

# The Predictive Power of Anisotropic Spatial Correlation Modeling in Housing Prices

Bing Zhu · Roland Füss · Nico B. Rottke

Published online: 9 October 2009  
© Springer Science + Business Media, LLC 2009

**Abstract** This paper develops a method to capture anisotropic spatial autocorrelation in the context of the simultaneous autoregressive model. Standard isotropic models assume that spatial correlation is a homogeneous function of distance. This assumption, however, is oversimplified if spatial dependence changes with direction. We thus propose a local anisotropic approach based on non-linear scale-space image processing. We illustrate the methodology by using data on single-family house transactions in Lucas County, Ohio. The empirical results suggest that the anisotropic modeling technique can reduce both in-sample and out-of-sample forecast errors. Moreover, it can easily be applied to other spatial econometric functional and kernel forms.

**Keywords** Spatial regression · Hedonic price model · Anisotropic spatial correlation · Simultaneous autoregressive model · Housing market

---

B. Zhu · N. B. Rottke  
Real Estate Management Institute (REMI), European Business School (EBS),  
International University Schloss Reichartshausen, 65201 Wiesbaden, Germany

B. Zhu  
e-mail: zhu.ebs@rem-institute.org

N. B. Rottke  
e-mail: rottke.ebs@rem-institute.org

R. Füss (✉)  
Union Investment Chair of Asset Management and Real Estate Management Institute (REMI),  
European Business School (EBS), International University Schloss Reichartshausen,  
65375 Oestrich-Winkel, Germany  
e-mail: roland.fuess@ebs.edu

## Introduction

Houses are heterogeneous goods and thus buying a house always means buying an entire range of housing characteristics. A house's fixed location means its value will tend to induce spatial autocorrelation and spatial heterogeneity. Hence, using traditional ordinary least squares (OLS)-based hedonic house price modeling will be inefficient if the assumption of uncorrelated error terms is violated. The OLS error terms can be spatially autocorrelated for three reasons. Firstly, some of a house's structural and neighborhood attributes are unobservable. Thus, some important characteristics may be omitted. Secondly, even if all or most of the information is available, some variables are difficult to quantify such as the quality of the public school system. Moreover, even after taking all available location variables into account house prices can still be correlated because of their relative proximity to each other (Valente et al. 2005). To overcome this problem we incorporate spatial dependence considerations into the traditional hedonic model.

Most approaches to modeling spatial dependence assume that the correlation structure is isotropic. Under this assumption, the dependence of pairs of observations is only a function of the distance between properties while the direction separating the properties is ignored. However, the intensity of spatial dependence often decreases differently in different directions. For example, land price gradients tend to be flatter along radial transportation routes than in other directions (Colwell and Munneke 2009). In some studies, the land price gradient is defined as a function of the direction around the center of the city under a piecewise linear function. Therefore, house prices and residuals of hedonic house price regressions will exhibit anisotropic spatial autocorrelation. They are thus functions of both the distance and the direction that separates points in space (Gillen et al. 2001).

Several studies have demonstrated the inadequacy of the isotropic assumptions, directly or indirectly. Gillen et al. (2001) test the difference in parameters of a one-dimensional semivariogram of house prices. They use data from January 1995 through March 1998 in both the north-south and east-west directions for twenty-one submarkets of Montgomery County, PA. Colwell and Munneke (2009) examine the structure of urban land prices. They show that land price gradients vary according to the direction from the commercial business district (CBD). However, although insufficiency of isotropic assumptions has been studied in these articles the gain in forecast accuracy obtained from explicitly anisotropic models has not been shown.

The purpose of this paper is to illustrate the incorporation of an anisotropic spatial correlation structure into a standard hedonic house price function based on a spatial econometric model. We use one of the most popular spatial models, the simultaneous autoregressive model (SAR)<sup>1</sup> with the simplest kernel form, the  $K$  nearest neighbor ( $K - NN$ ) kernel,<sup>2</sup> to demonstrate the practical use of anisotropic

<sup>1</sup> Other function specifications include the spatial lag model, the spatial error model, and the spatial Durbin model (see Can 1992; LeSage 1999; LeSage and Pace 2004).

<sup>2</sup> Other kernel forms include the negative exponential kernel (Dubin 1988), the Gaussian kernel, and so on. Thorsnes and McMillen (1998) use a semiparametric estimator to analyze the relationship between land values and parcel size in a sample of 158 undeveloped parcels in Portland, Oregon. They find that implementing the alternative kernels Gaussian, Epanechnikov, Quartic Triangular, and Uniform gives very similar results.

spatial modeling. However, it is equally applicable to more complex econometric models.

Our approach is closely tied to the anisotropic approach in non-linear scale-space image processing (Weickert 1996). In image processing, this local anisotropic technique preserves or even enhances anisotropic information such as edges, lines, or flow-like structures that may be blurred by using normal isotropic Gaussian smoothing methods. Although this anisotropic approach is based on filtering methodology and not on regression analysis, it defines the spatial correlation between point pairs in essentially the same way as spatial econometric estimators. We apply this idea to house price regressions with the aim of demonstrating how allowing for anisotropic spatial dependence produces more accurate predictions than OLS and conventional isotropic techniques.

The remainder of this paper is organized as follows. “The Simultaneous Autoregressive Model” introduces the SAR model with  $K - NN$  kernel that will act as our econometric vehicle for applying the anisotropic spatial correlation approach. “Modeling Anisotropic Spatial Dependence” discusses the main steps of this estimation technique. “Estimation Results of Anisotropic Spatial Modeling” gives our empirical results which are based on transaction data from properties in Lucas County, Ohio. The final section provides some conclusions.

## The Simultaneous Autoregressive Model

The hedonic housing price specification relates the market value of a property to the set of characteristics that determine its value. In this study, we consider a semi-logarithmic function form as follows<sup>3</sup>:

$$Y = \log(P) = \mathbf{X}\beta + \varepsilon \quad \text{with } \varepsilon \sim N(0, \sigma^2), \quad (1)$$

where  $P$  is the  $n \times 1$  vector of house prices,  $Y$  is a  $n \times 1$  vector of log of house price,  $\mathbf{X}$  is a  $n \times l$  matrix of house characteristics,  $\beta$  is a  $l \times 1$  vector of unknown hedonic coefficients, and  $\varepsilon$  is a  $n \times 1$  vector of residuals.

If the residuals are spatially dependent, then Eq. 1 becomes:

$$\begin{cases} Y = \mathbf{X}\beta + \varepsilon \\ \varepsilon = \rho \mathbf{W}\varepsilon + e \end{cases} \quad \text{with } e \sim N(0, \sigma^2), \quad (2)$$

where  $\mathbf{W}$  denotes a spatial weight matrix and  $\rho$  is the coefficient.  $\mathbf{W}$  is a standard matrix with each row summing to 1 and contains zeros on the main diagonal. The non-zero entries of the  $j$ th row of  $\mathbf{W}$  represent the observations whose errors interact with the error of the  $i$ th observation.

<sup>3</sup> Basu and Thibodeau (1998) state two reasons for selecting this functional form: 1) it is easy to interpret, and 2) it reduces heteroskedasticity. Bowen et al. (2001) demonstrate the second point by comparing different forms of functions.

If  $\mathbf{W}$  is known, we can estimate the regression coefficients by maximizing the log-likelihood value of Eq. 2. We can define the log-likelihood function as follows:

$$L = \frac{1}{2} \ln |(\mathbf{I} - \rho \mathbf{W})'(\mathbf{I} - \rho \mathbf{W})| - \frac{n}{2} \ln(2\pi\sigma^2) - \frac{(Y - \mathbf{X}\beta)'(\mathbf{I} - \rho \mathbf{W})'(\mathbf{I} - \rho \mathbf{W})^{-1}(Y - \mathbf{X}\beta)}{2\sigma^2} \quad (3)$$

Typically,  $\mathbf{W}$  is unknown. Standard methods derive the elements of  $\mathbf{W}$  on the basis of distance alone. We follow Pace and Gilley’s (1997, 1998) definitions:

$$w_{ij} = 1/N, \quad (4)$$

where  $w_{ij}$  denotes the weight of the interaction between location  $j$  and location  $i$ . The basic idea is to find the  $N$  nearest properties for the  $i$  th property within a search radius  $R$ , where  $N \in [m, n]$ , and  $m$  is the maximum and  $n$  is the minimum nearest properties.

This reliance on the assumption of isotropic spatial dependence in conjunction with the SAR model of Eq. 2 generates what we refer to as ISAR models. The alternative suggested here combines a weight matrix  $\mathbf{W}$  based on anisotropic spatial dependence with the SAR model of Eq. 2. We will refer to this model herein as the ASAR model.

### Modeling Anisotropic Spatial Dependence

If the spatial structure is anisotropic, the spatial dependence between the two comparable properties cannot simply be defined as a function of the distance between them. Coordinate transformation is the most common approach for this problem. We see that the global anisotropic technique, e.g., geometric anisotropic kriging, transforms the coordinate of each property uniformly using a certain rotation angle and a shrinking or dilating rate. But the local anisotropic method transforms the coordinates more flexibly. The main steps include gradient estimation, coordinate transformation, and iteration. The gradient estimation is to decide in which direction and to what degree the coordinates will be transformed at each individual point. Coordinate transformation is to reform the shape of the neighborhood from a circle to an ellipse to include more properties with higher spatial autocorrelation to the neighborhood. After that, we will iterate the previous two steps to generate a more desirable result.

