Perspectives on Spatial Data Analysis

Bearbeitet von
Luc Anselin, Sergio J. Rey

1. Auflage 2010. Buch. xviii, 290 S. Hardcover
ISBN 978 3 642 01975 3
Format (B x L): 15,5 x 23,5 cm
Gewicht: 1350 g

Chapter 2
Spatial Interaction and Spatial Autocorrelation: A Cross-Product Approach

Arthur Getis

This Chapter was originally published in:
Getis, A. (1991) Spatial Interaction and Spatial Autocorrelation: Across-Product Approach. Environment and Planning A 23:1269-1277. Reprinted with permission of PION Limited, London

Abstract A cross-product statistic is used to demonstrate that spatial interaction models are a special case of a general model of spatial autocorrelation. A series of traditional measures of spatial autocorrelation is shown to have a cross-product form. Several interaction models are shown to have a similar form. A general spatial statistic is developed which indicates that the relationship between the two types of models is particularly strong when the focus is on measurements from a single point.

2.1 Introduction

In casual conversation one rarely makes a distinction between those elements of our environment that are associated and those that interact. It is commonly believed that if tangible or intangible variables interact they are therefore in association with one another. Spatial scientists, however, have made in the technical literature a distinction between spatial association, which implies correlation, and spatial interaction. There is among them a deep-seated view that spatial interaction implies movement of tangible entities, and that this has little to do with spatial correlation. A literature on spatial autocorrelation has arisen that is nearly devoid of references to the literature on gravity and interaction models. Only on rare occasions will a spatial scientist use the words "spatial interaction" to refer to the ideas of the spatial associationists (Haining, 1978; Ord, 1975).

In this paper, I suggest that the family of spatial interaction models is a special case of a general model of spatial autocorrelation. The goal is to bring the two modeling "camps" together into a single group whose purpose is to develop further

A. Getis
Department of Geography, San Diego State University, San Diego, CA, USA
e-mail: arthur.getis@sdsu.edu

spatial models in a general way. In recent reviews of the interaction model and spatial autocorrelation literature, such as in Haynes and Fotheringham (1984) and in Anselin (1988), respectively, there is little recognition of the contributions of the other group. There has not been a discussion that shows that the two types of models can be described in a general way by the same spatial model. In order to solidify the relationship I will present a statistic that I have developed with the assistance of Ord that can be interpreted as either an indicator of spatial autocorrelation or a measure of spatial interaction.

There have been a number of generalizations of gravity and spatial interaction models (Tobler, 1983; Wilson, 1970). The most recent contribution is by Haynes and Fotheringham (1984), who write the general model as

$$T_{ij} = f(V_i, U_j, S_{ij}),$$

where T_{ij} is the interaction (tangible or intangible) between i and j, V_i and U_j represent vectors of origin and destination attributes, respectively, and S_{ij} represents a vector of separation attributes. By introducing constraints and specifying the form of the attributes, one can produce a model for validation. The relationship between the dependent and independent variables is often constrained. In some instances emphasis is on V_i (origin-specific, production-constrained gravity models), whereas in others the U_j are most important (destination-specific, attraction-constrained), and in some models there is a balance between the two (doubly constrained models). Fotheringham (1983) adds a further general term to the system, C_j, which represents a vector of competition variables. As he implies, however, C_j is a refinement of and a more detailed specification for U_j.

The historic background that has led to the current understanding of spatial autocorrelation models is very much different from that of interaction models. Spatial autocorrelation modeling has had a shorter history. Interaction modeling has been active for over 100 years although it was in the late 1950s when there was a resurgence of interest that has lasted to the current time. The field had already been reinfused with the theoretical energy of Wilson in the early 1970s when Cliff and Ord (1973) presented their ground-breaking explication of the spatial autocorrelation problem based on the work of Moran (1948), Geary (1954), and Whittle (1954).

Since 1973 the development of spatial autocorrelation models has been slow and tedious. The literature gives no evidence that Moran's I model, the join-count model, and the Geary model have been replaced or modified. Considerable progress is clear, however, in the development of regression models that include one or more spatial autocorrelation coefficients. In related developments, spectral models and especially variograms (Kriging) are being used to estimate the nature of autocorrelation in spatial data.

