Beyond Moran’s $I$: Testing for Spatial Dependence Based on the SAR Model

Hongfei Li*  Catherine A. Calder†  Noel Cressie‡

Department of Statistics
The Ohio State University
1958 Neil Avenue
Columbus, OH 43210

Department of Statistics Preprint No. 763, The Ohio State University
October 14, 2005

Abstract

The statistic known as Moran’s $I$ is widely used to test for the presence of spatial dependence in observations taken on a lattice. Under the null hypothesis that the data are independent and identically distributed normal random variates, the distribution of Moran’s $I$ is known, and hypothesis tests based on this statistic have been shown in the literature to have various optimal properties. Given its simplicity, Moran’s $I$ is also frequently used outside of the formal hypothesis testing setting in exploratory analyses of spatially referenced data. In this paper, we argue against this informal use of Moran’s $I$. We show that for data generated according to the spatial autoregressive (SAR) model, Moran’s $I$ is only a good estimator of the strength of the spatial dependence parameter when there is little spatial dependence in the data. Based on this observation, we develop an alternative to Moran’s $I$, which we call $APLE$ since it is an approximate profile likelihood estimator (APLE) of the SAR spatial dependence parameter. We show that $APLE$ can be used both as a test statistic.

*Graduate Student; hongfei@stat.ohio-state.edu
†Assistant Professor of Statistics; calder@stat.ohio-state.edu
‡Professor of Statistics; ncressie@stat.ohio-state.edu
and an estimator of the strength of spatial dependence. We include both theoretical and simulation-based mo-
tivation for using \textit{APLE} as an estimator, and we propose the \textit{APLE} scatterplot, an exploratory graphical tool
that is analogous to the Moran scatterplot. We demonstrate that the \textit{APLE} scatterplot is a better visual tool for
assessing the strength of spatial dependence in the data than the Moran scatterplot. In addition, Monte Carlo
tests based on both \textit{APLE} and Moran’s \textit{I} are introduced and compared. Finally, we include an analysis of the
well known Mercer and Hall wheat-yield data to illustrate the difference between \textit{APLE} and Moran’s \textit{I} when
they are used in exploratory analyses of spatial data.

\textbf{Key Words:} \textit{APLE} Statistic, \textit{APLE} Scatterplot, Monte Carlo Test, Moran Scatterplot, Profile Likelihood, Spa-
tial Autoregressive (SAR) Model
1 Motivation

The spatial autoregressive (SAR) model is commonly used to analyze spatial processes on a lattice. Following the notation of Ord (1975), we specify an SAR model for \( Z \equiv (Z(s_1), \ldots, Z(s_n))^t \), a vector of observations on (possibly irregular) lattice locations \( \{s_i : i = 1, \ldots, n\} \), by letting

\[
Z = \rho WZ + \epsilon. \tag{1}
\]

The matrix \( W \equiv \{w_{ij}\} \) is a known spatial-neighborhood matrix with elements \( w_{ii} = 0 \) for \( i = 1, \ldots, n \), and \( \epsilon \equiv (\epsilon(s_1), \ldots, \epsilon(s_n))^t \) is a vector of independently and identically distributed normal random variables, each with mean zero and variance \( \sigma^2 \). Since \( Z = (I - \rho W)^{-1} \epsilon \), it is clear that \( E(Z) = 0 \). In practice, data will almost always have to undergo some detrending. For the rest of the paper, we shall assume that this detrending has already taken place, and hence it is appropriate for \( Z \) to have mean 0.

In exploratory analyses of spatial data, a formal statistical model may not be explicitly assumed. However, we argue that in these situations the informal notion of spatial dependence is often implicitly based on a spatial autoregressive framework where the goal is to assess the predictive ability of neighboring values of the data. In order for informal assessments of the strength of spatial dependence to translate into this implied formal statistical modeling framework, exploratory spatial data analysis tools should be based on estimators of \( \rho \), the spatial dependence parameter in the SAR model (Equation 1). However, most estimators of \( \rho \) do not have a convenient form, and consequently, they are impractical to calculate in exploratory analysis. Therefore, instead of computing an estimate of \( \rho \), the statistic known as Moran’s \( I \), \( Z'WZ / Z'Z \), is commonly used in exploratory analyses (Moran, 1950; see also Section 2). Due to its simplicity and widespread use in exploratory analyses of spatially referenced data, it is tempting to interpret Moran’s \( I \) as an estimator of \( \rho \), although the early literature (Ord, 1975; Anselin, 1988) makes it clear that this interpretation of Moran’s \( I \) is not appropriate. Still, Moran’s \( I \) is often referred to as a coefficient of spatial autocorrelation making it tempting, in practice, to misuse it as an estimator of \( \rho \). There is clearly a need for an easy-to-compute, efficient estimator of \( \rho \) that can be used to explore the spatial-dependence structure of a process on a lattice.
In this paper, we derive a likelihood-based estimator of the spatial-dependence parameter $\rho$ in the SAR model (1). We use the notation $APLE$ below, since it is an approximate profile likelihood estimator (APLE) of $\rho$. We show by simulation that for data generated according to the SAR model, $APLE$ provides a better description of the strength of the spatial dependence in the process than Moran’s $I$, especially when the true value of $\rho$ is not close to 0. In addition to exploring the properties of $APLE$ itself, we also propose an $APLE$ scatterplot, which is analogous to the Moran scatterplot, a commonly used spatial exploratory analysis tool (Anselin, 1996). We show that $APLE$ not only gives an easy-to-compute, efficient estimator of $\rho$, it is also a better exploratory tool than Moran’s $I$.

