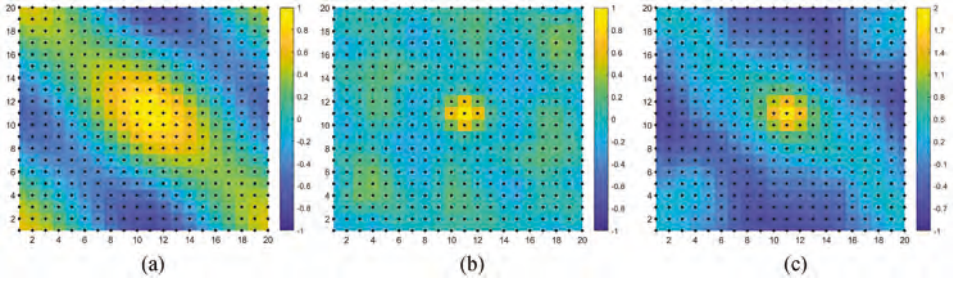

Figure 11. Inverse discrete Fourier transform of the power spectral densities corresponding to the three surfaces in Figure 2.

Figure 11 describes the inverse discrete Fourier transforms of the corresponding power spectral densities which also provide a means of estimating $R_X(i,j)$ according to equation (12):

$$R_X(i,\ j) = F^{-1}(S_X(p,\ q)) \tag{12}$$

where

$$F^{-1}(X) = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\omega_m^{-(i-1)p}\omega_n^{-(j-1)q}X_{i,\ j} \tag{13}$$

so that Figures 9 and 11 are equivalent. Both Figures 9 and 11 are shifted so that $C_X(0,\ 0)$ is at the center of the plot because this represents the correlation of a value with itself and so is always 1. The value of each pixel in Figures 9 and 11 thus expresses the correlation between the two points of the process X whose relative positions are described by the distance and direction of the pixel to the center of the plot. The sign of the values in Figure 9 corresponds to positive/negative dependence with larger absolute values indicating higher dependence. These three figures show how the dependence in the original surfaces varies by distance and direction. For example, Figure 9(a) depicts a surface in which spatial dependence varies more slowly than in the surfaces represented in Figure 9 (b) and 9(c) and also decreases more slowly in a northwest to southeast direction than in a northeast to southwest direction.

In Figure 10 the power spectral density is indicated by color with yellow pixels indicating large amplitude waves and blue pixels indicating small amplitude waves. The position of each pixel describes the frequency of the wave with pixels closer to the center indicating lower frequency waves. The pixel at (10, 10) in Figure 10(a), for example, corresponds to a large amplitude (yellow), low frequency (close to the center), wave and represents A main component of spatial dependence from the northwest to the southeast (orthogonal to the direction of the vector linking the pixel to the center of the figure). Similarly, the light blue pixel at (12, 10) corresponds to a medium amplitude, low frequency, wave in a northeasterly direction, suggesting a more gentle component of spatial dependence running northeast to southwest. These two main components of spatial dependence, a strong north-west to south-east dependenceand weaker north-east to south-west dependence are the main features of Figure 9(a) and 2(a).

In Figure 10(b) we see a large number of medium-amplitude waves suggesting that the pattern of spatial dependence in Figure 2(b) varies more rapidly over space but is generally weaker than in Figure 2(a). In Figure 10(c) we see a large amplitude wave at low frequency, similar to Figure 10(a), along with some medium amplitude waves at moderate frequency indicating that the original surface in Figure 2(c) has both strong dependence which varies slowly over space and weak dependence that varies rapidly over space. Figure 10(c) thus identifies that the distribution in Figure 2(c) results from a combination of two processes with different patterns of spatial dependence.

We now revisit the three distributions in Figure 5 and in particular the distribution in Figure 5(c) which results from a combination of processes which we were unable to identify with the spatial variogram. The power spectral densities corresponding to the three surfaces are shown in Figure 12 where Figure 12(a) indicates a spatial distribution

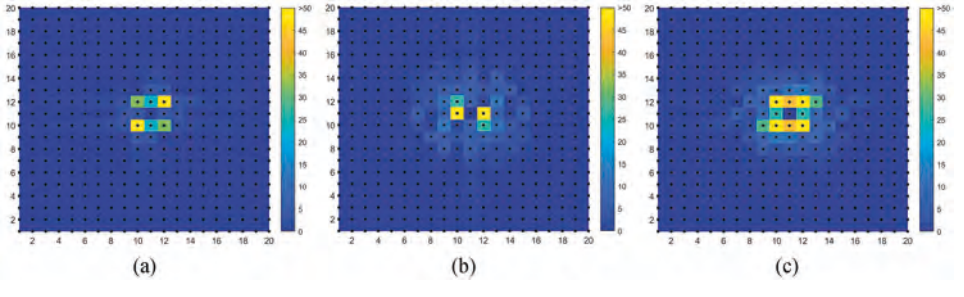

Figure 12. The power spectral densities corresponding to the three surfaces in Figure 5.

with strong dependence varying slowly from the north-west to the south-east. Figure 12 (b), in contrast, is indicative of a distribution with strong dependence in a north-south direction mixed with a confused set of weaker dependencies in various other directions varying more rapidly over space. Figure 12(c) indicates a mixture of these two distributions.

Consequently, power spectral density analysis provides an alternative means of investigating multi-scale spatial dependence. The relative amplitude represents the strength of the dependence, the frequency represents the variation in the dependence across space, and the direction of the wave represents the direction of the dependence. Furthermore, half the wave period, which is determined by its frequency, can be thought of as the 'scale' of the dependence. Half the wave period is the maximum distance where the contribution of the wave to the autocorrelation function is all positive/negative in a certain direction. If a high-amplitude wave dominates the autocorrelation function, the half period of the wave is double the distance from the origin to where the autocorrelation function first drops to 0 in a given direction. If several high amplitude waves jointly determine the autocorrelation function, then the dependence is multi-scale. Each wave represents a different degree of dependence and since different waves have different directions, anisotropy in dependence can also be reflected by the power spectral density map.