The common elements of the various spatial autocorrelation models are (1) a matrix of values representing the association between locations and (2) values representing a vector of the attributes of the various locations. To my knowledge, only Hubert and his associates Golledge, Costanzo, and Gale (Hubert and Golledge, 1982; Hubert et al., 1981, 1985) have developed a general form for the association of these elements. Their cross-product statistic, Γ, is written

$$\Gamma = \sum_{i,j} W_{ij} Y_{ij}, \tag{2.1}$$

where W_{ij} are elements of a matrix of measurements of spatial proximity of places i to places j, and Y_{ij} is a measure of the association of i and j on some other dimension. A slightly different form is

$$\Gamma = \sum_{i,j} W_{ij} Y_j \tag{2.2}$$

in which the relationship between the Y_i and Y_j is implicit rather than explicit as in (2.1). In this and in all subsequent formulations where we use summation signs, i does not equal j (that is, there is no self-association or self-interaction), unless otherwise indicated. In addition, in all subsequent formulations stationarity and isotropy are assumed where required. A common choice for Y_{ij} is

$$Y_{ij} = (x_i - x_j)^2, \tag{2.3}$$

where the x are the values observed for variate X_i. Clearly Y_{ij} could be some other measure of the association between i and j. For example, Hubert et al. (1981) propose $\cos(d_i - d_j)$ where d_i and d_j are angular directions at i and j. In the following paragraphs I shall identify briefly the differentiating elements of the various spatial autocorrelation models.

2.2 Cross-Product Spatial Autocorrelation Models

In this section I give a survey of the models of spatial autocorrelation. In each case attention is on the form of the model. The purpose is to show that nearly all of the models are simply just another specification of a cross-product statistic.

2.2.1 The Join-Count Models

These models require a 0,1 attribute scale. That is, some places display the attributes (1) whereas others do not (0). The Y_{ij} of the cross-product statistic differs according to the particular model of which there are three: (1) association of places with the attribute $Y_{ij} = x_i x_j$; (2) association of places with and without the attribute $Y_{ij} = (x_i - x_j)^2$; and (3) association of places without the attribute $Y_{ij} = (1 - x_i)(1 - x_j)$. The first and the third model exhibit a multiplicative form. Each of the models is constrained by allowing only a value of one for a success and zero for a failure. The model is evaluated against the expectation of the moments of X_i (see Cliff and Ord, 1973). There are no constraints on the weight matrix although in practice researchers usually choose a one-or-zero scheme to identify spatial proximity or no

Table 2.1 A comparison of various spatial models and the cross-product statistic

Model	W_{ij}	Y_{ij}	Restrictions		Scale
			W_{ij}	Y_{ij}	
Cross-product statistics					
$\Gamma = \sum \sum W_{ij} Y_{ij}$	W_{ij}	Y_{ij}	None	None	None
$\Gamma = \sum \sum W_{ij} Y_j$	W_{ij}	Y_j	None	None	None
Spatial autocorrelation models					
Joint count					
$BB = \frac{1}{2} \sum \sum W_{ij} x_i x_j$	W_{ij}	$x_i x_j$	0/1	0/1	$\frac{1}{2}$
$BB = \frac{1}{2} \sum \sum W_{ij} (x_i - x_j)^2$	W_{ij}	$(x_i - x_j)^2$	0/1	0/1	$\frac{1}{2}$
$BB = \frac{1}{2} \sum \sum W_{ij} (1 - x_i)(1 - x_j)$	W_{ij}	$(1 - x_i)(1 - x_j)$	0/1	0/1	$\frac{1}{2}$
Moran's					
$I = \frac{n \sum \sum W_{ij} (x_i - \bar{x})(x_j - \bar{x})}{W \sum (x_i - \bar{x})^2}$	W_{ij}	$(x_i - \bar{x})(x_j - \bar{x})$	None	None	$\frac{n}{W \sum (x_i - \bar{x})^2}$
Geary's					
$c = \frac{(n-1) \sum \sum W_{ij} (x_i - x_j)^2}{2W \sum (x_i \bar{x})^2}$	W_{ij}	$(x_i - x_j)^2$	None	None	$\frac{n-1}{2W \sum (x_i - \bar{x})^2}$
Semi-variance					
$\gamma = \frac{1}{2} \sum_{i=1}^{n-h} \sum_{j=i+h}^{n} W_{ij} (x_i - x_j)^2$	W_{ij}	$(x_i - x_j)^2$	1	None	$\frac{1}{2}$
Second-order					
$K(d) = \frac{\sum \sum W_{i,j}(d) x_i x_j}{(\sum x_i)^2 - \sum x_i^2}$	$W_{ij}(d)$	$x_i x_j$	0/1	Positive	$[(\sum x_i)^2 - \sum x_i^2]^{-1}$
Getis model					
$G_i(d) = [\sum_j W_{ij}(d) x_i x_j](\sum_j x_i x_j)^{-1}$	$W_{ij}(d)$	$x_i x_j$	0/1	Positive	$(\sum_j x_i x_j)^{-1}$
Spatial interaction models					
General gravity					
$T_{ij} = k x_i^{\alpha} x_j^{\tau} W_{ij}^{-\beta}$	$W_{ij}^{-\beta}$	$x_i^{\alpha} x_j^{\tau}$	None	Positive	k
Origin-specific, production-constrained					
$T_{ij} = (x_i x_j^{\alpha} W_{ij}^{-\beta})(\sum_j x_j W_{ij}^{-\beta})^{-1}$	$W_{ij}^{-\beta}$	$x_i x_j^{\alpha}$	None	Positive	$(\sum_j x_j W_{ij}^{-\beta})^{-1}$
General spatial models					
i-to-all-j model					
$G_i = (\sum_j x_i x_j W_{ij}^{-\beta})(\sum_j x_i x_j)^{-1}$	$W_{ij}^{-\beta}$	$x_i x_j$	None	Positive	$(\sum_j x_i x_j)^{-1}$
i-to-j model					
$G_{ij} = (x_i x_j W_{ij}^{-\beta})(x_i x_j)^{-1}$	$W_{ij}^{-\beta}$	$x_i x_j$	None	Positive	$(x_i x_j)^{-1}$

