

Price-Contour-Based Spatial Dummy Variables for Segmenting Market in the Geographic Information System Assisted Hedonic Modelling of Residential Property Prices

Abdul Hamid bin Hj. Mar Iman, Ph. D

*Centre for Real Estate Studies,
Faculty of Engineering and Geoinformation Science,
Universiti Teknologi Malaysia,
81310 Skudai, Johor, Malaysia.
Email: hamid@fkg.utm.my*

Abstract

Dummy variable has been traditionally used to discriminate one sub-market against the other in the property market analysis. It is basically used to analyse various phenomena that have spatial influences on these sub-markets. This paper discusses the use of price-contour based spatial dummy variables for delineating urban residential sub-markets in the prediction of residential property values. Geographic Information System (GIS) is used to create price contours and these contours are used as a basis for creating spatial dummy variables representing sub-markets. These sub-markets are then included in the hedonic regressions to predict the prices of residential units in various urban sub-markets in the study area. The attribute and spatial data for the analysis are analysed using ArcView 3.1. The results show that GIS integrative capability to analyse the physical, locational, socio-demographic, economic, and market factors helps to demarcate the residential sub-markets within a particular urban area and this has improved price predictions.

Key words: Residential market, price-contour, spatial dummy, neighbourhood characteristics, GIS-hedonic.

1. Introduction

The residential market consists of sub-markets and the evidence of their presence is manifested in the spatial price differentials across particular geographic areas (Jones et al., 2001; Dunse et al., 2002). These sub-markets signify differing physical, social, economic, and market elements operating in different living communities.

Many studies have been undertaken to explain the variation in residential prices across particular geographic areas and, thus, the existence of residential sub-markets in those areas. In essence, differing physical, social, economic, and market elements exist in various “communal places of living” which, for convenience, are called “neighbourhoods”. Different neighbourhoods have different characteristics and they form the major factors that give rise to the variation in the residential

prices.

The effects of sub-markets on residential prices have been examined in a number of housing studies (Tse and Love, 2000; Tse, 2002; Lake et al., 2000). By including variable groups such as structural, sub-markets, accessibility or environmental, Lake et al. (2000) found that as a group, neighbourhood variables accounted for most of the variation in property prices followed by structural variables, environmental variables and accessibility variables.

However, the definition of “neighbourhood” in itself is dicey (Ding and Knapp, 2002). Different definitions of neighbourhoods and, thus, the criteria used to describe them are given according to the specific problems under

consideration (Megbolugbe et al., 1996). Among the criteria used are political jurisdiction (Straszheim, 1975; Goodman, 1981); zip code (Bingham and Zhang, 1997; Schwartz, 1999); census tract (Simons et al., 1998; Chen, 1994; Rothenberg et al., 1991); census blocks (Can, 1990); and vicinity around individual properties (Kain and Quigley, 1970). In this study, a neighbourhood is simply defined as the physical boundaries that form a polygon of a housing estate or a “taman”, which is similar to the concept of census blocks or census tracts.

In the same way, various techniques have been applied to demarcate residential sub-markets (Watkins, 2001; Bourassa et al., 1999; and Chen, 1994; Chen, 1994). However, neither these nor other housing studies have claimed the sufficiency of the techniques they used. This study attempts to examine the usefulness of GIS-generated price-contour interpolations to demarcate residential sub-markets and to incorporate these sub-markets in the hedonic modelling of residential property prices. Besides, this study also uses GIS to generate additional spatial neighbourhood information to be input to the hedonic models, through overlays, buffering, and polygon intersections.

GIS-hedonic analysis can be used to assess residential property prices against their determinants by depicting these determinants and the resulting prices spatially. The hedonic model itself has been used to analyse housing markets since the 1970s. Commonly used in the U.S.A. for mass appraisal purposes, it is also used in other countries such as the U.K., Australia, and Switzerland to a limited extent. In Malaysia, there have been some efforts in developing property database with a view to using it for the mass appraisal of ratable properties within the jurisdiction of local authorities.

This paper begins with defining residential sub-markets and urban neighbourhood characteristics. The price-contour technique for demarcating residential sub-markets is discussed next. A review on the hedonic analysis of the residential market is then presented, followed by a discussion on data collection and analysis procedure. The study area is Johor Bahru and a brief outline of its residential market precedes a description of the data and the research method. The following section is the empirical analysis. The final section highlights the key findings and suggests areas for further research.