#### Gradient Estimation

Just as in image processing we calculate the gradient to measure spatial structure changes. In a unidirectional one-dimensional space the gradient is defined as the rate of change to distance, that is,  $\nabla\rho = \frac{\Delta\varepsilon}{d}$ , where  $\Delta\varepsilon$  is the difference between the estimated spatial trend  $\varepsilon$  in two locations and  $d$  is the distance between the properties. Note here that  $\varepsilon$  in Eqs. 1 and 2 includes the spatial structure information, the part of the house price that cannot be explained by variables in the house price specification.

Under the anisotropic assumption we need to consider direction. So we use gradient vector  $\nabla\bar{\rho}$  in a two-dimensional space under the geodetic system as follows:

$$\nabla\bar{\rho}(s_{ij}) = \frac{\Delta\varepsilon_{ij}}{\Delta\bar{D}_{ij}} = \begin{bmatrix} \widehat{\varepsilon}(s_j) - \widehat{\varepsilon}(s_i) \\ u(s_j) - u(s_i) \\ \widehat{\varepsilon}(s_j) - \widehat{\varepsilon}(s_i) \\ v(s_j) - v(s_i) \end{bmatrix}, \quad (5)$$

where  $\nabla\bar{\rho}(s_{ij})$  denotes the gradient vector at the location  $i$  in the direction of  $ij$ .  $\Delta\varepsilon_{ij}$  is the spatial change from location  $i$  to location  $j$ .  $\widehat{\varepsilon}(s_i)$  and  $\widehat{\varepsilon}(s_j)$  denote the estimated residuals in location  $i$  and  $j$ , respectively.  $\widehat{\varepsilon}(s_i) = y(s_i) - \widehat{\beta}x(s_i)$ , i.e. the spatial trend is measured as the estimated residuals in the OLS regression (i.e. the estimated part of the residuals in the SAR model).  $\Delta\bar{D}_{ij}$  denotes the distance vector from property  $s_i$  to property  $s_j$ . Under the geodetic system this is equal to the difference in latitude and longitude between the property pair,  $\Delta\bar{D}_{ij} = \begin{bmatrix} \Delta\bar{u}_{ij} \\ \Delta\bar{v}_{ij} \end{bmatrix} = \begin{bmatrix} u(s_j) - u(s_i) \\ v(s_j) - v(s_i) \end{bmatrix}$ , where  $u(s_i)$  and  $v(s_i)$  denote the latitude and longitude coordinates of property  $s_i$ , respectively.

If we define the gradient on a certain property as equal to the average gradient of each nearby property to the center property we can calculate the gradient according to Eq. 6<sup>4</sup>:

$$\nabla\bar{\rho}(s_i) = \frac{1}{N} \sum_{j=1}^N \nabla\bar{\rho}(s_{i,j}) = \begin{bmatrix} \frac{1}{N} \sum_{j=1}^N \widehat{\varepsilon}(s_j) - \widehat{\varepsilon}(s_i) \\ \frac{1}{N} \sum_{j=1}^N u(s_j) - u(s_i) \\ \frac{1}{N} \sum_{j=1}^N \widehat{\varepsilon}(s_j) - \widehat{\varepsilon}(s_i) \\ \frac{1}{N} \sum_{j=1}^N v(s_j) - v(s_i) \end{bmatrix}, \quad (6)$$

where  $N$  is the number of neighbors.

### Coordinate Transformation

Figure 1 illustrates the basic idea of this step using an artificial example. Suppose all the points in Fig. 1 represent houses with similar structural characteristics. Hollow circles represent properties that are higher-priced because they are near a picturesque river; black points represent properties that are lower-priced because they face a noisy street. If we suppose the center property (the gray point) also has a low price, then we can identify the gradient direction as the arrow in Fig. 1c. Figure 1a and c illustrates the theoretical isotropic and anisotropic neighborhoods, respectively. An isotropic neighborhood should be a circle, because distance is the only factor (Fig. 1a). A typical anisotropic neighborhood is an ellipse because it considers both distance and direction (Fig. 1c). In an isotropic neighborhood, both black points and hollow circles are included according to their distances from the center point. But the elliptical anisotropic neighborhood excludes most properties in the direction of large spatial structure changes. Thus, some hollow circles are intentionally excluded from the elliptical anisotropic neighborhood.

<sup>4</sup> See the numerical example in the Appendix 1 for detailed calculation steps.

We can determine such anisotropic neighborhoods by means of coordinate transformation. Figure 1b and c tells us how such transformation is achieved. We use the geodetic coordinate system as our original system, as in Fig. 1a. To construct a local coordinate system, as in Fig. 1b, we first rotate the horizon's main axis in the direction of the gradient, and then shrink the coordinate in the perpendicular direction of the gradient.

Suppose  $\theta$  is the rotation angle and  $\varphi$  is the shrinkage or dilation rate of the principal axis of the ellipse. Then, we can define the transformed coordinates as:

$$\begin{bmatrix} u^* \\ v^* \end{bmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \varphi \end{pmatrix} \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \cos \theta \cdot u + \sin \theta \cdot v \\ -\varphi \sin \theta \cdot u + \varphi \cos \theta \cdot v \end{bmatrix}, \quad (7)$$

where  $\begin{bmatrix} u^* \\ v^* \end{bmatrix}$  is the transformed coordinate under the new local coordinate system,

and  $\begin{bmatrix} u \\ v \end{bmatrix}$  is the latitude and longitude coordinates under the geodetic system.

From Fig. 1b, we see that within the perpendicular direction of the gradient property A moves to the position of A\* with shrinkage rate  $\varphi$ . However, in the principal direction of the gradient property B remains in the same location.

Under this transformed local coordinate system, we can construct an isotropic circle neighborhood (Fig. 1b). Note that this isotropic circle is equivalent to an elliptical anisotropic neighborhood in the original coordinate system as in Fig. 1c. In other words, coordinate transformation can help us construct an anisotropic neighborhood using the traditional isotropic spatial correlation definition (the function of the distance between the property pair). But the distances in each approach are different. The distances used in anisotropic approach are calculated

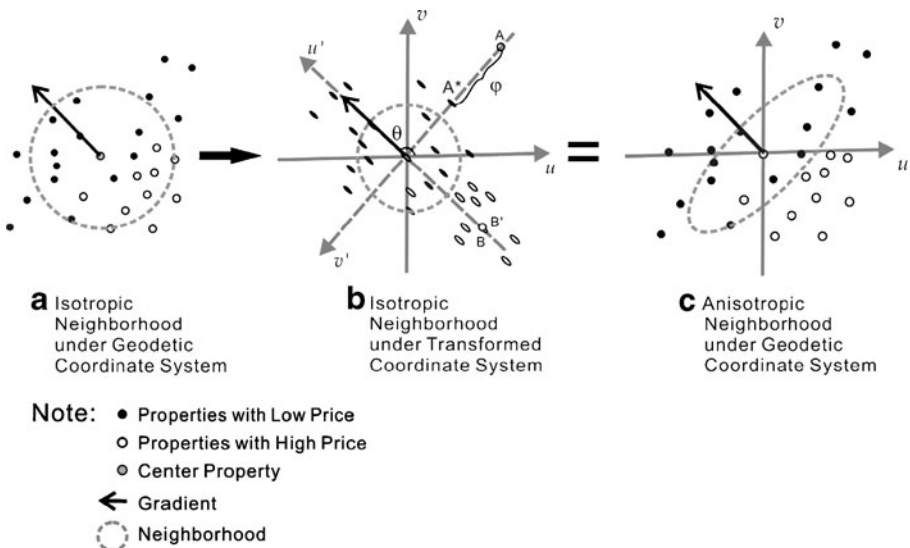


Fig. 1 Coordinate transformation

by using the transformed coordinates under the local coordinate system, as follows<sup>5</sup>:

$$\begin{aligned}
 d'(s_i, s_j) &= \sqrt{|\Delta\bar{u}_{ij}^*|^2 + |\Delta\bar{v}_{ij}^*|^2} \tag{8} \\
 &= \sqrt{(\cos \theta \cdot \Delta\bar{u}_{ij} + \sin \theta \cdot \Delta\bar{v}_{ij})^2 + (-\varphi \sin \theta \cdot \Delta\bar{u}_{ij} + \varphi \cos \theta \cdot \Delta\bar{v}_{ij})^2} \\
 &= \sqrt{(\Delta\bar{D}'_{ij} \cdot \bar{\tau}_u)^2 + (\varphi \Delta\bar{D}'_{ij} \cdot \bar{\tau}_v)^2},
 \end{aligned}$$

where  $\Delta\bar{u}_{ij}^*$  and  $\Delta\bar{v}_{ij}^*$  are the difference in the transformed coordinates from property  $s_i$  to property  $s_j$  under the local transformed coordinate system,  $\Delta\bar{D}_{ij}$  denotes the distance vector from property  $s_i$  to  $s_j$  under the geodetic system.  $\bar{\tau}_u, \bar{\tau}_v$  decide the rotation direction.  $\bar{\tau}_u$  is the unit vector of the horizon axis of the transformed local coordinate system; It is the same as the direction of the gradient, that is,  $\bar{\tau}_u = \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix} = \frac{\nabla\bar{\rho}(s_i)}{|\nabla\bar{\rho}(s_i)|}$ , where  $|\nabla\bar{\rho}(s_i)|$  is the absolute value of the gradient.  $\bar{\tau}_v = \begin{bmatrix} -\sin \theta \\ \cos \theta \end{bmatrix}$  is the unit vector of the vertical axis of the transformed local coordinate system, i.e. in the perpendicular direction to the gradient. Obviously,  $\bar{\tau}_u \cdot \bar{\tau}_v = 0$ .