2 Background

In this section, we provide a brief historical overview of the development of Moran’s $I$ and of previous work on the estimation of $\rho$ in the SAR model. In addition, we illustrate the consequence of incorrectly interpreting Moran’s $I$ as an estimator of $\rho$.

While it is named after P.A.P. Moran, in its current form, the statistic known as Moran’s $I$ was not developed until a full 20 years after Moran’s initial paper (Moran, 1950). Moran (1950) first proposed a test statistic to assess the degree of spatial autocorrelation between adjacent locations. In order to construct this statistic for $n$ independent variates $X_i$ ($i = 1, \ldots, n$) on lattice locations $\{s_i : i = 1, \ldots, n\}$, we define $\delta_{ij}$ to be an indicator such that $\delta_{ij} = 1$ if the $i$th and $j$th locations are adjacent, and $\delta_{ij} = 0$ otherwise. Assuming that the observations have constant mean, the obvious way to detrend is to define $Z_i = X_i - \bar{X}; i = 1, \ldots, n$. Originally (Moran, 1950), Moran’s test statistic was defined as

$$I^0 = \frac{\sum_{i=1}^n \sum_{j=1}^n \delta_{ij} Z_i Z_j}{\sum_{i=1}^n Z_i^2}.$$  \hspace{1cm} (2)

Subsequently, Cliff and Ord (1973) proposed a statistic to test for a more general form of spatial dependence in the residuals from a linear regression model. Their statistic was defined by analogy with an approach first advanced by Durbin and Watson (1950) in the context of testing for serial correlation in time series. If $Z$ is
the column vector of residuals from a linear regression model, and \( \mathbf{A} \) is a real symmetric matrix, the Durbin-Watson test statistic is \( d \equiv \mathbf{Z}'\mathbf{A}\mathbf{Z}/\mathbf{Z}'\mathbf{Z} \). A special case of this statistic is approximately equal to \( \sum_i (Z_i - Z_{i+1})^2 / \sum_i Z_i^2 \), which is \( 2(1 - I^0) \) in (2), for \( \delta_{ij} = 1 \) if \( j = i + 1 \) and 0 otherwise.

For testing whether \( \rho = 0 \) in the SAR model (1), Cliff and Ord (1981) proposed the test statistic

\[
I \equiv \frac{\mathbf{Z}'\mathbf{WZ}}{\mathbf{Z}'\mathbf{Z}},
\]

which they called Moran’s \( I \). Notice that \( \mathbf{Z}'\mathbf{WZ} = \mathbf{Z}'[(\mathbf{W} + \mathbf{W}')/2]\mathbf{Z} \), and hence Moran’s \( I \) is often defined with a symmetric matrix \( \mathbf{W} \). The statistic (3) has the same form as the Durbin-Watson statistic \( d \), but it is developed in a spatial context. Cliff and Ord informally argued that for values of \( \rho \) in the neighborhood of zero, their test coincides with the likelihood ratio test of the null hypothesis that \( \rho = 0 \) versus the alternative hypothesis that \( \rho \) is some given (nonzero) value. However, they did not consider the more general alternative hypothesis that \( \rho \neq 0 \). In addition, they showed that given certain regularity assumptions, the distribution of \( I \) is asymptotically normal. Other papers have demonstrated additional properties of Moran’s \( I \). Burridge (1980) showed that the test based on \( I \) is identical to the Lagrange multiplier test, and is therefore asymptotically equivalent to the likelihood ratio test. Furthermore, King (1981) showed \( I \) to be Locally Best Invariant (LBI) in the neighborhood of \( \rho = 0 \).