Note: BB black–black joins, *BW* black–white joins, *WW* white–white joins

spatial proximity. In Table 2.1 the cross-product characteristics of the models are identified.

2.2.2 Moran's I Models

The theoretical base for these models is interval-scale observations. There are two models here, differentiated only by the procedures for the evaluation of results. Unlike the join-count models, these are essentially a Pearson product-moment correlation coefficient model altered to take into consideration the effect of a spatial weight matrix. The cross-product, Y_{ij}, is the covariance, $(x_i - \bar{x})(x_j - \bar{x})$. The weight matrix has no restrictions. As in the Pearson statistic, Moran's measurement includes a scaling factor. No doubt the popularity of the Moran statistic is because of the asymptotic normal distribution of the model as n increases (Cliff and Ord, 1973). A roughly equivalent model based on a likelihood ratio statistic is by Haining (1977).

2.2.3 Geary's c Models

The two models here are similar to Moran's models except for the way in which the cross-product attributes are written. In this case the Y_{ij} is $(x_i - x_j)^2$. This is the same as the second join-count model. The variance is a scalar, and the weight matrix is as in the Moran models. The value 1 for c implies that there is no spatial autocorrelation.

2.2.4 The Semivariance Model

The semivariance is a geostatistical measure of autocorrelation based on a lattice of evenly spaced data points. Estimation of the semivariance, $\gamma(h)$, results from the sum of multiples of the values of pairs of points that are separated by a constant spatial lag h units of distance from one another in a single direction. Because of the supposed dependence between nearby data points, as h increases one would presume that the degree of autocorrelation would decline and the variance would increase to the level of the population at large. The model gets its name from the fact that the quantity is half the expected squared difference between two values. As h increases the trend of the $\gamma(h)$ values is called a variogram, not unlike the correlogram often found in studies that use Moran's I. For Hubert's statistic the value h is the equivalent of a one-or-zero weight matrix for a specified set of pairs of points that are h distance units apart in one direction (say east to west) and the values of Y_{ij} are of the form $(x_i - x_j)^2$. The variogram can be written in cross-product form as

$$\gamma(h) = \frac{1}{2} \sum_{i=1}^{n-h} \sum_{j=i+h}^{n} W_{ij}(x_i - x_j)^2. \tag{2.4}$$