The shrinkage rate  $\varphi$  can be defined as a function of gradient strength, i.e.,

$$\begin{aligned}
 \varphi &:= f\left(|\nabla\bar{\rho}(s_i)|\right) \text{ and} \\
 \begin{cases} f\left(|\nabla\bar{\rho}(s_i)|\right) < 1 \text{ if } |\nabla\bar{\rho}(s_i)| \neq 0 \\ f\left(|\nabla\bar{\rho}(s_i)|\right) = 1 \text{ if } |\nabla\bar{\rho}(s_i)| = 0 \end{cases}
 \end{aligned}$$

In this study, we define the function as follows:

$$f\left(|\nabla\bar{\rho}(s_i)|\right) = 1 / \sqrt{1 + |\nabla\bar{\rho}(s_i)|/b}, \tag{9}$$

where  $b$  is the parameter of the shrinkage rate equation indicating an increase in ellipse eccentricity with the unit change of gradient value. In image processing this equation is called the diffusivity function. It is an ad hoc function and can be applied in different forms (Weickert 1997). Here optimal value for parameter  $b$  is selected by maximizing the log-likelihood value in the regression.

The above steps<sup>6</sup> follow the anisotropic techniques of image processing. This technique helps to identify the neighbors with potentially higher autocorrelation to the center property. Therefore, the spatial weight matrix  $\mathbf{W}$  can better represent the structure information retained in the residuals of the house price specification.

<sup>5</sup> The distance in this study is calculated as the straight line between the latitude and longitude using the distance formula for the rectangular plane coordinate system. This is because, for a neighborhood, a spherical angle is negligible. In addition, we require only an approximate relative distance.

<sup>6</sup> Detailed calculation steps can be found in the numerical example in the Appendix 1.

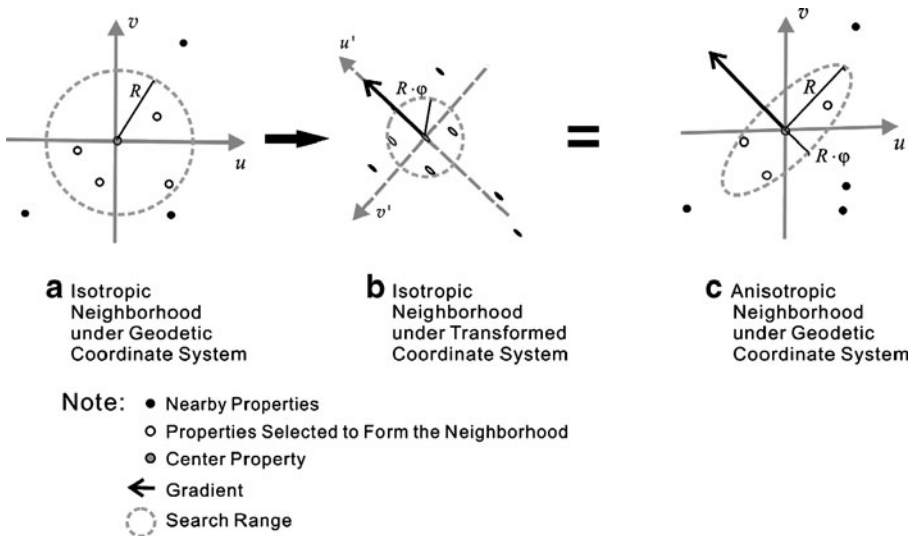


Fig. 2 Coordinate transformation with sparse properties

However, within image processing pixels are always in a regular grid. Houses, on the other hand, may be quite irregularly located. They may be very isolated or very tightly grouped. We define a search radius  $R$  and assume that properties located beyond such range have no spatial correlation with the center property. This is reasonable because spatial correlation usually declines as distance increases. Different radii will generate different prediction results for both ISAR and ASAR model. However, in spatial econometric models it is difficult to decide the optimal  $R$  before running the models. Therefore, we conduct the isotropic and anisotropic model under varied  $R$  and compare the results.

Scene I

For more isolated properties, most of the nearby properties usually have a longer distance to the center point than  $R$ . Figure 2a shows the isotropic neighborhood which is comprised of properties within this range.

Under the transformed local coordinate system we construct an isotropic neighborhood with a new search radius  $R'$ ,  $R' = R \cdot \varphi$  (Fig. 2b). This is equal to an elliptical neighborhood under the original coordinate system with a semimajor axis of  $R$  ( $R'/\varphi = R$ ) in the perpendicular direction of the gradient, and a semiminor axis of  $R \cdot \varphi$  in the principal direction of the gradient (Fig. 2c). Based on the new distances under the transformed local coordinate system, we select those properties with a smaller distance than  $R'$  to form the anisotropic neighborhood.

Scene II

For more tightly grouped properties, however, the maximum distance of the nearest  $N$  neighbors may be less than  $R$ .<sup>7</sup> In that case, we can construct the

<sup>7</sup> The numerical example in the Appendix 1 is obtained under this situation.



isotropic neighborhood using these  $N$  properties (Fig. 3a). We again establish a local coordinate system with the help of the gradient. To expedite the process, we choose to recalculate the distance of the nearest  $2N$  properties under the local coordinate system, instead of calculating the distance of all the properties in the district. According to the new distances, we then choose the  $N$  nearest properties (Fig. 3b).

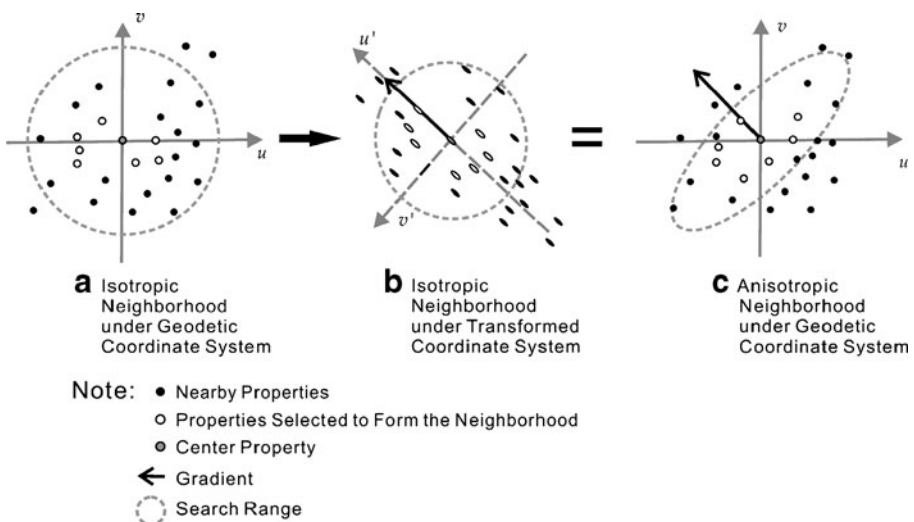
The properties perpendicular to the gradient will be our priority because their distance in this direction has been shortened. Figure 3c shows that the nearest properties chosen to form the neighborhood are different before and after the transformation.

*Scene III*

Figure 4 illustrates the case when the gradient is very large. Under this circumstance  $R'$  could be dramatically smaller because  $\varphi$  is small. Under the original coordinate system, the maximum distance of the nearest  $N$  properties may be smaller than  $R$ . Therefore the isotropic neighborhood is comprised of  $N$  properties (Fig. 4a). But after the transformation fewer properties are chosen because of the shortened anisotropic range  $R'$  (Fig. 4c). This is a reasonable consideration, because when the spatial structure changes greatly the neighborhood should be smaller.

*Scene IV*

In some special cases houses may be located so close to each other that they even share a common fence. In this situation, the gradient would be abnormally large because the distance is so short. Thus, the noise in  $\Delta\varepsilon_{ij}$  is amplified by the extremely



**Fig. 3** Coordinate transformation with intensive properties

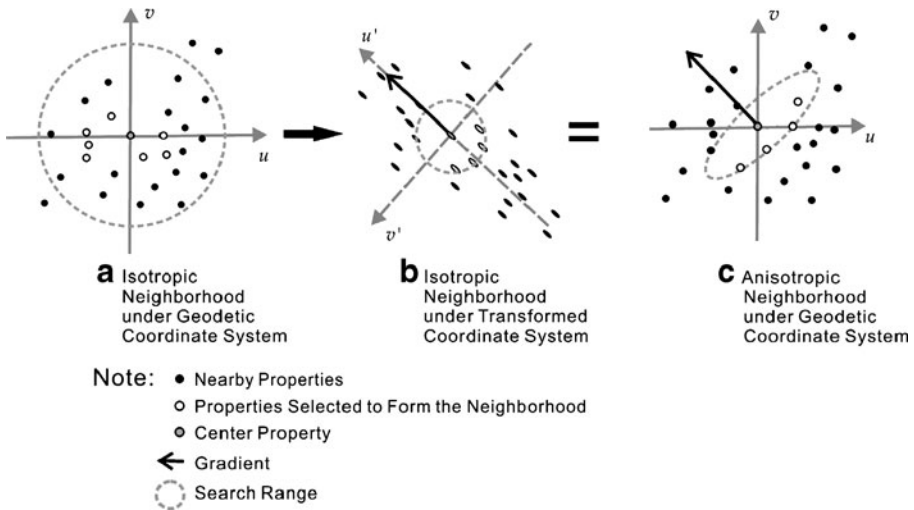


Fig. 4 Coordinate transformation with large gradients

short distance. So, if the distance between the coordinates of the property pair is less than  $135 \text{ m}^8$  we assume a gradient of  $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ . Within such a short distance we suppose that we would not see material changes in the spatial structure. This helps us reduce the influence of the noise in  $\Delta\varepsilon_{ij}$  when calculating the gradient.