By analogy to the calculation of the exact distribution of the Durbin-Watson \( d \) statistic for serial autocorrelation of regression residuals, Tiefelsdorf and Boots (1995) calculated the exact small-sample distribution of Moran’s \( I \) statistic under the spatial independence assumption (i.e., \( \rho = 0 \) in Equation 1) by numerical integration. For \( \rho \) not necessarily equal to zero in (1), Tiefelsdorf (2000) developed the distribution of Moran’s \( I \) through distribution theory for the ratio of quadratic forms. Moreover, Tiefelsdorf (2002) showed that the saddlepoint method applied to the ratio of quadratic forms in normal random variables can be used to obtain an accurate and computationally efficient approximation to the sampling distribution of Moran’s \( I \). Even though Moran’s \( I \) is by far the most widely used statistic for testing for the absence versus presence of spatial dependence, once the null hypothesis is rejected, the interpretation of Moran’s \( I \) is not straightforward. One of the main goals of this article is to assess \( I \) as an estimator of the strength of spatial dependence.
We now review briefly previous work on the estimation of $\rho$ in the SAR model (1). Ord (1975) derived a least squares estimator of $\rho$, namely $Z'WZ/Z'W'WZ$, which is not consistent. He then suggested a modified least squares estimator, which is the solution to the following quadratic equation in $\rho$:

$$Z'(I - \rho W)'W(I - \rho W)Z = 0.$$ 

Although this estimator is consistent, Ord illustrates that its efficiency relative to the maximum likelihood estimator (MLE) declines drastically as $\rho$ increases. However, the MLE of $\rho$ can only be obtained by numerically determining the maximum value of Equation 5 (Section 3.1 below), making it impractical to calculate when performing exploratory analyses.

In contrast to the MLE of $\rho$, Moran’s $I$ is straightforward to evaluate. As a result, it is commonly used in exploratory analyses of spatial data, and it is tempting to (mis)interpret it as an estimator $\rho$. To illustrate that Moran’s $I$ is not a good estimator of $\rho$, especially when $\rho$ is not near zero, and therefore is not a good exploratory tool, we performed a simulation study. This study consisted of generating data on a $10 \times 10$ regular square lattice according to the SAR model (1) for known (or true) values of $\rho = \rho^*$. For $\rho^*$ equal to 0, 0.1, 0.5, and 0.9, we generated 5,000 data sets consisting of samples from the corresponding SAR model. For each of these data sets, we then calculated Moran’s $I$ and compared it to $\rho^*$. In this analysis, we took $\sigma^2$ to be 1, and the spatial neighborhood matrix $W$ to be row standardized and defined according to a second-order neighborhood scheme. The distribution of Moran’s $I$ evaluated for each of the simulated data sets, along with the ‘true’ $\rho^*$, is given in Figure 1. Clearly, the distribution of Moran’s $I$ is centered around $\rho^*$ for $\rho^*$ near zero. However, for the other values of $\rho^*$ (including 0.1), the distribution is shifted away from $\rho^*$. These simulations demonstrate the potential for erroneous conclusions in statistical procedures where the strength of spatial dependence is being explored through Moran’s $I$.

Figure 1 Here
3 The APLE statistic and simulation analysis

Recall the simultaneously autoregressive (SAR) model (1) for data \( Z \equiv (Z(s_1), \cdots, Z(s_n))' \) on a lattice \( \{s_i : i = 1, \cdots, n\} \). Assuming that \( I - \rho W \) is invertible, it is straightforward to see that

\[
Z \sim N(0, (I - \rho W)^{-1}(I - \rho W')^{-1}\sigma^2).
\]

Consequently, the likelihood function of \( \rho \) (and \( \sigma^2 \)) can be easily obtained (see Section 3.1 below). In this section, using the likelihood function, we derive an approximate profile likelihood estimator (APLE) of \( \rho \), which is:

\[
APLE = \frac{Z'[(W + W')/2]Z}{Z'(W'W + \lambda'\lambda/n)Z},
\]

where \( \lambda \) is the vector of eigenvalues of the spatial neighborhood matrix \( W \), as an estimator of \( \rho \). In Section 3.1, we justify using APLE as an estimator of \( \rho \).

Comparing APLE to Moran’s \( I \), both statistics can be written as the ratio of quadratic forms. Since \( Z'[(W + W')/2]Z = Z'WZ \), it is clear that both statistics have the same numerator. However, their denominators differ; the denominator of APLE has a weighted form, \( \sum_i \sum_j a_{ij}Z_iZ_j \), with weights \( \{a_{ij}\} \) given by \( [W'W + \lambda'\lambda/n] \).

3.1 Derivation of the APLE statistic

Since under the SAR model, data \( Z \) are assumed to be distributed as \( N(0, (I - \rho W)^{-1}(I - \rho W')^{-1}\sigma^2) \), the likelihood function of \( \rho \) and \( \sigma^2 \) is given by

\[
L(\rho, \sigma^2) = (2\pi\sigma^2)^{-n/2} |(I - \rho W')(I - \rho W)|^{1/2} \exp \left\{ -\frac{Z'(I - \rho W')(I - \rho W)Z}{2\sigma^2} \right\}.
\]

Maximization of (5) with respect to \( \rho \) and \( \sigma^2 \) yields the maximum likelihood estimates (MLE) of these parameters. However, a closed-form solution to this maximization problem is not available (Ord, 1975). Consequently, the MLE is usually approximated numerically using either the Newton-Raphson or Fisher’s scoring algorithm.