The linkages between the power spectral densities and the degree and pattern of spatial dependencies exhibited by a distribution are derived from equation (12). However, another definition of the power spectral density shown in equation (7) links power spectral density to the processes that produced the data. A very simple surface, such as a pure 2D sinusoid, can be represented by a single wave; more complex surfaces need to be represented by more waves. Complex surfaces can be thought of as having a relatively small set of controlling high amplitude waves that describe the main components of spatial dependence in the data and a large number of low amplitude waves which describe local details and some of which can be treated as the trivial components of spatial dependence. Consequently, even complex surfaces can be represented accurately by a relatively small number of waves and the number of waves that can describe a surface accurately indicates the complexity of the surface (Horn *et al*. 1986, Gonzalez and Woods 2008, Solomon and Breckon 2011, Sonka *et al*. 2014). Further, we can compare the particular waves that describe two or more different surfaces to indicate any processes which are common to these surfaces. We now describe an example of this.

4. An example of the discrete Fourier transform applied to a 2-D surface

Figure 2 displays three different surfaces: Figure 2(a) and 2(b) describe surfaces derived from two independent spatial processes while Figure 2(c) describes a surface obtained from a mixture of these two processes. These three surfaces are repeated in row 1 of Figure 13. To demonstrate the application of a discrete Fourier transform to these surfaces, we first calculate the set of waves that represent these surfaces as described by equation (5) and then order these waves in terms of their amplitudes, highest to

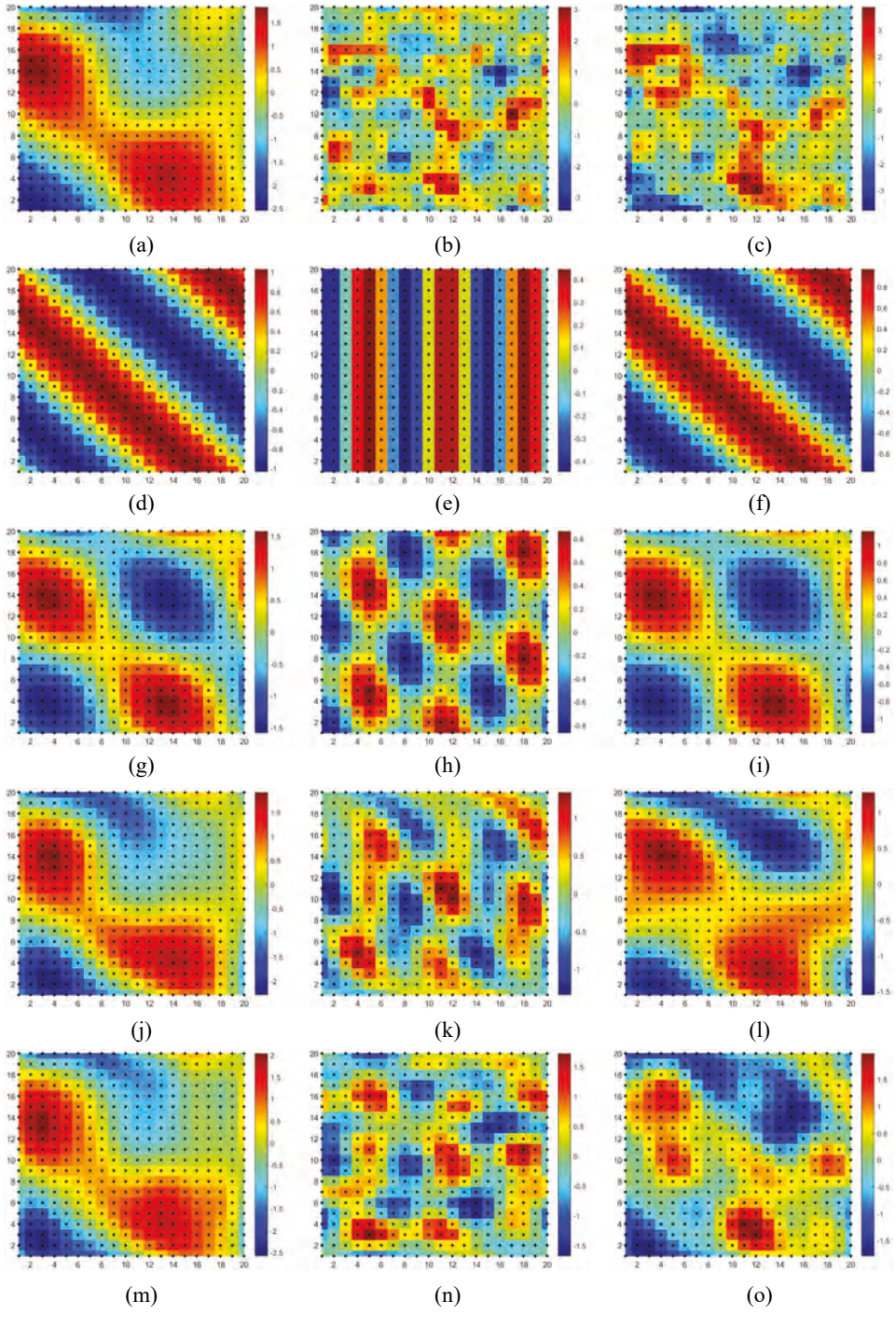

Figure 13. Surfaces constructed using various numbers of waves derived from the power spectral densities of Figure 2(a), 2(b) and 2(c). The first row contains the original three surfaces. The second row contains surfaces derived from the highest amplitude wave; the third row surfaces are generated from the highest two amplitude waves; the fourth row from the highest four amplitude waves; and the fifth row from the highest eight amplitude waves.

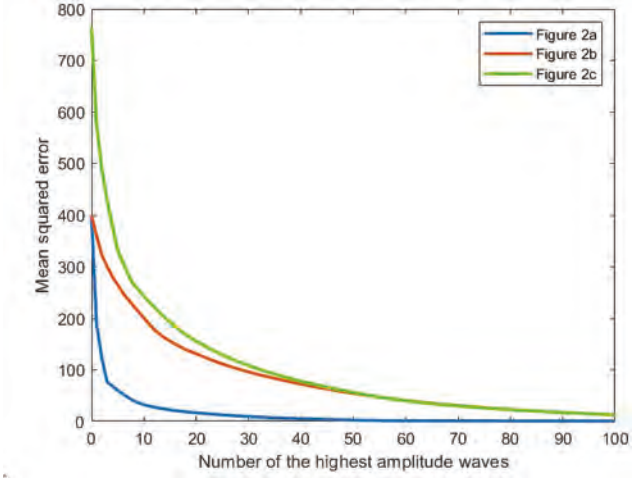

Figure 14. Mean squared error by the number of the highest amplitude waves related to Figure 13.

lowest. In Figure 13 we describe how each of the three surfaces can be represented by the highest amplitude wave (second row), the highest two amplitude waves (third row), the highest four amplitude waves (fourth row) and the highest eight amplitude waves (fifth row).