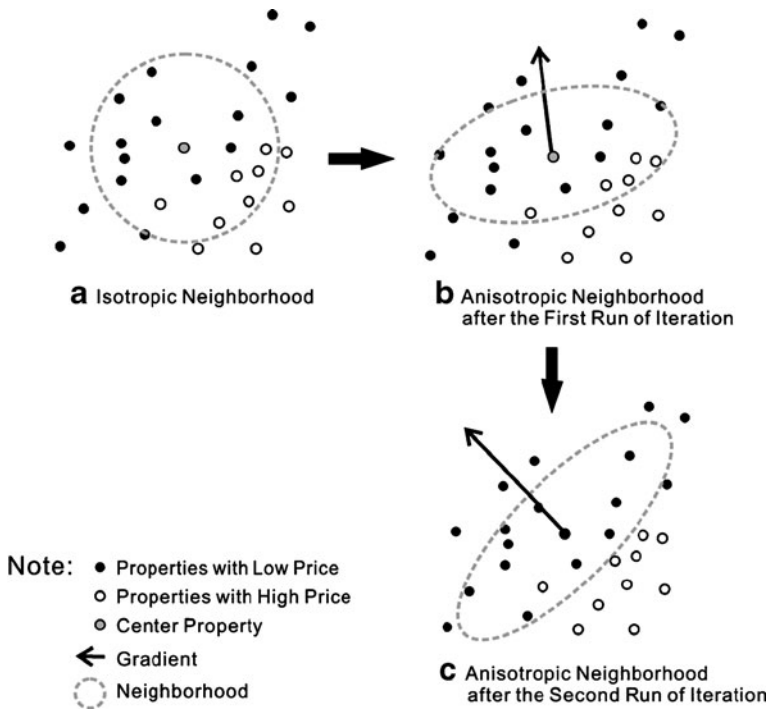
When the number as well as the distances of ‘nearest neighbors’ under the transformed local coordinate system are identified, we can use the isotropic definition in the last section to define the spatial weight matrix  $\mathbf{W}$ .

### Iteration

The purpose of this step is to use the prior information to modify the estimated gradient in order to ensure a precise neighborhood transformation. The shape of the transformed neighborhood is determined directly by the gradient (see Fig. 5). From Eqs. 5 and 6 it is apparent that the spatial trend  $\hat{\varepsilon}(s_i)$ , which determines the gradient, can be affected by the prior regression coefficients. However, the initial OLS coefficients are usually not efficient, because they are estimated under the independent error assumption.

Thus, to use the prior information and improve predictive power, it is necessary to conduct an iteration of the entire process. In our study, we find that the log-likelihood value varies only slightly after the second iteration. Therefore, we perform two iterations.

<sup>8</sup> The value depends on the spatial structure in the dataset. In our data, a value of 135 m fits the data best. Compared with other values, the anisotropic approach can best enhance the fitness of the model.



**Fig. 5** Different neighborhoods under different gradients

### Estimation Results of Anisotropic Spatial Modeling

The database we use here consists of 4,378 single-family home transactions from 1998:1 to 1998:4 in Lucas County, Ohio. The data are provided by the county auditor and contain structural housing descriptions (e.g., building year, lot size in square feet, room numbers), sale prices, sale time, and geographic coordinates.

Our sample data consists of two groups: We use 3,978 transactions for the estimation sample and another randomly selected 400 transactions (i.e. 10% of the database) are reserved for the prediction sample. Table 1 provides the descriptive statistics (mean and standard deviation) of these housing characteristics for the in- and out-of-sample dataset.

Figure 6 illustrates the estimated gradient using different numbers of neighbors. Each arrow represents the gradient at a certain point. The length of the arrow represents the value of the gradient; the direction of the arrow identifies the direction of the gradient. Figure 6a is derived using the nearest four properties, while Fig. 6b is derived using the nearest twenty properties.

Gradients based on larger neighborhoods tend to capture the average spatial trend and smooth the individual spatial structure. In contrast, using a smaller neighborhood retains more individual spatial characteristics. However, the gradients show

**Table 1** Descriptive statistics for sample data<sup>a</sup>

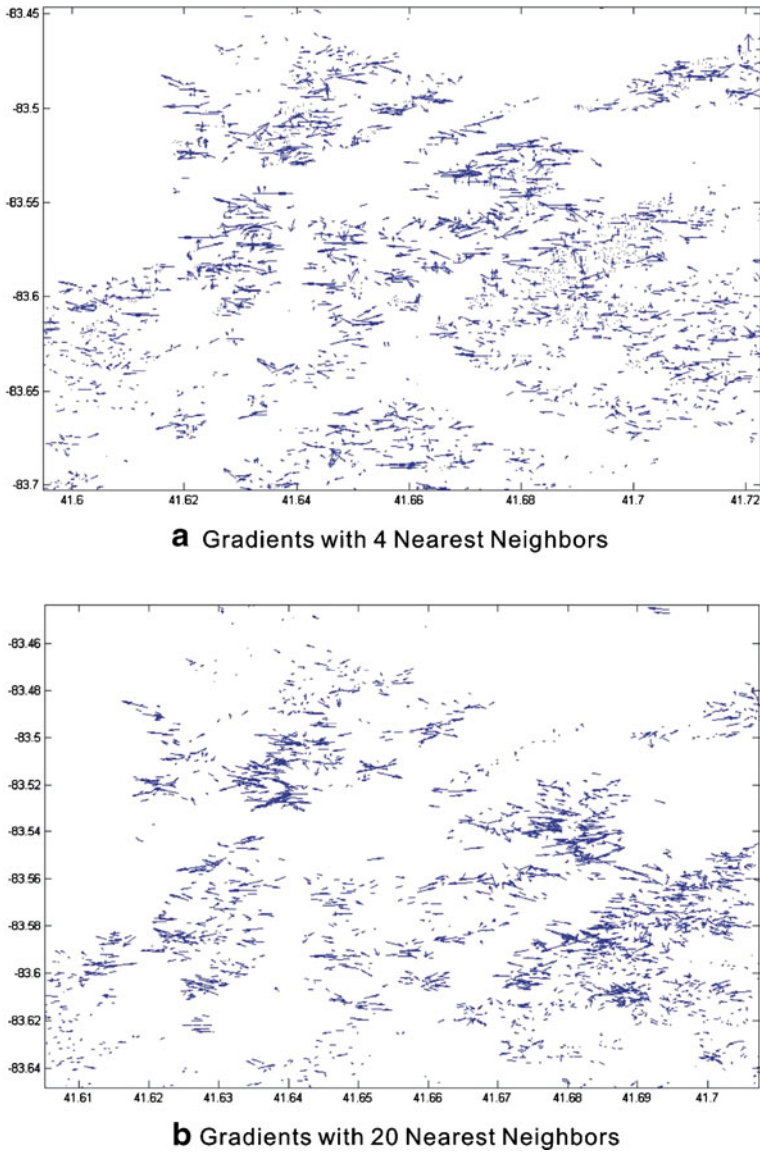
Variables	Estimation Sample		Prediction Sample	
	Mean	Standard Deviation	Mean	Standard Deviation
LP	11.085	0.787	11.048	0.758
AGE	0.638	0.274	0.650	0.253
AGES	0.482	0.375	0.486	0.345
LTS	8.938	0.772	8.871	0.723
LTSS	80.478	14.848	79.216	13.863
ROOM	6.077	1.297	6.025	1.292
BED	2.961	0.719	2.940	0.740
TLA	7.196	0.361	7.164	0.346
GAR	0.853	0.355	0.853	0.355
STOR	1.374	0.451	1.385	0.455
WALL_SIM	0.466	0.499	0.478	0.500
WALL_BRI	0.356	0.479	0.343	0.475
Q2	0.381	0.486	0.383	0.487
Q3	0.376	0.484	0.355	0.479
Q4	0.018	0.134	0.025	0.156
XCOORD	41.686	0.045	41.658	0.041
YCOORD	-83.602	0.079	-83.598	0.075

<sup>a</sup>The sample data consist of 4378 transactions in 1998 in Lucas County, Ohio; 400 transactions are randomly selected as the prediction sample; the rest are estimation sample. LP: log of sell price in US dollars; AGE: (sold year – building year)/100; AGES: AGE squared; LTS: log of lot size in square feet; LTSS: log of lot size squared in square feet; ROOM: number of rooms; BED: number of bedrooms; TLA: log of total living area in square feet; GAR: dummy variable for garage; STOR: number of stories; WALL\_SIM: dummy variable for using asbestos or composite tile shingles or stone or wood siding; WALL\_BRI: dummy variable for brick siding; Q2: houses sold in the second quarter; Q3: houses sold in the third quarter; Q4: houses sold in the fourth quarter; XCOORD: latitude; YCOORD: longitude

stronger noise and discontinuity. Thus, we cannot determine which number is best at this stage because there are no criteria for gradient accuracy. This can only be made by examining the final results.