Rather than maximizing the likelihood function directly, we use the profile likelihood to derive APLE. Since for fixed \( \rho \), \( \hat{\sigma}^2(\rho) = Z'(I - \rho W')(I - \rho W)Z/n \), we can substitute this estimate into the likelihood function.
\( \mathcal{L}(\rho, \sigma^2) \) (5), giving us the profile likelihood function of \( \rho \). After making this substitution, the negative profile loglikelihood function of \( \rho \) is

\[
\ell_P(\rho) \equiv -2 \ln \mathcal{L}(\rho, \hat{\sigma}^2(\rho)) = n(\ln(2\pi) + 1) - 2 \ln |I - \rho W| + n \ln \left[ \frac{Z'(I - \rho W')(I - \rho W)Z}{n} \right].
\]

(6)

By minimizing (6) with respect to \( \rho \), we can obtain the maximum profile likelihood estimator of \( \rho \). However, this estimator is again not available in closed form.

In order to derive an estimator of \( \rho \) in closed form, we consider the profile likelihood estimating equation obtained by putting the first derivative of (6) with respect to \( \rho \) equal to zero. Since \( \ln |I - \rho W| = \sum_{i=1}^{n} \ln(1 - \rho \lambda_i) \), where recall that \( \{\lambda_i\} \) are the eigenvalues of \( W \), this profile likelihood estimating equation for \( \rho \) can be written as:

\[
\sum_{i=1}^{n} \frac{\lambda_i}{1 - \rho \lambda_i} - n \left( \frac{Z'[\frac{W + W'}{2} - \rho W'W]}{Z'(I - \rho W')(I - \rho W)Z} \right) = 0.
\]

(7)

Then we approximate the first term on the left-hand side of Equation 7 using a Taylor series expansion, \( \sum_{i=1}^{n} \frac{\lambda_i}{1 - \rho \lambda_i} \approx \sum_{i=1}^{n} \lambda_i^2 \rho \), since the sum of the eigenvalues of \( W \) equals zero. After discarding terms of order \( \rho^2 \) and higher, we obtain an approximation to the profile likelihood estimating equation:

\[
\sum_{i=1}^{n} \lambda_i^2 \rho Z'Z - n Z' \left( \frac{W + W'}{2} - \rho W'W \right) Z = 0.
\]

Solving this equation for \( \rho \), we obtain the approximate profile likelihood estimator for \( \rho \), \( APLE \), given by (4).

### 3.2 Comparison of \( APLE \) with Moran’s \( I \) via simulation

In this section, we compare \( APLE \) and Moran’s \( I \) using a simulation study, where the data are simulated in the same manner as described in Section 2. From Figure 2, the sampling distribution of \( APLE \) is centered around \( \rho^* \), the true value of \( \rho \), regardless of whether \( \rho^* \) is near zero or not. This is in contrast to the behavior of Moran’s \( I \); its sampling distribution is only centered around \( \rho^* \) when \( \rho^* \) is near zero. Therefore, compared to Moran’s \( I \), \( APLE \) appears to provide a much better estimate of \( \rho \), especially when the true value of \( \rho \) is not close to 0.
3.3 The APLE scatterplot

Proposing an exploratory spatial data analysis (ESDA) tool, Anselin (1996) interpreted Moran’s $I$ as a regression coefficient in a regression of $WZ$ on $Z$. This interpretation provides a way to visualize the linear association between $Z$ and $WZ$ in the form of a bivariate scatterplot of $WZ$ against $Z$. Anselin referred to this plot as the Moran scatterplot. He also pointed out that the least squares slope in this regression is equal to Moran’s $I$, although its significance (using the standard $t$ test for the linear regression) is not appropriate.

Based on the same idea, $APLE$, given by (4), can be visualized as a least squares regression coefficient. Let

$$X = (W'W + \lambda^2 \lambda I/n)^{-\frac{1}{2}} Z, \quad (8)$$

and

$$Y = (W'W + \lambda^2 \lambda I/n)^{-\frac{1}{2}} ([W + W']/2) Z. \quad (9)$$

Then, $APLE$ can be thought of as a regression coefficient in a regression of $Y$ on $X$ through the origin; that is,