The results in Figure 13 indicate that the original surfaces in Figure 2 cannot be represented accurately by either one or two waves although they do indicate that Figure 2(a) and 2(c) shares a common structure whereas the surface in Figure 2(b) is different. Recall that Figure 2(a) has strong regional spatial dependence, Figure 2(b) has weaker and more local spatial dependence and the distribution in Figure 2(c) is a composite of these two. The surface in Figure 2(a) is summarized accurately with four waves but those in Figure 2(b) and 3(c), having greater spatial variation, require at least eight waves. Figure 14 shows the mean squared error when each of the original surfaces in Figure 2 is represented by a surface generated with a limited number of the by the number of the highest amplitude waves related to Figure 13. The mean square error decreases very quickly when the first few waves with the largest amplitude are added, and then the error decreases gradually as the number of waves increases.

5. A new multiscale measure of spatial dependence based on a discrete Fourier transform

According to equation (12), the inverse Fourier transform of the power spectral density, $F^{-1}[S_X(p,q)]$, can be used as a multiscale estimate of the spatial dependence exhibited by a surface of values. However, if we include all the waves in the power spectral density, as was done to construct Figure 11, the estimate will have a large variance and be subject to overfitting. To overcome this, only a limited number of the highest amplitude waves from the power spectral density need be included in the estimate with all other waves to zero. The highest amplitude waves describe the main components of spatial dependence in the

data so the loss of the smaller amplitude waves in the estimation of the spatial dependence only ignores some local details and trivial components of spatial dependence. Here we define $D_X(K)$ as the power spectral density containing only the top K amplitude waves from the original power spectral density $S_X(p,q)$. We then define $\Omega(K) = F^{-1}[D_X(K)]$as a multiscale estimate of the spatial dependence in the original surface using K waves. For example, $\Omega(K)$ with K =1, K =4, K =50, and K =100 for the three surfaces in Figure 2 are shown in Figure 15. Again, the values in Figure 15 are shifted so that $\Omega(K)$ (0,0) is at the center of the plot.

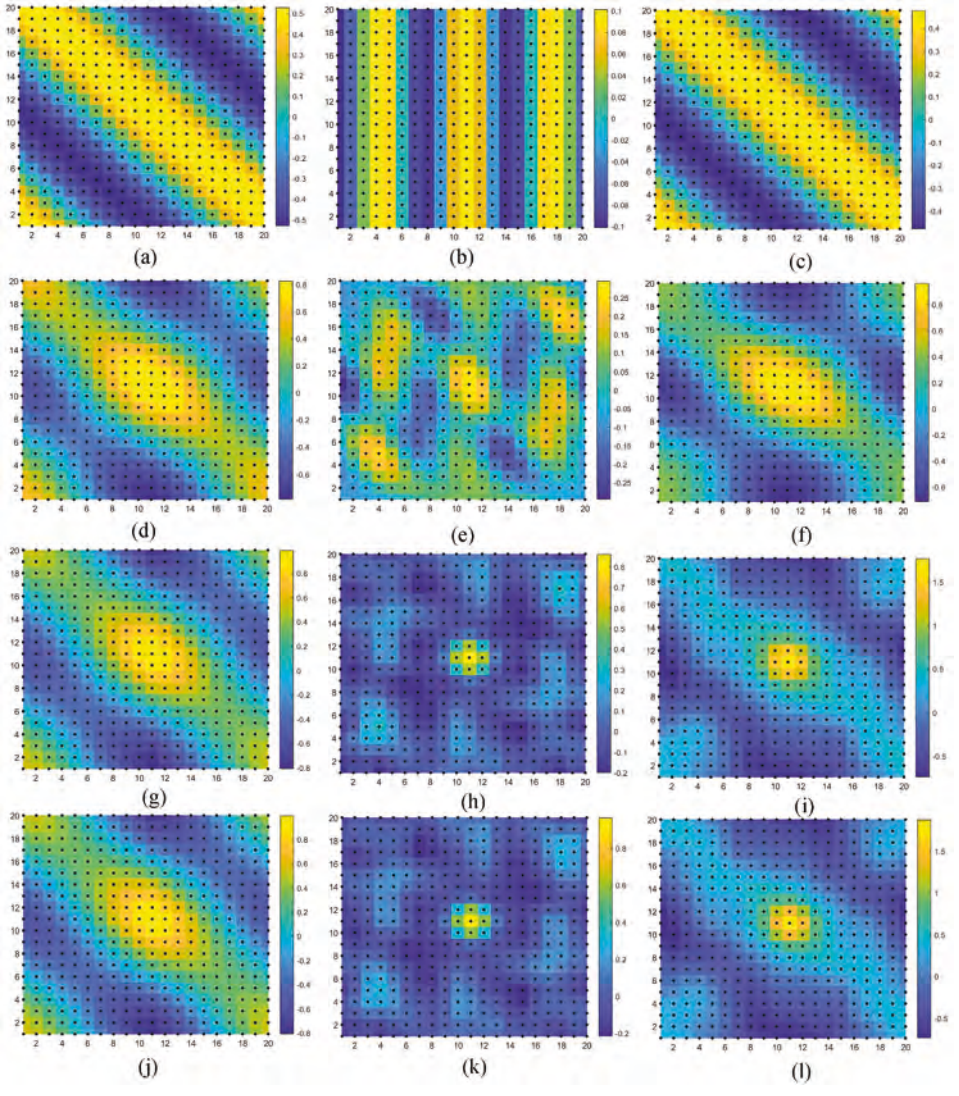

Figure 15. $\Omega(K)$ with K =1 (first row), 4 (second row), 50 (third row), 100 (fourth row) corresponding to the three surfaces in Figure 2.