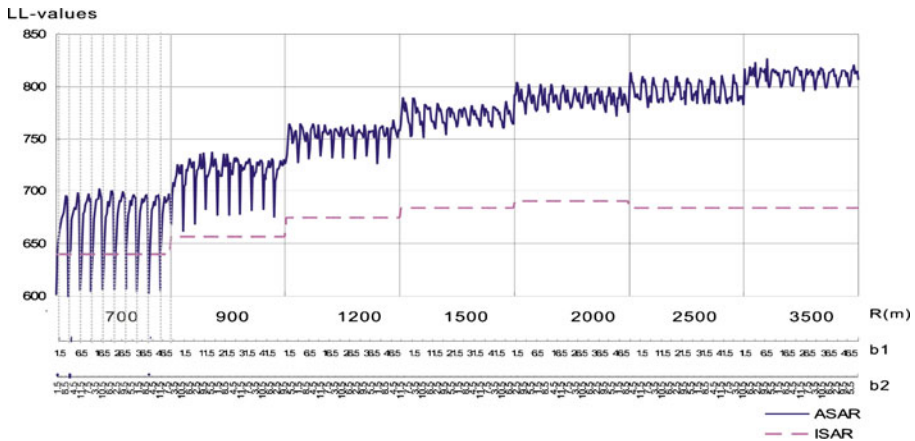
We find that the best results are achieved by using different neighborhood numbers to calculate gradient in each time of iteration. Our first run included up to the four nearest properties, and our second run included up to twenty properties. We posit that including both larger and smaller neighborhoods allows us to take advantage of both common and individual spatial information.<sup>9</sup>

<sup>9</sup>We use up to two times the nearest neighbors (i.e. 4/20 nearest neighbors in our case) to calculate the gradient, because in the anisotropic approach the final nearest 2/10 neighbors are selected from the nearest 4/20 neighbors, see Appendix 1. In doing so, we include all potential neighbors to calculate the gradient.



**Fig. 6** Gradients with 4/20 nearest neighbors

The input parameters of the isotropic approach include the range  $R$ , the maximum, and the minimum number of nearest properties. The anisotropic approach includes one additional parameter, namely parameter  $b$  in the shrinkage rate. In our dataset, best results in the ISAR model are obtained with a maximum of ten and a minimum of two properties. We use ten and two as the maximum and minimum number of neighbors in both isotropic and anisotropic approach.



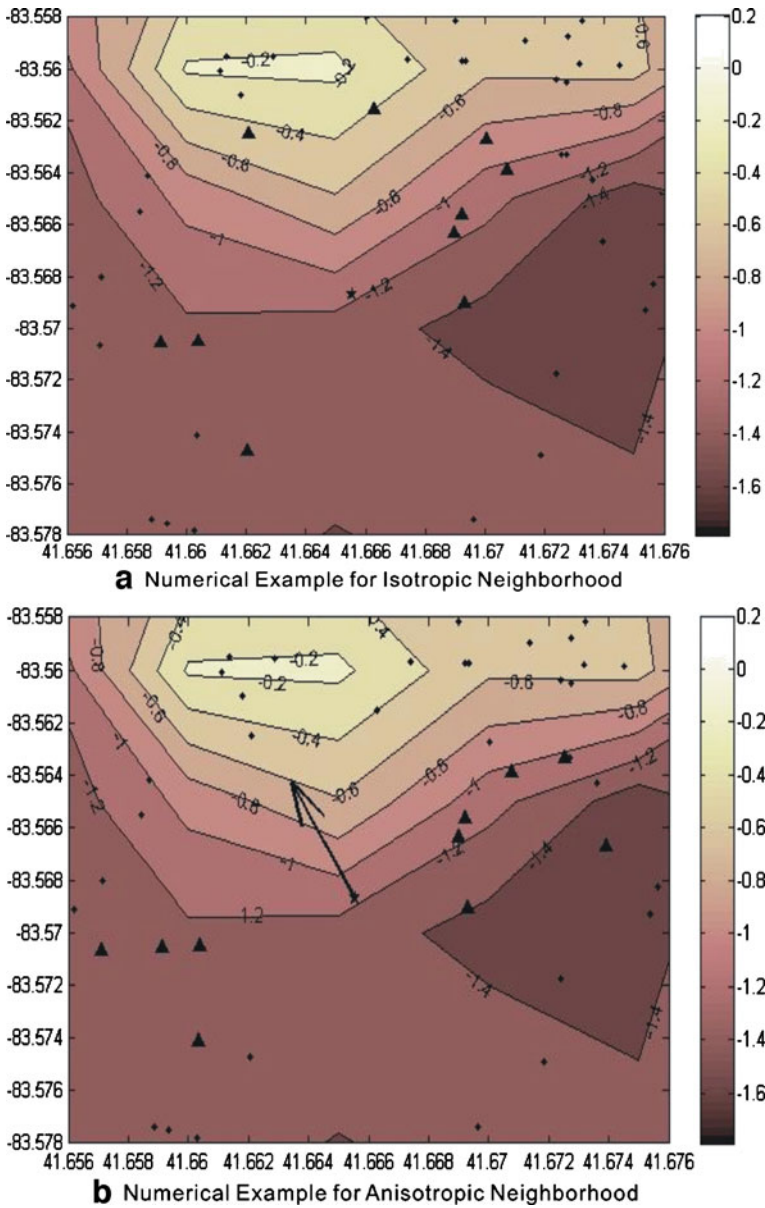
**Fig. 7** Comparison of log of maximized likelihoods with different radii and parameters ( $R$  is the neighborhood radius;  $b_1$  denotes parameter  $b$  in the diffusivity equation in the first run of iteration, from 1.5 to 46.5 and 5 as interval;  $b_2$  denotes the parameter  $b$  in the second run of iteration, from 1.5 to 11.5 and 1 as interval. The different likelihood values are sorted firstly by  $R$ , then by  $b_1$ , then by  $b_2$ , i.e. the first LL-value is generated by  $R=700$ ,  $b_1=1.5$  and  $b_2=1.5$ , the second by  $R=700$ ,  $b_1=1.5$  and  $b_2=2.5$ , et cetera, then twelfth LL-value is by  $R=700$ ,  $b_1=6.5$  and  $b_2=1.5$ , et cetera, then the 111 LL-value is by  $R=900$ ,  $b_1=1.5$  and  $b_2=1.5$ , et cetera)

The remaining parameters are the range  $R$  in the isotropic approach and  $R$ ,  $b_1$ ,  $b_2$  in the anisotropic approach.<sup>10</sup> Figure 7 illustrates the sensitivity of these parameters to the maximized log-likelihood (LL) values from both the isotropic and anisotropic approach. The different likelihood values are sorted first by  $R$ , then by  $b_1$ , and then by  $b_2$ . From Fig. 7 we can see that, with optimal  $b_1$  and  $b_2$ ,<sup>11</sup> the ASAR model always yields higher LL-values than the ISAR model for  $R$  between 700 to 3,500 m. The difference increases as  $R$  increases. After 2,500 m, the LL-value of the ISAR model does not change much because most of the ten nearest properties are within a radius of 2,500 m. For each value of  $R$  there is little variation in the LL value which implies that parameter  $b_1$  does not affect the final LL-values too much. This is because the purpose of the first iteration is to provide a more precise gradient but not to generate final estimation result. In contrast, the final LL-values are more sensitive to parameter  $b_2$ . However, the degree of the sensitivity decreases as  $R$  increases.

Figure 8 provides empirical examples of the isotropic and anisotropic neighborhoods. Different grey intensities reflect the spatial trends  $\hat{\varepsilon}(s_i)$ . Properties chosen to form a neighborhood are identified by  $\blacktriangle$ , while the center property is denoted by  $\star$ . Other outside properties are denoted by  $\blacklozenge$ .

<sup>10</sup> Because we iterate twice and use different neighborhoods in each iteration step to measure the gradient (i.e. a smaller neighborhood in the first run and a larger one in the second run), we have two parameter  $b$ :  $b_1$  in the first iteration and  $b_2$  in the second.

<sup>11</sup> The optimal  $b_1$  and  $b_2$  are selected by iterating the procedure in the given range from 1 to 50 for  $b_1$  and from 1 to 12 for  $b_2$  and comparing the log-likelihood value.



**Fig. 8** Comparison of property selection in neighborhood by ISAR and ASAR models (★ represents the center property; ▲ represents the neighborhood; ◆ represents other properties; → represents the gradient)

In the isotropic  $K$ -NN approach, the selected neighbors for property ★ are located over several regions where  $\hat{\varepsilon}(s_i)$  differs from  $-1.4$  to  $-0.4$ . In contrast, the anisotropic  $K$ -NN approach makes the shape of the neighborhood much flatter in the gradient direction. The  $\hat{\varepsilon}(s_i)$  of the neighbors changes from  $-1.4$  to  $-1$ .