2.2.5 Second-Order Spatial Autocorrelation

In a measure of spatial autocorrelation I developed earlier (Getis, 1984) the distance between x_i and x_j, is d. The d value generates a weight matrix of ones for all pairs of points found within d of one another and gives zeroes for all other pairs of points. The result is a cumulative measure of spatial autocorrelation for each distance. The measure taken over many distances creates a cumulative correlogram. The main difference between the second-order approach and the variogram is its cumulative nature, the second-order model does not depend on a lattice of points. The model for an area of size A is given by the expression

$$K(d) = \left(A \sum_{i,j} W_{i,j} x_i x_j \right) \left[\left(\sum X_i \right)^2 - \sum X_i^2 \right]^{-1}, \tag{2.5}$$

where the elements $W_{i,j}$ of the matrix are one or zero, with a one attributed to those j within d of i, and the $Y_{i,j}$ matrix contains $x_i x_j$ pairs. The X variable has a natural origin and $x_i \geq 0$. Clearly, the cross-product statistic describes the numerator and the denominator is a scalar that describes the sum of all $x_i x_j$ pairs, revealing that the measure $K(d)$ is a proportion.

2.2.6 Spectral Analysis

Although I suspect that it is possible to squeeze a spectral view of spatial autocorrelation into a cross-product form, spectral analysis is fundamentally different from the analytical models presented above. In spectral techniques it is assumed that there is a series of frequencies making up distinct periodicities in spatial data. The mathematics for identifying the harmonics are more complicated than those embodied in cross-product analysis. Spectra result from the addition of successive harmonics of a cosine wave. Spectral analysis is an effective analytical device if one is willing to assume that spatial autocorrelation is a consequence of some sort of vibratory motion or accumulation of wave-like forces.

2.2.7 The Spatial Autoregressive Model

A first-order autoregressive model is given by

$$Y_i = \alpha + \rho \sum W_{i,j} Y_j + \epsilon_i. \tag{2.6}$$

For a spatial autoregressive interpretation ρ is the spatial autocorrelation coefficient, W_{ij} is an element of the spatial weight matrix, and ϵ is the uncorrected, normally distributed, nonspatially autocorrelated, homoscedastic error term. The $W_{i,j} Y_j$ is a spatial variable which we construct from the dependent variable itself, and the system is stationary. Thus, the model represents the spatial dependence structure of Y. This is not a model of spatial autocorrelation per se but a model of the effect of spatial autocorrelation on a dependent variable. The main difference with the models described above is that the coefficient ρ is a parameter that relates the spatial dependence form of Y with itself, whereas Moran's I, for example, is strictly a value representing the spatial autocorrelation characteristic of variable Y. In fact, the numerators of both I and ρ are the covariance.

2.2.8 A Cross-Product View of Spatial Autocorrelation

The point of the above exercise is that the numerators of the autocorrelation models are essentially cross-product statistics (see Table 2.1). The W_{ij} matrix is not constrained or, if it is, the constraint is usually because of some maximum-distance rule, contiguity, or another condition that focuses attention on a specified set of interacting locations. In Table 2.1, the values of the Y_{ij} are entered into the equation in a multiplicative way, as a squared difference, or as a covariation. All other parts of the equations define the base or scalar for the calibration of the various statistics.

Hubert et al. (1981) imply that for testing purposes scales in the formulations are unnecessary. Scales are generally included in the various measures of spatial autocorrelation in order to satisfy assumptions that allow for statistical tests on well-known probability distributions. Hubert (1977) has developed a randomization technique of matrix manipulation that allows one to make statements of statistical significance without making distributional assumptions. Thus not only have we defined a family of cross-product statistics, but if we were to follow Hubert's advice we would use the same type of evaluation procedure for every formulation of Y_{ij}.

2.3 Interaction Models

I shall write the formulas for two common gravity and interaction models:

$$T_{ij} = kP_i^y P_j^\alpha d_{ij}^{-\beta} \tag{2.7}$$

and

$$T_{ij} = A_i O_i W_j^{\alpha_i} d_{ij}^{-\beta_i}. \tag{2.8}$$

The first is the general unconstrained gravity model where the P_i and P_j represent the magnitude of the variable under study at i and j, d_{ij} is the distance separating i and j, the exponents on the P variable are sometimes used to differentiate the effect of the origin from that of the destination. The exponent on the distance value represents the friction of distance. The k is a scalar or constant of proportion.