The surfaces in Figure 15 are multiscale representations of the spatial dependence exhibited in the original three surfaces in Figure 2. To compute a single-scale dependence measure from one of these surfaces, in order to compare with traditional single-scale measures such as *Moran's I*, it is necessary to compare the values in one cell with the values in neighboring cells. For example, consider calculating *Moran's I* with a spatial weights matrix in which the weight is 1 if two cells share a common edge or point and 0 otherwise, the 'queen's case'. To produce an equivalent value to that of *Moran's I* calculated in this way, we use the same 'queen's case' definition of neighboring cells in Figure 15 and define $\delta_8(K)$ as the average value of the 8 adjacent locations around the center of the shifted Ω(K) map. To compare $\delta_8(K)$ and *Moran's I* as equivalent single-scale measures of spatial dependence, we use a mean-stationary Gaussian random field to generate 100 surfaces of data each with a unique dependence parameter *h* varying from 0.1 to 10 in steps of 0.1. For each of these surfaces we calculate *Moran's I* (queen's case) and $\delta_8(K)$ with $K = 2$, $K = 4$, $K = 8$, and $K = 16$ and depict the relationships between these values in Figure 16. We can produce a value of $\delta(K)$ corresponding to any definition of a spatial weights matrix used to compute *Moran's I* by simply applying the same weights matrix to the way $\delta(K)$ is derived.

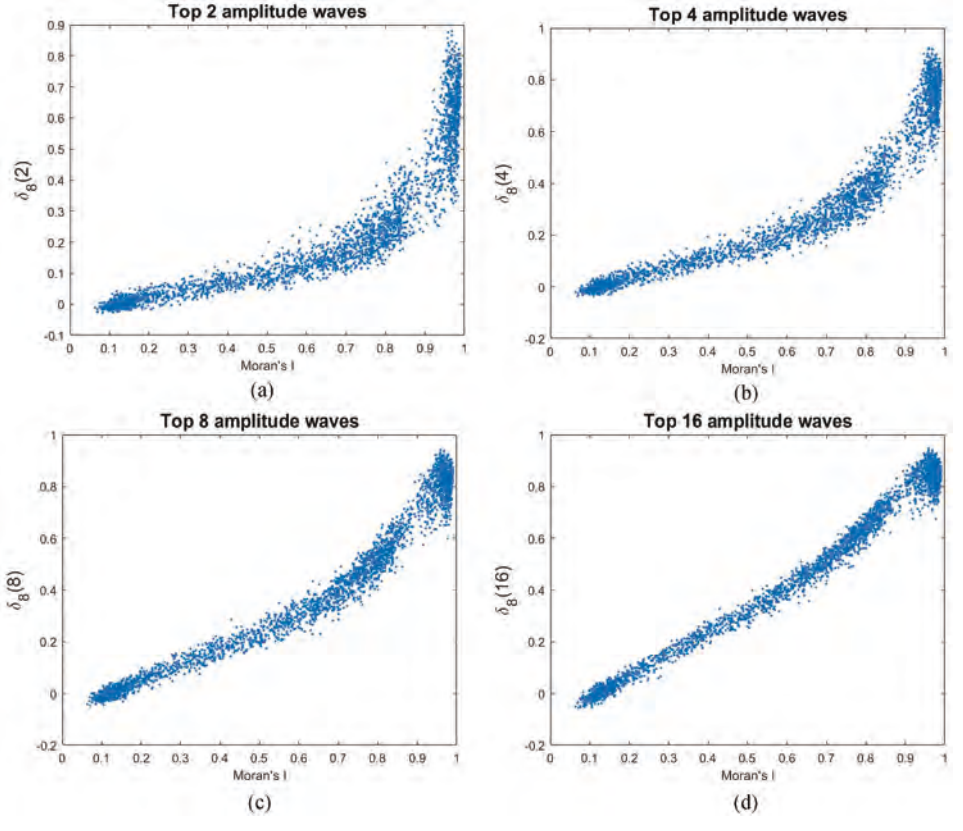

Figure 16. Scatterplots of $\delta_8(K)$ $K = 2,4,8,16$ against *Moran's I* for 100 distributions with different degrees of spatial dependence.

The scatterplots show that *Moran's I* and $\delta_8(K)$ are highly correlated even when *K* is small, supporting the proposition that Ω(K) provides an alternative, but multiscale, measure of spatial dependence. The larger is *K*, obviously, the closer is the relationship between *Moran's I* and $\delta_8(K)$ but it is unnecessary to include all the waves in the power spectral density to represent spatial dependence. The advantages of $\Omega(K)$ over *Moran's I* are that: (i) the former is a multiscale measure of spatial dependence; and (ii) It is unnecessary to *a priori* decide on a specific form for the spatial weights matrix (how many neighbors to use) which can have a critical impact on the value of *I* computed. However, it should be recognized that *Ω(K)* is a more complex measure than Moran's I and to be consistent, it would be fairer to compare *Ω(K)* to *I(K)* although the latter statistic has not yet been developed.

6. Inference about multiscale spatial dependence

To this point we have demonstrated that multiscale spatial dependence can be described by a surface of values we term *Ω(K)* as described in Figure 15. We have further shown that this surface can be used to obtain a single scale measure of spatial dependence, by using a spatial weights matrix applied to the values of *Ω(K)* and that this value would be equivalent to a value of *Moran's I* computed from the original surface using the same spatial weights matrix. It now remains to be demonstrated how we can declare if there is any significant multiscale spatial dependence in the original surface using *Ω(K)*.

Since the strength of any spatial dependence in the spatial domain is indicated by the amplitude of a wave in the Fourier domain, the statistical significance of an amplitude's magnitude over that expected by noise can be used to determine the presence of significant spatial dependence. To demonstrate this, we employ a Monte Carlo simulation procedure to test the amplitudes of a set of waves derived in the Fourier transform of a spatial data set. The null hypothesis H_0 is that there is no spatial dependence in terms of whatever processes produced the observed data – any observed level of spatial dependence is created by chance. We calculate the power spectral density related to the spatial data set and sort the waves from the highest amplitude to the lowest amplitude. Then we use a spatial random permutation on the data and repeat the previous step to obtain the sorted amplitude waves from this randomized distribution. We set the spatial position of the realization randomly, calculate the corresponding power spectral density and again sort the waves according to amplitude, highest to lowest. We repeat this randomization *M* times to obtain a distribution of the sorted amplitudes which yields *M* values of the amplitude for each order of amplitude. That is, we have *M* values of the amplitudes of the highest amplitude wave for each distribution, *M* values of the amplitude of the second highest amplitude waves and so forth. For each level of amplitude, we then compare the position of the amplitude for the observed distribution in a ranked list of all the amplitudes and the proportion of values lying above that of the observed distribution is the p-value associated with this wave. The *p-values* of all the waves corresponding to the three surfaces in Figure 10, which are the three power spectral densities related to Figure 2(a-c), are shown in Figure 17.