$$APLE = \frac{Y'X}{X'X}.$$

We propose a visualization tool based this decomposition of $APLE$, which is analogous to the Moran scatterplot and which we call the $APLE$ scatterplot. The $APLE$ scatterplot consists of plotting points $\{(X_i, Y_i) : i = 1, \ldots, n\}$, where $X = (X_1, \ldots, X_n)'$ is given by (8) and $Y = (Y_1, \ldots, Y_n)'$ is given by (9). Superimposed on the scatterplot of $X$-$Y$ points is the regression line through the origin whose slope is given by $APLE$. To illustrate the $APLE$ scatterplot, we simulated data $Z$ from $N(0, (I - \rho^* W)^{-1}(I - \rho^* W')^{-1} \sigma^2)$, which is the SAR model (1) with $\rho = \rho^*$, on a $10 \times 10$ square lattice with $\rho^* = 0.5$ and $\sigma^2 = 1$. For this simulated dataset $Z$, Moran’s $I$ equals 0.207, while $APLE$ equals 0.483, which is much closer to $\rho^* = 0.5$. Figure 3 illustrates both the $APLE$ and Moran scatterplots corresponding to these simulated data. As we would expect,
the least squares line corresponding to APLE is virtually indistinguishable from the line with slope \( \rho^* = 0.5 \), represented by the solid line. However, there is a substantial difference between the straight line obtained from Moran’s \( I \) and the solid line obtained from the true value of \( \rho \).

Figure 3 Here

4 APLE as a test statistic

While our underlying motivation for the APLE statistic is the need for an easy-to-compute estimator of \( \rho \), APLE also can be used as a test statistic. To illustrate this use of APLE, consider testing the hypothesis \( H_0 : \rho = \rho_0 \) versus \( H_1 : \rho \neq \rho_0 \) for data generated according to the SAR model (1). We show in Appendix A that the test based on APLE can be derived from the Lagrange multiplier test statistic, up to first-order terms in a Taylor-series expansion. A related approach has been used to motivate the use of Moran’s \( I \) as a test statistic (e.g., Anselin, 1988). Moreover, a score statistic derived in the case where \( \sigma^2 \) is a nuisance parameter also can be shown to yield, up to first-order terms, the APLE statistic (see Appendix B).

Rather than comparing exact or large sample tests based on APLE and Moran’s \( I \), we compare the two test statistics within the Monte Carlo testing framework. To set up a Monte Carlo hypothesis test, we first define the null and alternative hypotheses \( H_0 : \rho = \rho_0 \) and \( H_1 : \rho \neq \rho_0 \), respectively. We let \( \{U^i_{\rho_0} : i = 1, \ldots, K\} \) denote \( K \) values of the statistic (either APLE or Moran’s \( I \)) generated by independently simulating data \( K \) times under the null hypothesis \( \rho = \rho_0 \). When simulating the data, we can assume, without loss of generality, that \( \sigma^2 = 1 \) since both APLE and Moran’s \( I \) do not depend on \( \sigma^2 \). For large \( K \), we obtain the acceptance region \( \{U_{\rho_0}([K\alpha/2]), U_{\rho_0}(1+[K(1-\alpha/2)])\} \) (e.g., Hope, 1968; Kornak et al., 2005) at the \( \alpha \) significance level, where \([K\alpha/2]\) denotes the smallest number that is larger than \( K\alpha/2 \) and \( U_{(i)} \) denotes the \( i \)th order statistic.

In order to compare the Monte Carlo tests based on APLE and Moran’s \( I \), we consider each test for a range
of values of $\rho$ under the null hypothesis. For $\rho_0 \in \{-0.9, -0.8, \cdots, 0.9\}$, we perform the Monte Carlo test of $H_0 : \rho = \rho_0$; that is, we obtain the acceptance region $\left(U_{\left(K\alpha/2\right)}, U_{\left(1+K(1-\alpha/2)\right)}\right)$ based on simulating $K = 5,000$ data sets, like the data set generated in Section 3.3, assuming that $\rho = \rho_0$. Then, for each value of $\rho_0$, we determine the upper and lower bounds of the acceptance region for both Monte Carlo tests (i.e., based on $I$ and on $APLE$) with $\alpha = 0.05$, which are illustrated in Figure 4. The solid curves in the figure are obtained by linearly interpolating the upper and lower bounds of the acceptance regions from the $APLE$-based Monte Carlo test for values for $\rho_0$ between $-1$ and $1$, and the dashed curves correspond to linear interpolations of the upper and lower bounds of the acceptance regions for the Monte Carlo tests based on Moran’s $I$ for values for $\rho_0$ between $-1$ and $1$.