2. Residential Sub-Markets and Identifying Urban Neighbourhood Characteristics

The theoretical literature in housing economics suggests that neighbourhood characteristics are important in determining the value of a housing unit (Tse and Love, 2000). Thus, to obtain accurate estimates of the impact

of its components on residential prices, it is necessary to include neighbourhood indicators but the level of aggregation of these variables is unclear. These characteristics, in turn, are used to segment residential market into smaller sub-markets which reflect either supply- or demand-related factors. For example, sub-markets may be defined by structural type (e.g. single-family detached, terraced residential, town-house, condominium); structural characteristics (land size, neighbourhood size, property age); or neighbourhood characteristics (e.g. ethnic groups, crime areas, etc.). These characteristics can be either determined a priori or using some statistical methods (see Bourassa et al., 1999).

The concept of neighbourhood has been analysed in many disciplines but there is little consensus on what constitutes a neighbourhood. Neighbourhood can be viewed as elements which give a sub-market its particular physical identity, including boundaries, land use, building groups, streets, public spaces, particular landscapes - natural and artificial, and specific areas such as a community area, local shopping centre or an important public building. A broader view looks at urban neighbourhood characteristics as a blend of physical, environmental, socio-economic, and market factors operating within urban areas (Furuseth, 2000). Unfortunately, complete and fine-grained data on many of the above elements are seldom available. Consequently, only a few of them are considered in urban studies. In this study, the set of variables representing neighbourhood characteristics are identified based on the local conditions affecting the residential market. These variables represent the property type, plot size, location, demographic, and market factors. Section 5.2 discusses these variables further.

However, an issue remains to be resolved: how could these factors be used to segment neighbourhood boundaries to improve predictions of residential property prices? Discussion follows.

3. Price-Contour Techniques for Demarcating Residential Sub-markets

Price-contour market segmentation techniques are an attempt to address the above issue. Specifically, price-contour interpolation technique is used to segment a particular residential market through property re-grouping into more homogeneous polygons on the basis of similarity of the selling prices (Hamid, 2001). [Figure 1 illustrates this technique.] This technique is similar to using a “dummy” variable to model the effects of two different groups of independent variables on a particular dependent variable (see Koutsoyiannis, 1986, Kennedy, 1992). The only difference is that, the dummy variables are created from spatial interpolations of property prices.

Selling prices are used as the basis for identifying sub-markets on the assumption that property prices are the actual manifestation of a spectrum of factors that influence buyers' choices for properties they buy. These considerations should have effectively included all the factors mentioned under Section 2.0. In other words, the myriad of factors that influence buyers' choices for properties are all merged into a single enveloping factor – market price. Therefore, it is logical to perceive market price as a general differentiating factor for neighbourhood. The point of interest here is whether the resulting interpolations significantly represent different sub-market profiles and, thus, can improve hedonic modelling.

Suppose one considers the spatial occurrences of residential property prices to be fairly uniform and, therefore, it is not necessary to group them into different sub-markets. Then, the hedonic model can be specified as:

$$P = a + bX + u \quad (1)$$

where P is the selling price, a , b , u are the regression intercept, coefficients, and error term respectively, and X is residential distance from the nearest town. In reality, however, properties are likely to demonstrate spatial dependence of selling prices among the neighbouring units (Pace and Gilley, 1997; Basu and Thibodeau, 1998; Dubin et al., 1999). Thus, neighbouring prices of similar property types tend to lie quite closely to each other, creating what is called a “clustering” phenomenon (Ripley, 1981; Wiltshaw, 1996), although they may be physically belong to different geographic boundaries (see Figure 1). Figure 1 shows an example of residential properties located with respect to the locations of two towns, with prices of the properties shown as contours over the area.

In such a situation, the use of equation (1) for estimating and predicting residential prices may not be reliable

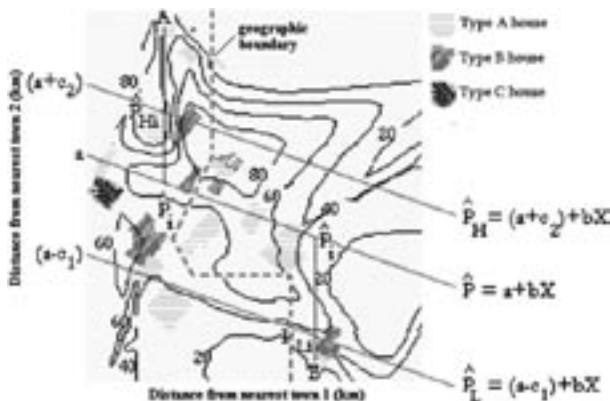


Figure 1 : An illustration of property grouping into different sub-markets

because residential units in the low-price sub-markets will tend to be under-estimated while those in the high-price sub-markets tend to be over-estimated. To improve modelling, especially for prediction, the properties need to be re-grouped into more homogeneous sub-markets.