In Figure 17(a), there are three waves with *p-values* below 0.05 indicating that the spatial dependence present in the distribution in Figure 2(a) can be represented by these three waves; the rest are irrelevant. In Figure 17(b), there are 15 significant

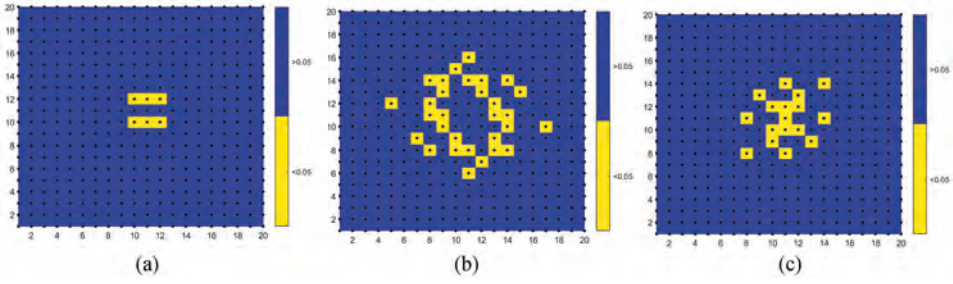

Figure 17. *P-values* corresponding to the three power spectral density surfaces in .Figure 10, which, in turn, relate to the three spatial distributions in Figure 2.

waves and in Figure 17(c), there are 8 significant waves. Since waves in opposite directions depict the same wave, the power spectral densities in Figure 10 and their inference in Figure 17 are centrally symmetric. The yellow pixels in Figure 17 indicate the corresponding pixels in Figure 10 that are significant: any spatial dependence in the original surface is mainly determined by these waves. The dark blue pixels indicate the corresponding pixels in Figure 10 which are the trivial components of spatial dependence and which do not contribute any significant spatial dependence in the original surface. In Figure 17(a), for example, there are six yellow pixels which represent three significant waves with *p-values* below 0.05 indicating that the spatial dependence present in Figure 2(a) can be represented by these three waves; the rest are irrelevant. In Figure 17(b) there are 15 significant waves and in Figure 17(c) there are 8 significant waves. The more complex the dependence structure in the original spatial distribution, the more significant waves will be identified through the analysis of the power density spectra. Identification of the significant waves in this manner allows us to discard the remaining waves and represent the multiscale spatial dependence structure in a distribution with a small number of waves which reduces the parameterization of the system and eliminates the trivial components of spatial dependence from the analysis. Of course, if the spatial dependence in the original surface was single scale, only one wave would be identified as significant in the power spectral density analysis.

7. Constructing an objective, multiscale spatial weights matrix

Spatial weights matrices are a foundation of many types of spatial analysis, such as calculating spatial dependence, spatial interpolation, cluster analysis, and spatial regression models (Anselin 1988, Aldstadt and Getis 2006, Lu and Wong 2008, Getis 2009). The classical derivation of a spatial weights matrix is based on a subjective definition of either discrete neighbors or a rate of weighting decay as distance increases so that the results of any spatial analytical technique that uses a spatial weights matrix must also contain this same degree of subjectivity. Consequently, since a single spatial weights matrix can only reflect spatial relationships at a specific spatial scale, any results from the use of such a spatial weights matrix have limited application. For example, if a statistic used to measure spatial dependence containing a traditional spatial weights matrix indicates no

significant spatial dependence, this does not indicate that there is no spatial dependence in the data, only that there is none at the particular spatial scale defined by the spatial weights matrix.

Consequently, in this section we propose a method using the power spectral density to construct a spatial weights matrix which yields the maximum degree of spatial dependence that can be calculated from the data at all spatial scales. In essence, the elements of this new spatial weights matrix define the degree of connection between each pair of cells that maximizes the overall spatial dependence between the data values in those cells. Such a spatial weights matrix has the advantages that it is objective given that it always yields the maximum degree of spatial dependence that can be found at any spatial scale, it is scale-free, and it can identify spatial dependence in complex distributions where other multiscale techniques fail to identify any spatial dependence.

To demonstrate the derivation of this new spatial weights matrix, consider the complex spatial distribution shown in Figure 18(a). Figure 18(b) shows the multi-scale *Moran's I* plot for this distribution which indicates that there is no significant spatial dependence in this distribution at any spatial scale: *Moran's I* is always very low and the associated *p-values* are always in excess of 0.4. This is also the inference that would be drawn if we were to use discrete weights such as the 'queen's case' and 'rook's case' definitions of neighbours.

Now, consider the power spectral density of the data in Figure 18(a) which is shown in Figure 19(a) along with the associated *p-values* of each wave in Figure 19(b). If every wave in the power spectral density were insignificant, there would be no spatial dependence in the data in Figure 18(a) at any spatial scale. In this case, five waves are significant so there is some significant spatial dependence in the data. We now want to uncover the spatial scale of this dependence through calculating a spatial weights matrix that maximizes whatever dependence there is in the data.

To do this, the five significant waves are retained and all other waves are set to 0 to obtain a simplified power spectral density, $D_X(K), K = 5$, from which a spatial weights matrix, each element of which is w_{ij}, can be derived by using

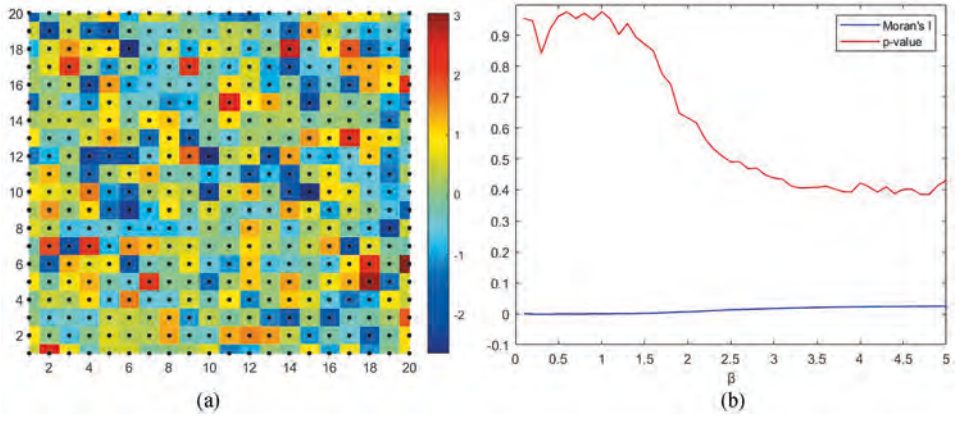

Figure 18. (a) A spatial data distribution and (b) *Moran's I* and associated β values calculated at different spatial scales for this distribution.