Figure 4

Using Figure 4, we can approximate the acceptance region for both the $APLE$- and Moran’s $I$-based Monte Carlo tests. To illustrate, consider the null hypothesis $\rho_0 = 0.5$. The intersections of the vertical line at $\rho = 0.5$ with the solid and dashed curves indicate the acceptance regions (on the vertical axis) for both the $APLE$ and Moran’s $I$-based tests, respectively. These acceptance regions are represented by braces to the left of the vertical axis: the acceptance region for the test based on $APLE$ is represented as interval $A$ and the acceptance region for the test based on Moran’s $I$ is represented as interval $B$. Comparing these two acceptance regions, we see that the true value of $\rho$ falls inside interval $A$, but outside interval $B$. As a result, it is difficult to interpret the test based on Moran’s $I$ for $\rho_0 = 0.5$. Generally, by including the gray $45^\circ$ line on the plot, it is apparent that only for values of $\rho_0$ near zero will the gray line pass through the acceptance region based on Moran’s $I$; only for these values of $\rho$ near 0 will the acceptance region for the Moran’s $I$-based test include $\rho_0$. For other values of $\rho_0$, Moran’s $I$ falls outside the acceptance region, and therefore, the test is difficult to interpret. This undesirable property of the Moran’s $I$-based Monte Carlo test is not surprising since Moran’s $I$ is not a good estimator of $\rho$ for values of $\rho$ away from zero, as illustrated by the simulation study in Section 3.2.
By inverting the acceptance regions obtained from the two Monte Carlo tests, we can obtain confidence intervals for \( \rho \). Figure 5 illustrates this procedure. Suppose we have a statistic \( \hat{\rho} = 0 \), where \( \hat{\rho} \) can be APLE or Moran’s \( I \). If we draw a horizontal line at \( \hat{\rho} = 0 \), then the intersections of this line with the acceptance region curves (illustrated by the two squares in Figure 5) defines the corresponding \( 1 - \alpha = 0.95 \) percent confidence interval for \( \rho \). From Figure 5 we can determine that for values of \( \hat{\rho} \) near 0, both the APLE- and Moran’s \( I \)-based Monte Carlo tests provide nearly identical confidence intervals, which both include 0. However, for \( \hat{\rho} = 0.5 \), the confidence interval based on APLE includes \( \rho = 0.5 \), while the confidence interval based on Moran’s \( I \) does not include \( \rho = 0.5 \). As with the previous figure, the 45° dashed line is included in Figure 5 in order to identify the range of values of \( \rho \) where the confidence interval based on Moran’s \( I \) does not cover the true value of \( \rho \). The fact that the confidence intervals based on APLE include the true value of \( \rho \) for all values of \( \rho \) between \(-1\) and \(1\) provides further evidence that APLE is a better statistic upon which to base inference than Moran’s \( I \), whenever \( \rho \) is not close to 0.

---

**Figure 5** Here

---

5 **Illustrative example**

In this section, we test for the presence of spatial dependence using both APLE and Moran’s \( I \) in the famous wheat-yield data of Mercer and Hall (1911). These data were used by both Whittle (1954) and Besag (1974) to illustrate their models for spatial dependence. The wheat yields are from 10.82 feet × 8.50 feet plots and were collected in the summer of 1910. There were a total of 500 of these plots arranged in a lattice with 20 rows (8.50 feet) running east to west and 25 (10.82 feet) columns running north to south.

Cressie (1993) showed the presence of a north-south trend in the data. Consequently, we subtracted the column median to remove this trend and then subtracted the overall mean. We then assumed that the residuals
follow the zero mean SAR model given by (1). Based on the analysis in Cressie (1985), we chose a row-
standardized spatial neighborhood matrix $W$ with first- and second-nearest neighbors.

For these data, Moran’s $I = 0.212$ and $APLE = 0.502$. Figure 6 provides an $APLE$ scatterplot for the data. In addition, in order to display the uncertainty in the slope, a 95 percent confidence cone is included on the $APLE$ scatterplot. This confidence cone is constructed by shading the area between the lines through the origin with slopes equal to the upper and lower confidence bounds derived from the $APLE$-based Monte Carlo test described in Section 4 for a sample size of $n = 500$ and using $K = 5,000$ simulated data sets. By examining this $APLE$ scatterplot, we can visualize the estimate of $\rho$ given by $APLE$ and assess our confidence in this estimate. Given that the confidence cone does not include the horizontal line $\rho = 0$, we reject the hypothesis that $\rho = 0$.

6 Conclusion

In this paper, we examined Moran’s $I$ and the new $APLE$ statistic both in terms of testing and estimation of spatial dependence assuming an SAR model. Although Moran’s $I$ is widely recognized as an estimator of spatial dependence, we showed by simulation that it can be misleading, especially when the spatial dependence parameter $\rho$ is not near zero. To test for the strength of spatial dependence in the hypothesis $H_0 : \rho = \rho_0$, where $\rho_0$ may not equal zero, we derived a Monte Carlo test based on the statistic $APLE$. Properties of this test were obtained from simulation and shown to be superior to the corresponding test based on Moran’s $I$. Finally, our analysis of the Mercer and Hall (1911) wheat-yield data demonstrated the use of $APLE$ as an exploratory spatial data analysis tool.