The characteristics of interaction measures that help differentiate them from autocorrelation measures are (1) a focus on a single ij relationship; (2) the use of exponents to adjust variables; (3) constraints to draw attention to one or more of the variables. In terms of the cross-product statistic there are significant similarities between them. In Table 2.1, (2.7) is rewritten to conform to the nomenclature of the cross-product statistic. Note that no summation sign is used in (2.7) or in Table 2.1. The focus in interaction modeling is on a single association, although the derivation of the parameters usually depends on the empirical data of all associations. The point, however, is that the form of the measure is similar to measures of spatial autocorrelation. The T_{ij} is simply one value that could be used in the development of a spatial autocorrelation statistic. The elements of a W_{ij} matrix contain the values of $d_{ij}^{-\beta}$. The Y_{ij} are simply the association values between the places i and j. As in the spatial autocorrelation statistics, the Y_{ij} are defined in any of a number of ways. The various constraints placed on the values at the i places can easily be accommodated in a cross-product statistic. Thus, the exponents that are used in interaction models represent more advanced development than in autocorrelation models, but there is nothing standing in the way of the use of exponents to enhance spatial autocorrelation measures (Cliff and Ord, 1969).

Equation (2.8), the origin-specific production-constrained interaction model, has been rewritten in Table 2.1 to conform to the cross-product model. It is clear that even with the complexity characteristic of many interaction models, the general form remains that of a cross product.

2.4 A General Spatial Statistic[1]

The statistic developed below contains the elements of the cross-product statistic but instead of it being a summary measure over an entire set of data it focuses on a single point as in spatial interaction measures. As it is developed here, the translation from spatial autocorrelation to interaction is not without problems.

The statistic is given by the equation

$$G_i(d) = \Big[\sum_j W_{ij}(d)x_j\Big]\Big(\sum_j x_j\Big)^{-1}, \tag{2.9}$$

where W_{ij} is a one-or-zero spatial weight matrix with ones for all links defined as being within distance d of a given place i and all other links are zero. The variable X has a natural origin and is positive. The numerator is a cross product and the denominator is the sum of all the x other than x_i. If S is equal to $x_1 + \cdots + x_n$, it follows directly that

$$K(d) = \left[\sum_j x_i(S - x_i)G_i(d)\right]\left(S^2 - \sum_i x_i^2\right)^{-1} \tag{2.10}$$

so that $G_i(d)$ represents a partition of $K(d)$ to provide an index for the ith location.

Making use of a permutations argument and recognizing that the denominator is invariant under permutations, we can consider the statistic as

$$G_i = \left[\sum_j Q_j x_j\right]\left(\sum_j x_j\right)^{-1},$$

where $Q_j = 1$ if $W_{ij} = 1$, otherwise $Q_j = 0$. This means that $P(Q_j = 1)$ is equal to $W(n-1)^{-1}$ where $W = \sum_j W_{ij}(d)$. Then

[1] This section was developed with J K Ord.

$$E(G_i) = \left[\sum_j E(Q_j)x_j\right]\left(\sum_j x_j\right)^{-1} = W(n-1)^{-1} \tag{2.11}$$

and

$$E(G_i^2) = \left(\sum_j x_j\right)^{-2}\left[\sum_j x_j^2 E(Q_j^2) + \sum_{j,k} x_j x_k E(Q_j Q_k)\right]$$

so that $E(Q_j^2) = E(Q_j)$ as $Q_j = 0$ or 1, and $E(Q_j Q_k) = W(W-1)(n-1)^{-1}(n-2)^{-1}$ (that is, hypergeometric). This yields

$$E(G_i^2) = \left(\sum_j x_j\right)^{-2}\left\{\left[(n-1)^{-1}W\sum_j x_j^2 + \frac{W(W-1)}{(n-1)(n-2)}\left[\left(\sum_j x_j\right)^2 - \sum_j x_j^2\right]\right]\right\}$$

so

$$\begin{aligned} var\,(G_i) &= E(G_i^2) - E^2(G_i) \\ &= \left(\textstyle\sum_j x_j\right)^{-2}) \Big[(n-1)^{-1}(n-2)^{-2}W(n-1-W)\textstyle\sum_j x_j^2 \\ &\quad + (n-1)^{-2}(n-2)^{-1}W(\textstyle\sum_j x_j)^2\Big]. \end{aligned}$$

If we put $(\sum_j x_j)(n-1)^{-1} = Y_1$ and $(\sum_j x_j^2)(n-1)^{-1} - Y_1^2 = Y_2$, then

$$var\,(G_i) = \frac{W(n-1-W)}{(n-1)^2(n-2)}\frac{Y_2}{Y_1^2}. \tag{2.12}$$

In this paper we will not further discuss properties of $G_i(d)$ except to say that $G_i(d)$ is normal as $n \to \infty$ (from properties of sampling without replacement, that is, a Moran-type argument). In a subsequent paper (Getis and Ord, 1992), characteristics of $G_i(d)$ will be discussed for the case when normality cannot be assumed.