One way is by creating sub-markets on the basis of similarity of residential prices surpassing the geographic boundaries. As shown in Figure 1, three sets of *isovalue* polygons are drawn to form three categories of sub-markets: low-price sub-market (RM 20-40/sq.ft.), medium-price sub-market (RM40-60/sq.ft.) and high-price sub-market (RM 60-80/sq.ft.) over the two geographically separated areas. In order to better estimate and predict residential property prices in these newly delineated sub-markets, equation (1) needs to be modified as follows:

$$P = a + bX + c_j D_{N_j} + u \quad (2)$$

where P , a , b , X , and u are as defined before; and D_{N_j} is a set of “dummy” variables representing price-contour sub-markets ($j=1,2,\dots,m$), and c is the coefficient of spatial dummy variable ($j=1,2,\dots,m$). Based on Figure 1 and the rules on using “dummy” variables, equation (2) is specified as follows:

$$P = a + bX + c_1 D_{N_1} + c_2 D_{N_2} + u \quad (3)$$

where a , b , P , X , and u are as defined before; D_{N_1} is a dummy representing residential properties in the low-price sub-market; D_{N_2} is a dummy representing residential properties in the high-price sub-market; c_1 and c_2 are their respective coefficients. To avoid perfect multicollinearity, one of the sub-markets – medium-price sub-market – is simply made the control group.

Equation (3) systematically differentiates among residential properties in the three sub-markets through the estimations of intercepts. For residences in low-price sub-market, P_L , $c_1 = 1$ and $c_2 = 0$. Therefore, equation (3) becomes:

$$P_L = a + bX + c_1 D_{N_1} + u \quad (4a)$$

For residences in the high-price sub-market, P_H , $c_1 = 0$ and $c_2 = 1$. Thus, equation (3) becomes:

$$P_H = a + bX + c_2 D_{N_2} + u \quad (4b)$$

For the medium-price sub-market, $c_1 = c_2 = 0$ and, equation (3) reduces back to equation (1). In Figure 1, the estimated equation of (1) is $P^ = a + bX$; that of equation (4a) is $P_L^ = (a - c_1) + bX$; and that of equation (4b) is $P_H^ = (a + c_2) + bX$. Obviously, equations (1), (4a), and (4b) give different model estimations. *Ceteris*

paribus, the difference in the intercept between equations (1) and (4a) is $(a - c_1) - a = -c_1$ while the difference in the intercept between equations (1) and (4b) is $(a + c_2) - a = c_2$.

Using the basic equation (3), residential prices in different sub-markets may then be predicted more accurately using different equations. This will depend on the specific location of individual properties. Residences in areas other than the low-price or high-price sub-markets are predicted using equation (1); those in the low-price sub-markets using equation (4a); and those in the high-price sub-markets using equation (4b).

Since $c^1 < a < c_2$, it can be shown that when residential prices are spatially clustered and, thus, forming some kind of sub-markets, the use of a hedonic model represented by equation (1) may under- or over-estimate value predictions with larger magnitudes compared to predictions using separate equations of (4a) and (4b). Figure 1, illustrates this point. A and B are the actual price/sq. ft. based on market transaction. A is in the high-price sub-market while B is in the low-price sub-market. The more appropriate predictor models are, therefore, $P_H^{\wedge} = (a + c_2) + bX$ for A and $P_L^{\wedge} = (a + c_1) + bX$ for B. It can be shown that predictions of A and B using $P^{\wedge} = a + bX$ will be less accurate.

The prediction error of A using $P^{\wedge} = a + bX$ is:

$$\begin{aligned} P_i^{\wedge} - A &= (a + bX_i) - A \\ &= a + bX_i - A \end{aligned} \quad (5a)$$

where P_i is the price of a specific property at site i (see Figure 1)

The prediction error of A using $P_H^{\wedge} = (a + c_2) + bX$ is:

$$\begin{aligned} P_{Hi}^{\wedge} - A &= ((a + c_2) + bX_i) - A \\ &= a + c_2 + bX_i - A \end{aligned} \quad (5b)$$

Subtracting equation (5b) from equation (5a), we have:

$$\begin{aligned} P_i^{\wedge} - P_{Hi}^{\wedge} &= (a + bX_i - A) - (a + c_2 + bX_i - A) \\ &= a - a + bX_i - bX_i - A + A - c_2 \\ &= -c_2 \end{aligned}$$

This means, the prediction of A using equation (1) has under-estimated its actual value (indicated by a negative sign). For B, the prediction error using $P_L^{\wedge} = (a - c_2) + bX$ is:

$$\begin{aligned} P_{Li}^{\wedge} - B &= ((a - c_1) + bX_i) - B \\ &= a - c_1 + bX_i - B \end{aligned} \quad (5c)$$

Subtracting equation (5c) from equation (5a), we have:

$$\begin{aligned} P_i^{\wedge} - P_{Li}^{\wedge} &= (a + bX_i - B) - (a - c_1 + bX_i - B) \\ &= a - a + bX_i - bX_i - B + B + c_1 \\ &= +c_1 \end{aligned}$$

This means, the prediction of B using equation (1) has over-estimated its actual value (indicated by a positive sign).