Table 2 presents one set of the estimation results for the OLS hedonic regression as well as the ISAR<sup>12</sup> and ASAR models. The OLS model has an adjusted  $R^2$  of 0.766. The ASAR model raises the LL-value from 684.15 to 817.28 compared with the ISAR model. Most exogenous variables in the OLS regression are statistically significant except for the “number of rooms” variable, the “number of bedrooms” variable, the “brick siding wall” dummy variable and the “sold in the forth quarter” dummy variable.

In the ISAR model, the “number of bedrooms” variable, the dummy variable “simple siding wall” and the dummy variable “sold in the forth quarter” are statistically insignificant; other variables are significant at a 10% significance level. In the ASAR model, only the variable “number of bedrooms” is statistically insignificant; all other variables are significant at a 10% level. Both spatial dependent matrices in these two spatial models are highly significant.

Spatial autocorrelation models (the ISAR and ASAR models) provide a more logical explanation than the OLS model for the coefficient signs. For example, the coefficients for variable age in OLS regression indicates that for houses under 28 years old, age is positively correlated to the price, which is counterintuitive since the age of such houses is usually negatively correlated to its price. All the significant coefficients in spatial autocorrelation models have reasonable signs. The dependence matrix has a positive coefficient meaning that the spatial structure is positively autocorrelated.

Table 3 compares the prediction accuracies under different neighborhood ranges. Mean squared errors (MSE) is used to compare forecasting accuracy by different models (Dubin 1988; Pace and Gilley 1998; Bao and Wan 2004). For the ASAR model, in-sample MSEs are always smaller than those from ISAR model. At a radius of 2,500 m, the ASAR model can reduce the MSE by 53.38% compared to the OLS model and by 7.53% compared to the ISAR model. Furthermore, the out-of-sample MSE from the ASAR model is 55.38% less than in the OLS model and 12.69% less than in the ISAR model. Similar results hold for other ranges.

BT represents the percentage of times when the squared errors by the ASAR model are less than those of the ISAR model (Basu and Thibodeau 1998). We found that in most cases the ASAR model yields more accurate estimation results compared with the ISAR model both in-sample and out-of-sample. At a range of 2,500 m, for instance, ASAR in-sample estimations were 57.79% more accurate than ISAR estimations, while out-of-sample predictions were 59.75% more accurate than ISAR predictions.

To further test the ISAR and ASAR models forecasting performances, we use Wilcoxon’s (1945) signed-rank test and the Morgan-Granger-Newbold zero correlation test (Granger and Newbold 1977) which are used by Bao and Wan (2004). In many cases, we found that both tests reject the hypothesis of equal predictive accuracy at a highly significant level for the in- and out-of-sample predictions.

<sup>12</sup> When building the ISAR model we referred to the Spatial Econometrics Library by LeSage.



**Table 2** Regression results <sup>a</sup> (Neighborhood Range = 2,500 m;  $b_1=0.5$  and  $b_2=2.5$ )

Variables	OLS	ISAR	ASAR
AGE	0.885 <sup>b</sup> (0.106)	-0.181 <sup>d</sup> (0.095)	-0.185 <sup>c</sup> (0.093)
AGES	-1.579 <sup>b</sup> (0.075)	-0.441 <sup>b</sup> (0.068)	-0.393 <sup>b</sup> (0.067)
LTS	1.702 <sup>b</sup> (0.121)	0.544 <sup>b</sup> (0.033)	0.388 <sup>b</sup> (0.104)
LTSS	-0.076 <sup>b</sup> (0.006)	-0.021 <sup>b</sup> (0.002)	-0.013 <sup>c</sup> (0.005)
ROOM	-0.001 (0.010)	0.019 <sup>b</sup> (0.010)	0.017 <sup>c</sup> (0.007)
BED	-0.023 (0.014)	-0.006 (0.010)	0.003 (0.010)
TLA	0.657 <sup>b</sup> (0.033)	0.551 <sup>b</sup> (0.025)	0.525 <sup>b</sup> (0.024)
GAR	0.289 <sup>b</sup> (0.018)	0.139 <sup>b</sup> (0.013)	0.139 <sup>b</sup> (0.013)
STOR	0.206 <sup>b</sup> (0.017)	0.083 <sup>b</sup> (0.013)	0.072 <sup>b</sup> (0.012)
WALL_SIM	-0.084 <sup>b</sup> (0.018)	-0.034 <sup>b</sup> (0.013)	-0.021 <sup>d</sup> (0.012)
WALL_BRI	-0.017 (0.020)	0.017 (0.014)	0.029 <sup>d</sup> (0.014)
Q2	0.077 <sup>b</sup> (0.017)	0.066 <sup>b</sup> (0.012)	0.065 <sup>b</sup> (0.011)
Q3	0.080 <sup>b</sup> (0.017)	0.072 <sup>b</sup> (0.012)	0.071 <sup>b</sup> (0.011)
Q4	0.034 (0.047)	0.048 (0.032)	0.066 <sup>d</sup> (0.031)
CONSTANT	-2.999 <sup>b</sup> (0.601)	3.862 <sup>b</sup> (0.066)	4.791 <sup>b</sup> (0.646)
rho		0.849 <sup>b</sup> (0.003)	0.880 <sup>b</sup> (0.009)
$R^2$	0.767		
$R^2_{adjusted}$	0.766		
ML		684.15	817.28

<sup>a</sup> Calculations are based on the neighborhood range of 2,500 m and parameters  $b_1=0.5$  and  $b_2=2.5$ ; standard errors are in parentheses; rho is the coefficient of dependence matrix; ML is the log value of maximum likelihood; the OLS model assumes that house prices or corresponding residuals are independent of each other; ISAR is the traditional dependence matrix based on the assumption that the correlation between two comparable properties is a function of distance; ASAR is the SAR model with an anisotropic dependence matrix based on the assumption that the decay of dependence changes both in distance and direction

<sup>b</sup>, <sup>c</sup>, and <sup>d</sup> denote significance at the 1%, 5% and 10% levels, respectively

**Table 3** Predictive accuracy of the OLS, ISAR, and ASAR models under different neighborhood ranges<sup>a</sup>

R	MSE		BT	Wilcoxon's Signed-Rank Test	Morgan-Granger-Newbold Zero Test	
	OLS	ISAR				ASAR
<b>In-Sample</b>						
700		0.0739	0.0707	44.29%	-8.094 <sup>b</sup>	5.272 <sup>b</sup>
900		0.0734	0.0696	56.25%	-8.780 <sup>b</sup>	-0.092
1200		0.0730	0.0687	55.38%	-8.350 <sup>b</sup>	-1.818 <sup>d</sup>
1500	0.1448	0.0728	0.0681	55.86%	-9.213 <sup>b</sup>	-10.459 <sup>b</sup>
2000		0.0728	0.0681	55.93%	-14.394 <sup>b</sup>	-8.678 <sup>b</sup>
2500		0.0730	0.0675	56.91%	-9.802 <sup>b</sup>	-17.907 <sup>b</sup>
3500		0.0731	0.0676	55.86%	-9.826 <sup>b</sup>	-21.442 <sup>b</sup>
<b>Out-of-Sample</b>						
700		0.0902	0.0807	58.75%	-4.283 <sup>b</sup>	-4.415 <sup>b</sup>
900		0.0903	0.0814	63.50%	-7.703 <sup>b</sup>	-6.183 <sup>b</sup>
1200		0.0898	0.0811	64.50%	-5.928 <sup>b</sup>	-6.197 <sup>b</sup>
1500	0.1804	0.0898	0.0807	65.00%	-6.675 <sup>b</sup>	-6.059 <sup>b</sup>
2000		0.0917	0.0842	63.75%	-6.737 <sup>b</sup>	-3.456 <sup>b</sup>
2500		0.0922	0.0805	62.50%	-5.486 <sup>b</sup>	-5.698 <sup>b</sup>
3500		0.0918	0.0824	66.25%	-6.575 <sup>b</sup>	-8.739 <sup>b</sup>

<sup>a</sup> Calculations are based on the neighborhoods with the maximum ten nearest neighbors and minimum two nearest neighbors, but with different radii; R stands for the radius, in meters; MSE is the mean of squared errors; BT is the percentage of times that the squared error in ASAR model is smaller than that in the ISAR model. Both tests are used to determine whether the ASAR model contains significantly more useful information than the ISAR model. Both Wilcoxon's Signed-Rank Test and Morgan-Granger-Newbold Zero Test are used to determine whether the ASAR model contains significantly more useful information than the ISAR model

<sup>b</sup>, <sup>c</sup>, and <sup>d</sup> denote significance at the 1%; 5%; and 10% level, respectively

Unlike some global anisotropic neighborhood transformation methods, ours is more flexible because it is based on a local coordinate system transformation. Each neighborhood can be transformed differently according to the gradients in each point.

In order to compare our method with the global anisotropic method, we use the geometric anisotropic kriging method (GAKriging).<sup>13</sup> Table 4 shows the forecast accuracy for the OLS method, the isotropic kriging model (kriging), and the GAKriging. GAKriging produces greater MSEs and lower BTs than the isotropic kriging method for in-sample and out-of-sample predictions. In contrast, the ASAR model, when compared with the ISAR model, improves all four indices in most cases.

<sup>13</sup> The isotropic Kriging method has been applied on housing price spatial correlation analyses by many studies, see Dubin (1998); Basu and Thibodeau (1998); Gillen et al. (2001); Bourassa et al. (2007). Differently with the spatial econometric approach, which defines the spatial autocorrelation by spatial weight matrix, the Kriging method derives the correlation by semivariogram.