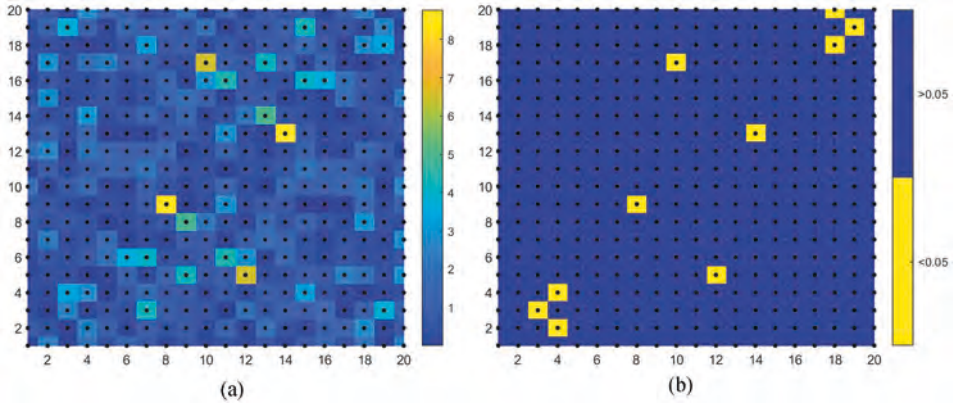

Figure 19. (a) The power spectral density for the distribution in Figure 18(a); (b) The corresponding for each wave.

$$w_{ij} = \begin{cases} \Omega(K)\left(xc_j - xc_i + 1,\ yc_j - yc_i + 1\right) - min(\Omega(K)), i \neq j \\ 0, i = j \end{cases} \tag{13}$$

where (xc_i, yc_i) is the coordinate of i. As each element in $\Omega(K)$ represents the correlation between the data at two locations, the set of weights derived from equation (13) then allows the identification of any significant spatial dependence in the data at any spatial scale. Because the values in $\Omega(K)$ can be negative, we subtract the minimum value in $\Omega(K)$ to ensure all the spatial weights are non-negative. The diagonal of the weights matrix is defined as zero. Following standard procedure with spatial weights matrices, we use row normalization on the matrix to make the sum of each row equal 1. This multiscale spatial weights matrix is shown in Figure 20(a) alongside an equivalent single-scale weights matrix derived from an inverse distance-squared weighting procedure.

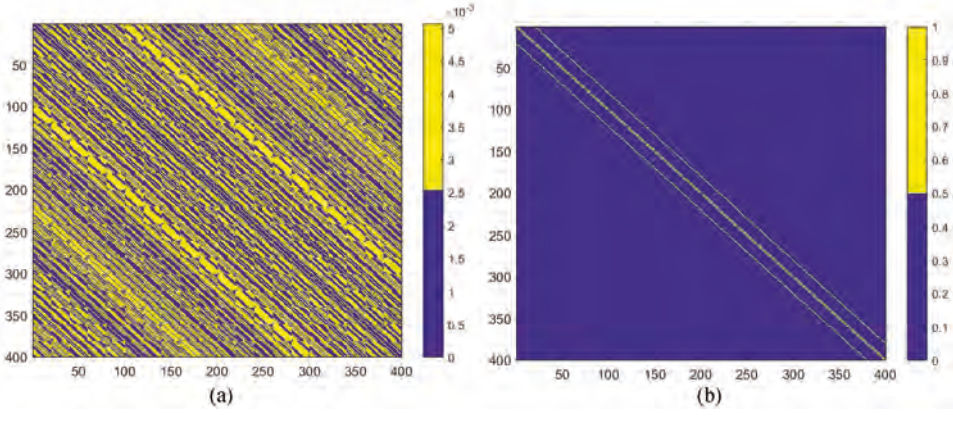

Figure 20. (a) The weights matrix from equation (13); (b) Weights from an inverse distance squared procedure.

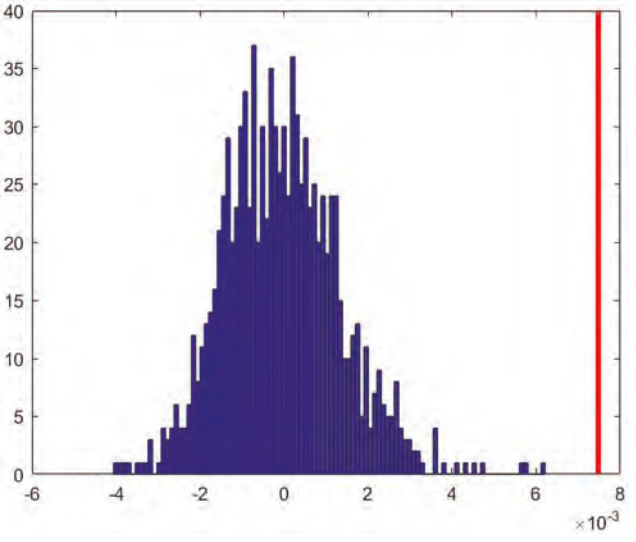

Figure 21. Monte Carlo simulated distribution of *Moran's I*. The red line is the *Moran's I* based on the data in Figure 19(a).

Both matrices have dimensions 400 × 400 showing the weights between every pixel and the remaining 399. In Figure 20(b) the weights are relatively large for neighboring east-west pixels (the center diagonal line) and are also relatively largely for pixels numbered 20 above and 20 below as these are neighbors to the north and south (the other two diagonal lines). All other weights are close to zero because the inverse distance square weighting only assigns significant weights to pixels in very close proximity to each other. By contrast, the weights derived from equation (13) shown in Figure 20(a) are much less differentiated across space although the impact of neighboring cells can still be seen from the diagonal lines of similar color running northeast to southwest across the map. This new set of spatial weights also depicts multi-scale dependencies so that higher weights can be found between pixels not in close proximity to each other.