2.4.1 Further Development of the Statistic

The difficulty with the statistic shown in (2.9) is in its dependence on a one-or-zero weight or distance matrix. Further development of the statistic would allow i to equal j and the substitution of $d_{ij}^{-\beta}$ for W_{ij}. For example, the following formulation would replace (2.9):

$$G_i = \sum_j x_i x_j d_{ij}^{-\beta} \left(\sum_j x_i x_j \right)^{-1}, \qquad \beta > 0, \tag{2.13}$$

where $i = j$ is allowed. In (2.13) there is an obvious correspondence between both the cross-product statistic and the general form of the interaction model. The expected value would be based on the assumption that all x values were similar. Thus,

$$E(G_i) = \frac{1}{n} \sum_j d_{ij}^{-\beta}, \qquad \beta > 0. \tag{2.14}$$

As with $G_j(d)$, the new statistic G_i would have a value as follows: $0 \leq G_i \leq 1$. If i were not equal to j then the denominator of (2.14) would be $(n-1)$. Tests based on the statistic would answer the fundamental question: "are the association and the interaction between i and all j greater than chance would have it?"

A variation on (2.13) and (2.14) would focus on the single relationship between a single i and a single j. These equations are

$$G_{ij} = \frac{x_i x_j d_{ij}^{-\beta}}{x_i x_j}, \qquad \beta > 0 \tag{2.15}$$

and

$$E(G_{ij}) = d_{ij}^{-\beta}. \tag{2.16}$$

Equations (2.15) and (2.16) complete the merger of correlation and interaction formulations.

2.4.2 *Interpretation of $G_i(d)$*

In order to test hypotheses, for example, if all x_i are set to one, the pattern of x_j represents a condition of no spatial autocorrelation. In this case, the null hypothesis is: there is no difference (and thus no spatial autocorrelation) among the x_j within distance d of i. By substituting a one for each x_j, we find (2.9) and (2.12) become

$$E[G_i(d)] = \frac{W}{n-1} \tag{2.17}$$

and

$$E[varG_i(d)] = \frac{(n-1-W)^2}{(n-1)^2(n-2)}. \tag{2.18}$$

The estimated $G_i(d)$ is found by solving (2.19) by using the observed x_j values. If

$$Z = \frac{G_i(d) - E[G_i(d)]}{\{E[varG_i(d)]\}^{1/2}}$$

is positively or negatively greater than some specified level of significance, then positive or negative spatial autocorrelation are obtained. A large positive Z implies that large values (values above the mean x_j) are spatially associated. A large negative Z means that small x_j are spatially associated with one another.

When $G_i(d)$ represents a measure of interaction, the model is expanded from (2.9) to

$$G_i(d) = \left[\sum_j W_{ij}(d) x_i x_j \right] \left[\sum_j x_i x_j \right]^{-1}. \tag{2.19}$$

A null hypotheses might call for interaction no greater (or less) than one might expect when all x_j are equal. The expectations are as in (2.17) and (2.18). Rejection of the null hypothesis would indicate that there is greater (or less) interaction than expected.

2.5 Conclusion

In verbal terms, the key words differentiating the two types of models are *interaction* and *association*. The interaction implied in gravity models refers to the possible *movement* of elements at i to or from places j. In the spatial autocorrelation model, the *link* between i and j is a correlation in the sense of places having common or different specified characteristics. As the development of the spatial autocorrelation model has a statistical origin, one usually considers association as having positive or negative statistical significance. For interaction models, statistical significance is less important and prediction is more important. For interaction modelers, interest is in the flow between places, whether or not the flow are greater or less than those predicted by a normal random variable model. In this paper we were able to show that the cross-product statistic of Hubert et al. (1981) allows for a unification of the two types of models. This was accomplished by means of the development of a spatial autocorrelation statistic that serves as a measure of spatial interaction as well. An advantage to the approach taken here is that the way is now paved for the development of statistical tests on interaction theory.

Acknowledgements The author would like to thank James O Huff, Norbcrt Oppenheim, and J Keith Ord for their valuable advice.