**Table 4** Predictive accuracy of global anisotropic approach<sup>a</sup>

	OLS	Kriging	GAKriging
In-sample			
MSE	0.1148	0.0784	0.0790
BT			49.50%
Out-of-sample			
MSE	0.1804	0.1841	0.1891
BT			47.75%

<sup>a</sup> MSE is mean squared error; BT is the percentage of times that the squared error is smaller than that in the OLS model; kriging refers to the kriging method, as per Bourassa et al. (2007); GAKriging refers to geometric anisotropic kriging, which is also referred to as Easykrig 3.0 when calculating one- and two-dimensional semivariograms

## Conclusions

Based on the anisotropic technique in scale-space image processing we propose an approach to modeling anisotropic autocorrelation in house prices. We apply the proposed anisotropic method to a simultaneous autoregressive model with  $K$  nearest neighbor kernel. The method can also be used for other spatial econometric models, such as the spatial lag model, the spatial Durbin model, and the general spatial model. It is also easy to pair with other kernel forms, such as a Gaussian kernel and a spherical kernel.

The basic steps include gradient estimation, neighborhood transformation, and iteration. In our study, we fitted a database consisting of transaction data from Lucas County, Ohio, to the model. We conducted both isotropic and anisotropic simultaneous autoregressive model using different search radii. In each circumstance, the anisotropic approach yields higher maximized log-likelihood values and reduces both in-sample and out-of-sample MSEs. In most cases, both Wilcoxon's signed-rank test and the Morgan Granger-Newbold test show that the anisotropic model performs substantially better than the standard isotropic model. Taking the results from 2,500 m neighborhood radius, for example, in-sample MSEs are reduced by 7.53% compared with a simultaneous autoregressive model with isotropic weight matrix and by 53.38% compared with an OLS model without spatial considerations. Out-of-sample MSEs are decreased by 55.38% and 12.69%, respectively.

Unlike the isotropic spatial dependence matrix the anisotropic spatial dependence matrix considers both distance and spatial structure information, measured by the direction and strength of the gradient. This is accomplished by two kinds of transformations: first, the shape of the neighborhood is altered from a circle to an ellipse; second, at some locations where the spatial structure develops extremely uneven the radius of the neighborhood is shortened to include fewer neighbors. These two transformations are accomplished automatically and differently at each location according to the gradient.

Compared with some global isotropic approaches (e.g. Geometric Anisotropic Kriging), the advantage of our local approach is twofold: First, the method can adapt to local characteristics more flexibly because the neighborhood is transformed differently at each point. Second, parameter estimates and prediction accuracy can be improved to the desired degree by iterating on the spatial dependence matrix without arbitrary choices by the researcher.

The importance of the anisotropic approach also depends on the urban growth pattern. It is difficult to expect a dramatic improvement in each region, especially when the spatial structure develops isotropic or nearly isotropic in each direction. In contrast, when the urban growth pattern has stripes or flow-like structures, e.g. rivers, commercial corridors along with major transportation arteries, a greater enhancement can be expected.

Because the gradient is the direct determinant of the neighborhood transformation it plays an important role in the proposed method. Properties are not usually located symmetrically in a grid, as are points in image processing. So we use an approximate method to avoid solving the non-linear diffusion filtering equation directly.

Future work in this area may want to improve the gradient calculation to better reflect spatial structure changes. Another objective could be to extend the method in the spatial-temporal dimension. Spatial econometric modeling is advantageous for temporal house price autocorrelation analyses. It would be useful to apply the anisotropic method to the temporal dimension in order to improve panel analyses such as those used to construct house price indices.

**Acknowledgements** The authors are grateful to James P. Lesage for allowing us to use the transaction data from Lucas County, Ohio. We would also like to thank Joachim Zietz, Klaus Uhlenbruck, and the anonymous referee for useful suggestions and helpful comments on previous versions of the paper. This manuscript has also benefited from the comments by participants at Immobilien- Forschungssymposium, EBS, 2008; ARES 2009 conference. We bear responsibility for any remaining errors.

## Appendix 1 Numerical example of anisotropic neighborhood definition

Using the properties in Fig. 8 as examples, we illustrate here how to construct an anisotropic neighborhood. The center property A has latitude and longitude of (49.665556, -83.5686870).

Table 5 shows the latitude, longitude, and spatial trends  $\hat{\varepsilon}(s_i)$  of property A and the nearest twenty properties. When defining an isotropic neighborhood, we must first use latitudes and longitudes to calculate the straight line distances under the geodetic coordinate system from each of the twenty properties to property A (Table 7, column 3). These properties are then sorted by ascending distance (Table 7, column 2) and we construct the isotropic neighborhood by using the ten nearest properties within a range of 0.027 (c.a. 2500 m), i.e., properties 1 to 10 (identified in Fig. 8a). The following steps show how to construct an anisotropic neighborhood.

**Table 5** Property descriptions

Property	Latitude $u(s_i)$	Longitude $v(s_i)$	Spatial Trend $\hat{\varepsilon}(s_i)$
A	41.665556	-83.5686870	-1.170891
A1	41.669310	-83.5690280	-1.191223
A2	41.668978	-83.5663320	-0.724011
A3	41.669209	-83.5656100	-1.046372
A4	41.660378	-83.5705240	-0.772146
A5	41.659139	-83.5705730	-0.990804
A6	41.662057	-83.5747290	-1.535109
A7	41.670743	-83.5638990	-0.931271
A8	41.662094	-83.5624770	0.318190
A9	41.666295	-83.5615320	0.550640
A10	41.670043	-83.5627040	0.105691
A11	41.672400	-83.5717510	-1.948181
A12	41.660337	-83.5740990	-0.872759
A13	41.658450	-83.5655000	-1.064081
A14	41.658698	-83.5641470	-0.701871
A15	41.657130	-83.5680390	-0.529503
A16	41.661816	-83.5610050	0.299630
A17	41.673925	-83.5666800	-1.157957
A18	41.657092	-83.5706570	-1.024243
A19	41.672542	-83.5633070	-1.088521
A20	41.671855	-83.5748730	-1.170891

**Gradient Estimation**

We first calculate the two-dimensional gradient from the nearest twenty properties to property A using Eq. 5. Table 2a shows spatial changes  $\Delta\varepsilon_{Aj}$ , distance vectors  $\Delta\bar{D}_{Aj}$ , and the estimated gradients. For example, the gradient from property A to property A1 can be calculated as follows:

$$\begin{aligned} \nabla\bar{\rho}(s_{A1}, s_A) &= \frac{\Delta\varepsilon_{A1}}{\Delta\bar{D}_{A1}} = \frac{\begin{bmatrix} \hat{\varepsilon}_{A1} - \hat{\varepsilon}_A \\ u_{A1} - u_A \\ \hat{\varepsilon}_{A1} - \hat{\varepsilon}_A \\ v_{A1} - v_A \end{bmatrix}}{\begin{bmatrix} -1.191223 + 1.170891 \\ 41.669310 - 41.665556 \\ -1.191223 + 1.170891 \\ -83.569028 + 83.568687 \end{bmatrix}} = \begin{bmatrix} -0.020332 \\ 0.003754 \\ -0.020332 \\ -0.000341 \end{bmatrix} \\ &= \begin{bmatrix} -5.419533 \\ 59.624633 \end{bmatrix}. \end{aligned}$$

The difference between the two longitudes is 0.000341 (about 20 m). Because of the small distance the gradient is unusually large while the spatial condition is unlikely to exhibit large changes. If the difference in latitude or

**Table 6** Gradient estimation

Property	Spatial change $\Delta \varepsilon_{Aj}$	Distance vector $\Delta_{-} D_{Aj}$	Gradient vector $\nabla_{-} \rho(s_{Aj})$
A1	-0.020332	[0.003754, -0.000341]	[-5.415933, 0]
A2	0.446880	[0.003422, 0.002355]	[130.590382, 189.758083]
A3	0.124519	[0.003653, 0.003077]	[34.086948, 40.467865]
A4	0.398745	[-0.005178, -0.001837]	[-77.007558, -217.063221]
A5	0.180087	[-0.006417, -0.001886]	[-28.064048, -95.486210]
A6	-0.364218	[-0.003499, -0.006042]	[104.092088, 60.281068]
A7	0.239620	[0.005187, -0.004788]	[46.196189, 50.045872]
A8	1.489081	[-0.003462, -0.006210]	[-430.121605, 239.787600]
A9	1.721531	[0.000739, 0.007155]	[0, 240.605275]
A10	1.276582	[0.004487, 0.005983]	[284.506782, 213.368198]
A11	-0.777290	[0.006844, -0.003064]	[-113.572474, 253.684729]
A12	0.298132	[-0.005219, -0.005412]	[-57.124392, -55.087190]
A13	0.106810	[-0.007106, 0.003187]	[-15.030910, 33.514166]
A14	0.469020	[-0.006858, 0.004540]	[-68.390259, 103.308457]
A15	0.641388	[-0.008426, 0.000648]	[-76.120080, 0]
A16	1.470521	[-0.003740, 0.007682]	[-393.187361, 191.424203]
A17	0.012934	[0.008369, 0.002007]	[1.545509, 6.444625]
A18	0.146648	[-0.008464, -0.001970]	[-17.326043, -74.440421]
A19	0.082370	[0.006984, 0.005380]	[-11.790739, 15.310428]
A20	0.832537	[0.006299, -0.006186]	[132.169775, -134.584107]

longitude is less than 0.0015 (about 135 m) we set the gradient equal to zero; hence,  $\nabla \bar{\rho}(s_{A1}) = \begin{bmatrix} -5.419533 \\ 0 \end{bmatrix}$ . The 20 estimated gradient vectors are in Table 6.