The value of *Moran's I* using this new spatial weights matrix is .0075. To assess the significance of this value we derive an experimental distribution for *I* by rearranging the data in Figure 18(a) randomly over space and re-calculating *I* each time. This distribution is shown in Figure 21 where the observed value of *I* is shown as a red line to the right of the distribution indicating that there is significant spatial dependence in the original data that was not identified by the multiscale *Moran's I* plot.

8. Summary and limitations

The measurement of spatial dependence is a common operation in the analysis of spatial data, either as a means of assessing the degree to which similar data values are clustered in space or as a means of detecting spatially autocorrelated residuals in a regression analysis. Both are critical operations that have important ramifications for other types of spatial analysis If significant spatial dependence is detected within a data set, for example, this would necessitate a recalculation of the effective sample size in any analysis of those

data; if significant spatial dependence is detected within set of residuals from a regression model, the inference about the parameter estimates from that model would be suspect and some type of spatial regression model may need to be employed. However, the problem with any measure of spatial dependence based on traditional spatial weights matrices (be they discrete or continuous) is that it can only detect spatial dependence at a single spatial scale. Consequently, significant spatial dependence might go undetected if such statistics are employed.

This paper describes this problem with a simulated data set and then highlights the limitation of two existing multi-scale measures of spatial dependence – a multi-scale *Moran's I* and a spatial variogram. It then describes a novel approach to building a multi-scale spatial weights matrix based on a discrete Fourier transform. This new weights matrix is shown to be able to detect spatial dependence across any spatial scale and can detect significant dependence which other measures cannot. The approach is, however, not without its own limitations. The discrete Fourier transform approach to analysing spatial data, which involves viewing the resulting data as the product of a set of interacting waves, is arguably far less intuitive than the interpretation of spatial data as the manifestations of processes at discrete points in space. Also, the technique is limited to views of spatial data as distributed across regular-spaced grids; irregular distributions of data would need to be converted to rasters in order to apply the Fourier transform. That said, the technique is able to identify spatial dependence at whatever spatial scale it manifests itself and can also identify spatial dependence that other multiscale methods cannot. Given the prevalence of *a priori* defined spatial weights matrices as essential ingredients not only in measuring spatial dependence but also in many types of spatial analysis such as spatial interpolation, cluster analysis and spatial regression modeling, it should be of broad concern that such matrices are limited to a single scale of analysis. The replacement of single-scale weighting matrices by multiscale spatial weights matrices could therefore bring new insights into large areas of spatial analysis.

Notes

1. Note that although here we use *Moran's I* as our measure of spatial dependence, our findings and comments would apply equally to other measures of spatial dependence such as Geory's C which also depends on an *a priori* defined spatial weighting matrix.
2. Here, in accordance with the literature, we use 'autocorrelation' as a measure of spatial dependence.

Acknowledgements

The first author would like to acknowledge the financial support of the award of grants from the US National Science Foundation (1841403, 1758786, and 2117455), the China Data Institute, and the Future Data Lab. The second author would like to acknowledge the financial support of grant 2117455 from the National Science Foundation.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the NSF [1758786,1841403,2117455]; China Data Institute [NA]; Future Data Lab [NA].

Notes on contributors

Hanchen Yu is a postdoctoral researcher in Center for Geographic Analysis at Harvard. His research interests include spatial analysis, spatial econometrics and spatial statistics. He contributed to the idea, study design and methodology, and manuscript writing of this paper.

A. Stewart Fotheringham is Regents' Professor of Computational Spatial Science in the School of Geographical Sciences and Urban Planning at Arizona State University. His research interests are in the analysis of spatial data using statistical, mathematical and computational methods. He contributed to the study design, interpretation of the results, and the writing of this paper.

Data and codes availability statement

The data and codes that support the findings of this study are available with the identifier(s) at the link https://www.doi.org/10.6084/m9.figshare.13574468.

References

Aldstadt, J. and Getis, A., 2006. Using AMOEBA to create a spatial weights matrix and identify spatial clusters. *Geographical Analysis*, 38 (4), 327–343. doi:10.1111/j.1538-4632.2006.00689.x

Anselin, L., 1988. *Spatial econometrics: methods and models*. Vol. 4. Dordrecht: Springer Science & Business Media.

Anselin, L., 1995. Local indicators of spatial association—LISA. *Geographical Analysis*, 27 (2), 93–115. doi:10.1111/j.1538-4632.1995.tb00338.x

Anselin, L., 2019. A local indicator of multivariate spatial association: extending *Geary's C*. *Geographical Analysis*, 51 (2), 133–150. doi:10.1111/gean.12164

Bandyopadhyay, S., Lahiri, S.N., and Nordman, D.J., 2015. A frequency domain empirical likelihood method for irregularly spaced spatial data. *The Annals of Statistics*, 43 (2), 519–545. doi:10.1214/14-AOS1291

Bauman, D., *et al*., 2018. Optimizing the choice of a spatial weighting matrix in eigenvector-based methods. *Ecology*, 99 (10), 2159–2166. doi:10.1002/ecy.2469

Cliff, A. and Ord, K., 1972. Testing for spatial autocorrelation among regression residuals. *Geographical Analysis*, 4 (3), 267–284. doi:10.1111/j.1538-4632.1972.tb00475.x

Cressie, N., 1993. *Statistics for spatial data*. New York: John Wiley & Sons.

Deb, S., Pourahmadi, M., and Wu, W.B., 2017. An asymptotic theory for spectral analysis of random fields. *Electronic Journal of Statistics*, 11 (2), 4297–4322. doi:10.1214/17-EJS1326

Delgado, M.A. and Robinson, P.M., 2015. Non-nested testing of spatial correlation. *Journal of Econometrics*, 187 (1), 385–401. doi:10.1016/j.jeconom.2015.02.044

Fauchald, P., Erikstad, K.E., and Skarsfjord, H., 2000. Scale-dependent predator–prey interactions: the hierarchical spatial distribution of seabirds and prey. *Ecology*, 81 (3), 773–783.