In future work, we will look at the effect of keeping higher-order terms in the Taylor-series approximation that give the APLE. We will also generalize $APLE$ for other models of spatial dependence, and where covariates are included specifically in the model.
Appendix A: Lagrange Multiplier Test

Burridge (1980) showed that Moran’s $I$ is identical to the Lagrange multiplier test statistic for the test with null hypothesis $\rho = 0$ versus the alternative hypothesis $\rho \neq 0$. In this section, we show that $APLE$, up to first-order, can be derived from the Lagrange multiplier test under a general null hypothesis $\rho = \rho_0$, where $\rho_0$ may not be zero.

For the test $H_0 : \rho = \rho_0$ versus $H_1 : \rho \neq \rho_0$, the loglikelihood function of $\rho$ and $\sigma^2$ is given by

$$
\ell(\rho, \sigma^2) = \ln L(Z|\rho, \sigma^2)
= -\frac{n}{2} \ln(2\pi\sigma^2) + \ln |I - \rho W| - \frac{Z'(I - \rho W')(I - \rho W)Z}{2\sigma^2}.
$$

(10)

Under the null hypothesis, imposing the restriction that $\rho = \rho_0$ yields the restricted loglikelihood

$$
\ell_R(\rho, \sigma^2) = \ell(\rho, \sigma^2) - \lambda(\rho - \rho_0),
$$

where the Lagrange multiplier $\lambda$ is to be evaluated. The estimating equations are given by

$$
\frac{\partial \ell_R}{\partial \sigma^2} = -\frac{1}{2\sigma^2} \left\{ n - \frac{1}{\sigma^2} Z'(I - \rho W')(I - \rho W)Z \right\} = 0,
$$

(11)

$$
\frac{\partial \ell_R}{\partial \rho} = -\lambda - \sum_{i=1}^n \frac{\lambda_i}{1 - \rho \lambda_i} + \frac{1}{2\sigma^2} Z'\{(W + W') - 2\rho W'W\}Z = 0,
$$

(12)

and $\rho - \rho_0 = 0$.

(13)

Equations 11 and 13 imply that

$$
\hat{\sigma}^2 = \frac{1}{n} Z'(I - \rho_0 W')(I - \rho_0 W)Z
$$

and

$$
\rho = \rho_0.
$$

From Equations 12 and 13, the restricted MLE of $\lambda$ is

$$
\hat{\lambda} = -\sum_{i=1}^n \frac{\lambda_i}{1 - \rho_0 \lambda_i} + \frac{1}{2\sigma^2} Z'\{(W + W') - 2\rho_0 W'W\}Z,
$$

(14)

which is the Lagrange multiplier test statistic. If $\hat{\lambda}$ differs significantly from zero, this indicates that the data do not support the null hypothesis.
To estimate $\rho$ using the Lagrange multiplier framework, in Equation 14, we replace $\rho_0$ with $\rho$, set $\hat{\lambda} = 0$, and solve for $\rho$. Using a Taylor-series expansion, we derive $\sum_{i=1}^{n} \lambda_i^{2} / (1 - \rho \lambda_i) \triangleq \sum_{i=1}^{n} \lambda_i^{2} \rho$, as before. Then Equation 14 implies:

$$2\hat{\sigma}^{2} \sum_{i=1}^{n} \lambda_i^{2} \rho = \mathbf{Z}' \{(W + W') - 2\rho W'W\} \mathbf{Z}. \quad (15)$$

Substituting $\hat{\sigma}^{2}$, obtained from Equation 11, into Equation 15 and keeping only first-order terms in $\rho$, we obtain an estimator of $\rho$,

$$\frac{\mathbf{Z}'(W + W')/2 \mathbf{Z}}{\mathbf{Z}'(W'W + \lambda \lambda I/n) \mathbf{Z}},$$

which is APLE, given by (4).

### Appendix B: Score Test

In this section, we show that APLE, up to first order terms, can be derived from the score statistic. The loglikelihood function of $\rho$ and $\sigma^2$ is given in Equation 10. For the test $H_0 : \rho = \rho_0$ versus $H_1 : \rho \neq \rho_0$, $\rho$ is the parameter of interest and $\sigma$ is the nuisance parameter. In our context, the score statistic, as defined by Cox and Hinkley (1974), is

$$Q_S = Q_S(\hat{\sigma}_0) \quad (16)$$

$$= \{U_{,\rho}(\rho_0, \hat{\sigma}_0) - b.(\rho_0, \hat{\sigma}_0)U_{,\sigma}(\rho_0, \hat{\sigma}_0)\}^{2}J_{,\rho\rho}(\rho_0, \hat{\sigma}_0),$$

where $\hat{\sigma}_0$ is the MLE of $\sigma$ under $\rho = \rho_0$. In Equation 16,

$$U_{,\rho}(\rho_0, \hat{\sigma}_0) = \left. \frac{\partial \ell(\rho, \sigma)}{\partial \rho} \right|_{\rho = \rho_0, \sigma = \hat{\sigma}_0},$$