Using Eq. 6, the gradient at location  $S_A$  is calculated as the average gradient of the nearest neighbor properties:

$$\nabla \bar{\rho}(s_A) = \frac{1}{20} (\nabla \bar{\rho}(s_{A1}) + \nabla \bar{\rho}(s_{A2}) + \dots + \nabla \bar{\rho}(s_{A20})) = \begin{bmatrix} -26.8191 \\ 53.0670 \end{bmatrix},$$

as in Fig. 8b. The strength of  $\nabla \bar{\rho}(s_A)$  is  $|\nabla \bar{\rho}(s_A)| = \sqrt{(-26.8191)^2 + 53.0670^2} = 65.373$ . We can represent the direction of  $\nabla \bar{\rho}(s_A)$  as a unit vector  $\bar{\tau}_u = \frac{\nabla \bar{\rho}}{|\nabla \bar{\rho}(s_A)|} =$

$$\begin{bmatrix} \frac{-26.8191}{65.373} \\ \frac{53.0670}{65.373} \end{bmatrix} = \begin{bmatrix} -0.43184 \\ 0.90195 \end{bmatrix}, \text{ and the perpendicular direction as } \bar{\tau}_v = \begin{bmatrix} -0.90195 \\ -0.43184 \end{bmatrix}.$$

**Coordinate Transformation**

We can define the shrinkage rate  $\varphi$  as a function of the gradient strength using Eq. 9. For  $b=2.5$ , the shrinkage rate  $\varphi$  for property A is given as  $\varphi = 1/\sqrt{1 + 65.373/2.5} = 0.192$ . When  $\bar{\tau}_u$ ,  $\bar{\tau}_v$ , and  $\varphi$  are known, the distances in the

**Table 7** Distances under the geodetic and local coordinate systems

Property	Geodetic Coordinate System		Local Coordinate System	
	Rank	Distance $d$	Rank	Distance $d'$
A1	<b>1</b>	<b>0.003769</b>	<b>5</b>	<b>0.002020</b>
A2	<b>2</b>	<b>0.004154</b>	<b>1</b>	<b>0.001024</b>
A3	<b>3</b>	<b>0.004776</b>	<b>3</b>	<b>0.001498</b>
A4	<b>4</b>	<b>0.005494</b>	<b>2</b>	<b>0.001192</b>
A5	<b>5</b>	<b>0.006688</b>	<b>4</b>	<b>0.001650</b>
A6	<b>6</b>	<b>0.006982</b>	12	0.004102
A7	<b>7</b>	<b>0.007059</b>	<b>7</b>	<b>0.002460</b>
A8	<b>8</b>	<b>0.007110</b>	18	0.007096
A9	<b>9</b>	<b>0.007193</b>	16	0.006184
A10	<b>10</b>	<b>0.007479</b>	11	0.003698
A11	11	0.007499	14	0.005785
A12	12	0.007518	<b>10</b>	<b>0.002967</b>
A13	13	0.007788	15	0.006011
A14	14	0.008225	17	0.007094
A15	15	0.008451	13	0.004437
A16	16	0.008544	20	0.008544
A17	17	0.008606	<b>6</b>	<b>0.002408</b>
A18	18	0.008690	<b>8</b>	<b>0.002473</b>
A19	19	0.008818	<b>9</b>	<b>0.002484</b>
A20	20	0.008829	19	0.008314
Range		0.027000		0.005184

local transformed coordinate system can be calculated using Eq. 8. For example, the new distance between properties A and A1 can be calculated as:

$$\begin{aligned}
 d'(s_A, s_{A1}) &= \sqrt{\left(\Delta \bar{D}_{A1}' \cdot \bar{\tau}_u\right)^2 + \left(\varphi \Delta \bar{D}_{A1}' \cdot \bar{\tau}_v\right)^2} \\
 &= \sqrt{\left(\left[\begin{matrix} 0.003754 \\ -0.000341 \end{matrix}\right]' \times \left[\begin{matrix} -0.43184 \\ 0.90195 \end{matrix}\right]\right)^2 + \left(0.192 \times \left[\begin{matrix} 0.003754 \\ -0.000341 \end{matrix}\right]' \times \left[\begin{matrix} -0.90195 \\ -0.43184 \end{matrix}\right]\right)^2} \\
 &= 0.002020.
 \end{aligned}$$

Similarly, the distance between properties A and A2 can be recalculated as:

$$\begin{aligned}
 d'(s_A, s_{A2}) &= \sqrt{\left(\Delta \bar{D}_{A2}' \cdot \bar{\tau}_u\right)^2 + \left(\varphi \Delta \bar{D}_{A2}' \cdot \bar{\tau}_v\right)^2} \\
 &= \sqrt{\left(\left[\begin{matrix} 0.003442 \\ 0.002355 \end{matrix}\right]' \times \left[\begin{matrix} -0.43184 \\ 0.90195 \end{matrix}\right]\right)^2 + \left(0.107 \times \left[\begin{matrix} 0.003442 \\ 0.002355 \end{matrix}\right]' \times \left[\begin{matrix} -0.90195 \\ -0.43184 \end{matrix}\right]\right)^2} \\
 &= 0.001024.
 \end{aligned}$$

The new distances between the twenty properties and property A under the local coordinate system are given in Table 7, column 5. Comparing the distances under both the geodetic and the local coordinate system, we can observe ordering changes. For example, property 17 is the seventeenth nearest property to property A under the geodetic coordinate system, but it ranks sixth under the local coordinate system.

For the local coordinate system, we can calculate range  $R'$ , to determine which properties make up the revised neighborhood, as follows:  $R' = R\varphi = 0.027 \times 0.192 = 0.005184$ . In this case, the nearest ten properties under both the geodetic coordinate system and transformed local coordinate system are with their own range (0.027 and 0.005184, respectively). So ten properties are chosen by both isotropic and anisotropic approach, but they are different ten properties.

## Reference

- Bao, H. X. H., & Wan, A. T. K. (2004). On the use of spline smoothing in estimating hedonic housing price models: empirical evidence using Hong Kong data. *Real Estate Economics*, 32(3), 487–507.
- Basu, S., & Thibodeau, T. G. (1998). Analysis of spatial autocorrelation in house prices. *Journal of Real Estate Finance and Economics*, 17(1), 61–85.
- Bourassa, S. C., Cantoni, E., & Hoesli, M. (2007). Spatial dependence, housing submarkets, and house price prediction. *Journal of Real Estate Finance and Economics*, 35(2), 143–160.
- Bowen, W. M., Mikelbank, B. A., & Prestegaard, D. M. (2001). Theoretical and empirical considerations regarding space in hedonic housing price model applications. *Growth and Change*, 32(4), 466–490.
- Can, A. (1992). Specification and estimation of hedonic housing price models. *Regional Science and Urban Economics*, 22, 453–474.
- Colwell, P. F., & Munneke, H. J. (2009). Directional land value gradients. *Journal of Real Estate Finance and Economics*, 39(1), doi:10.1007/s11146-007-9104-0.
- Dubin, R. A. (1988). Estimation of regression coefficients in the presence of spatially autocorrelated error terms. *Review of Economics and Statistics*, 70, 446–474.
- Dubin, R. A. (1998). Predicting house prices using multiple listing data. *Journal of Real Estate Finance and Economics*, 17(1), 35–59.
- Gillen, K., Thibodeau, T. G., & Wachter, S. (2001). Anisotropic autocorrelation in house prices. *Journal of Real Estate Finance and Economics*, 23(1), 5–30.
- Granger, C. W. J., & Newbold, P. (1977). *Forecasting economic time series*. New York: Academic.
- LeSage, J. P. (1999). *Spatial Econometrics*. <http://www.spatial-econometrics.com/html/wbook.pdf>.
- LeSage, J. P., & Pace, R. K. (2004). Models for spatially dependent missing data. *Journal of Real Estate Finance and Economics*, 29(2), 233–254.
- Pace, R. K., & Gilley, O. W. (1997). Using the spatial configuration of the data to improve estimation. *Journal of Real Estate Finance and Economics*, 14(3), 333–340.
- Pace, R. K., & Gilley, O. W. (1998). Generalizing the OLS and grid estimators. *Real Estate Economics*, 26(2), 331–347.
- Thorsnes, P., & McMillen, D. P. (1998). Land value and parcel size: a semiparametric analysis. *Journal of Real Estate Finance and Economics*, 17(3), 233–244.
- Valente, J., Wu, S., Gelfand, A. E., & Sirmans, C. F. (2005). Apartment rent prediction using spatial modeling. *Journal of Real Estate Research*, 27(1), 105–136.
- Weickert, J. (1996). *Anisotropic diffusion in image processing*. Stuttgart: Teibner-Verlag.
- Weickert, J. (1997). A review of non-linear diffusion filtering. In B. ter Haar Romeny, et al. (Eds.), *Scale-space theory in computer vision, lecture notes in computer science* (pp. 3–28). Berlin: Springer.
- Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics*, 1, 80–83.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.