Fuentes, M., 2007. Approximate likelihood for large irregularly spaced spatial data. *Journal of the American Statistical Association*, 102 (477), 321–331. doi:10.1198/016214506000000852

Geary, R.C., 1954. The contiguity ratio and statistical mapping. *The Incorporated Statistician*, 5 (3), 115–146. doi:10.2307/2986645

Getis, A., 2009. Spatial weights matrices. *Geographical Analysis*, 41 (4), 404–410. doi:10.1111/j.1538-4632.2009.00768.x

Getis, A. and Aldstadt, J., 2004. Constructing the spatial weights matrix using a local statistic. *Geographical Analysis*, 36 (2), 90–104. doi:10.1111/j.1538-4632.2004.tb01127.x

Getis, A. and Ord, J.K., 1992. The analysis of spatial association by use of distance statistics. *Geographical Analysis*, 24 (3), 189–206. doi:10.1111/j.1538-4632.1992.tb00261.x

Gonzalez, R.C. and Woods, R.E., 2008. *Digital image processing*. Hoboken, NJ: Prentice Hall.

Guinness, J., 2019. Spectral density estimation for random fields via periodic embeddings. *Biometrika*, 106 (2), 267–286. doi:10.1093/biomet/asz004

Hay, G.J., *et al*., 2001. A multiscale framework for landscape analysis: object-specific analysis and upscaling. *Landscape Ecology*, 16 (6), 471–490. doi:10.1023/A:1013101931793

Horn, B., Klaus, B., and Horn, P., 1986. *Robot vision*. Cambridge, MA: MIT Press.

Kelejian, H.H. and Prucha, I.R., 2001. On the asymptotic distribution of the Moran I test statistic with applications. *Journal of Econometrics*, 104 (2), 219–257. doi:10.1016/S0304-4076(01)00064-1

Kelejian, H.H. and Robinson, D.P., 1992. Spatial autocorrelation: a new computationally simple test with an application to per capita county police expenditures. *Regional Science and Urban Economics*, 22 (3), 317–331. doi:10.1016/0166-0462(92)90032-V

Lu, G.Y. and Wong, D.W., 2008. An adaptive inverse-distance weighting spatial interpolation technique. *Computers & Geosciences*, 34 (9), 1044–1055. doi:10.1016/j.cageo.2007.07.010

Matsuda, Y. and Yajima, Y., 2009. Fourier analysis of irregularly spaced data on Rd. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71 (1), 191–217. doi:10.1111/j.1467-9868.2008.00685.x

Meisel, J.E. and Turner, M.G., 1998. Scale detection in real and artificial landscapes using semivariance analysis. *Landscape Ecology*, 13 (6), 347–362. doi:10.1023/A:1008065627847

Moran, P.A., 1950. Notes on continuous stochastic phenomena. *Biometrika*, 37 (1/2), 17–23. doi:10.1093/biomet/37.1-2.17

Oman, S.D. and Mateu, J., 2019. The latent scale covariogram: a tool for exploring the spatial dependence structure of nonnormal responses. *Journal of Computational and Graphical Statistics*, 28 (1), 127–141. doi:10.1080/10618600.2018.1482766

Ord, J.K. and Getis, A., 1995. Local spatial autocorrelation statistics: distributional issues and an application. *Geographical Analysis*, 27 (4), 286–306. doi:10.1111/j.1538-4632.1995.tb00912.x

Perraudin, N. and Vandergheynst, P., 2017. Stationary signal processing on graphs. *IEEE Transactions on Signal Processing*, 65 (13), 3462–3477. doi:10.1109/TSP.2017.2690388

Rao, S.S., 2018. Statistical inference for spatial statistics defined in the Fourier domain. *The Annals of Statistics*, 46 (2), 469–499.

Rogerson, P.A., 2011. Optimal geographic scales for local spatial statistics. *Statistical Methods in Medical Research*, 20 (2), 119–129. doi:10.1177/0962280210369039

Rogerson, P.A. and Kedron, P., 2012. Optimal weights for focused tests of clustering using the local Moran statistic. *Geographical Analysis*, 44 (2), 121–133. doi:10.1111/j.1538-4632.2012.00840.x

Solomon, C. and Breckon, T., 2011. *Fundamentals of digital image processing: a practical approach with examples in Matlab*. Hoboken, NJ: John Wiley & Sons.

Sonka, M., Hlavac, V., and Boyle, R., 2014. *Image processing, analysis, and machine vision*. Toronto: Nelson Education.

Ver Hoef, J.M., Cressie, N., and Barry, R.P., 2004. Flexible spatial models for kriging and cokriging using moving averages and the Fast Fourier Transform (FFT). *Journal of Computational and Graphical Statistics*, 13 (2), 265–282. doi:10.1198/1061860043498

Westerholt, R., *et al*., 2018. A statistical test on the local effects of spatially structured variance. *International Journal of Geographical Information Science*, 32 (3), 571–600. doi:10.1080/13658816.2017.1402914

Westerholt, R., Resch, B., and Zipf, A., 2015. A local scale-sensitive indicator of spatial autocorrelation for assessing high-and low-value clusters in multiscale datasets. *International Journal of Geographical Information Science*, 29 (5), 868–887. doi:10.1080/13658816.2014.1002499

Wiener, N., 1930. Generalized harmonic analysis. *Acta Mathematica*, 55, 117–258. doi:10.1007/BF02546511

Wu, J., *et al*., 2000. Multiscale analysis of landscape heterogeneity: scale variance and pattern metrics. *Geographic Information Sciences*, 6 (1), 6–19.

Wu, Q., *et al*., 2020. Multi-scale identification of urban landscape structure based on two-dimensional wavelet analysis: the case of metropolitan Beijing, China. *Ecological Complexity*, 43, 100832. doi:10.1016/j.ecocom.2020.100832

Yao, T. and Journel, A.G., 1998. Automatic modeling of (cross) covariance tables using fast Fourier transform. *Mathematical Geology*, 30 (6), 589–615. doi:10.1023/A:1022335100486

Zhang, N. and Zhang, H., 2011. Scale variance analysis coupled with Moran's I scalogram to identify hierarchy and characteristic scale. *International Journal of Geographical Information Science*, 25 (9), 1525–1543. doi:10.1080/13658816.2010.532134