$$U_{,\sigma}(\rho_0, \hat{\sigma}_0) = \left. \frac{\partial \ell(\rho, \sigma)}{\partial \sigma} \right|_{\rho = \rho_0, \sigma = \hat{\sigma}_0},$$

and $J_{,\rho\rho}(\rho_0, \hat{\sigma}_0) = (J_{,\rho\rho} - J_{,\rho\sigma}J_{,\sigma\sigma}^{-1}J_{,\rho\sigma})^{-1}|_{\rho = \rho_0, \sigma = \hat{\sigma}_0},$

where $J.$ is the $2 \times 2$ information matrix with diagonal elements $J_{,\rho\rho}$, $J_{,\sigma\sigma}$, and off-diagonal element $J_{,\rho\sigma}$. It is clear that $U_{,\sigma}(\rho_0, \hat{\sigma}_0) = 0$; hence, in this special case, the score statistic does not depend on $b.$, a function of
both the parameter of interest under the null hypothesis and the MLE of the nuisance parameter (see page 324 of Cox and Hinkley (1974) for the exact form of $b$). Consequently, Using the score statistic (16) simplifies to:

$$Q_S = U_{\rho}(\rho_0, \hat{\sigma}_0)^2 J_{\rho\rho}(\rho_0, \hat{\sigma}_0).$$  

(17)

We know $Q_S$ is asymptotically distributed as $\chi^2_1$ (Cox and Hinkley, 1974). Therefore, for null hypothesis $H_0 : \rho = 0$ versus the alternative hypothesis $H_1 : \rho \neq 0$, we accept $H_0$ if $U_{\rho}(\rho_0, \hat{\sigma}_0^2)$ is close to 0. Using the loglikelihood of $\rho$ and $\sigma^2$ as given by Equation 10, it can be shown that

$$U_{\rho}(\rho_0, \hat{\sigma}_0^2) = -\sum_{i=1}^n \frac{\lambda_i}{1 - \rho \lambda_i} + \frac{Z' [(W + W')/2] Z - \rho Z' W W' Z}{\hat{\sigma}_0^2}. \quad (18)$$

We note in passing that upon substituting $\rho_0 = 0$ and $\hat{\sigma}_0^2 = Z' Z / n$ into Equation 18, we obtain the test statistic

$$U_{\rho}(0, Z' Z / n) = n \times \frac{Z' [(W + W')/2] Z}{Z' Z},$$

which is equivalent to Moran’s $I$ up to a constant.

For testing the general null hypothesis $H_0 : \rho = \rho_0$ versus the alternative $H_1 : \rho \neq \rho_0$, we accept $H_0$ if $U_{\rho}(\rho_0, \hat{\sigma}_0^2)$ is close to 0, where $\hat{\sigma}_0^2 = Z' (I - \rho_0 W)' (I - \rho_0 W) Z / n$. To get an estimate of $\rho$, we apply the same strategy as in Appendix A and set $U_{\rho}(\rho_0, \hat{\sigma}_0^2) = 0$. This yields the estimating equation, Equation 7, whose solution, up to first order, is the statistic APLE.
References


Figures

Figure 1: Distribution of Moran’s $I$ for simulated data sets generated from the SAR model (1) with $\rho^*$ (vertical bold line) equal to 0, 0.1, 0.5, and 0.9. The histograms are based on 5,000 simulated data sets.
Figure 2: A comparison of the sampling distribution of APLE and Moran’s I with the true value $\rho^*$ (vertical bold line) equal to 0, 0.1, 0.5 and 0.9, based on 5,000 simulations.
Figure 3: *APLE* and Moran scatterplots for simulated data generated according to an SAR model with $\rho$ equal to 0.5. The solid line is a line through the origin with slope 0.5; the dashed line is a line through the origin with slope given by the statistic (*APLE* on the left plot and $I$ on the right plot).
Figure 4: Acceptance region for $\rho$. The two solid curves represent the upper and lower bounds of the acceptance region based on $APLE$; the two dashed curves represent the upper and lower bounds of the acceptance region based on Moran’s $I$. For illustration, intervals A and B are the acceptance regions corresponding to the $APLE$- and Moran’s $I$-based Monte Carlo tests, respectively, when $\rho = 0.5$. 
Figure 5: Illustration of the procedure for obtaining confidence intervals for $\rho$ using APLE- and Moran’s $I$-based Monte Carlo tests. Confidence interval A corresponds to both types of test when $\hat{\rho} = 0$. Confidence interval B is constructed from an APLE-based Monte Carlo test with $\hat{\rho} = 0.5$. 
Figure 6: *APLE* scatterplot for the Mercer and Hall wheat-yield data. A confidence cone, derived by inverting the acceptance region of an *APLE*-based Monte Carlo test is also included; see the text for details.