As for explanatory purposes, the sub-markets reclassification in our case here does not alter the hedonic prices of residential attributes, *ceteris paribus*. This is reflected in all equations (1), (4a), and (4b) where the slope of the regression, b , is unchanged.

To sum up the discussion in this section, the following conclusion can be made: when there is evidence of price clustering as a result of spatial dependence among the neighbouring residential properties, value-contour based re-definition of residential sub-markets may be able to improve property value predictions. This is because, effectively, properties in different sub-markets are estimated using separate hedonic equations.

4. GIS-Hedonic Analysis

4.1 Hedonic Specification

Residential property is an example of a heterogeneous utility-bearing product consisting of a bundle of attributes, each of which is integral to the residential unit. Each residential user is assumed to derive utility directly from the residential characteristics and chooses the residential unit which maximises this utility. In this framework, each residential unit is completely defined by a vector of characteristics which encompasses physical, locational, demographic, market characteristics, etc.

One salient feature of the hedonic approach, according to Rosen (1974) is that the implicit prices of attributes of a utility-bearing heterogeneous product can be derived by jointly estimating the locus of reduced-form demander's bid and supplier's offer functions (Figure 2).

In Figure 2, suppose the residential unit Z is composed of n attributes. The reduced-form functions are given by:

$$P = f(Z_1, Z_2^*, \dots, Z_n^*) \quad (6a)$$

where P is price of the good and Z_1, Z_2^*, \dots, Z_n^* are characteristics of the good, with asterisks (*) denoting optimum quantities of the characteristics involved. A demander's bid function for a specific characteristic, $\theta^i = f(Z_1, Z_2^*, \dots, Z_n^*, U^*, I)$, shows the price the buyer is willingness to pay for the varying amounts of the characteristic (Z_1), given the optimal quantities of other characteristics (Z_2^*, \dots, Z_n^*), utility (U^*), and income

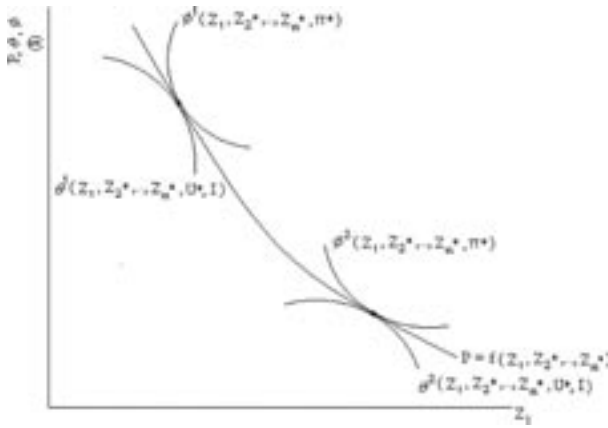


Figure 2 : Hedonic, Demanders' Bid, and Sellers' Offer Functions of Property Attributes (adapted from Rosen, 1974; King and Sinden, 1988).

(I). The seller's offer function with regard to a particular characteristic, $\phi^i = f(Z_1, Z_2^*, \dots, Z_n^*, \Pi^*)$, shows the price the seller is willing to accept for the varying amounts of the characteristic (Z_1), given the optimal quantities of other characteristics (Z_2^*, \dots, Z_n^*) and profit (Π^*).

In general, the hedonic price function represents equilibrium per unit prices of property parcels of various types, identified by the level of Z , resulting from the interaction of buyers and sellers in the property market (Xu et al., 1993). This implies that demand and supply jointly determine per unit property prices. Hence, the hedonic function is a joint envelope (in characteristic space) of demanders' bid curves and sellers' offer curves.

Based on Rosen's (1974) work, the residential market is modelled as a market for differentiated products due to attribute differences of the various property parcels (adapted from Palmquist 1984), with the assumption of equilibrium market condition (Freeman, 1979; Palmquist, 1989). This is expressed as follows (error terms ignored):

$$P_i^D(Z) = F_i(Z_1, \dots, Z_n, Y_1) \quad (7a)$$

$$P_i^S(Z) = G_i(Z_1, \dots, Z_n, Y_2) \quad (7b)$$

$$P_i(Z) = P_i^S(Z) \quad (7c)$$

where $P_i(Z)$ is the implicit market price of attribute Z_i ($i = 1, \dots, n$); Y_1 and Y_2 are exogenous demand and supply shift variables, respectively and; superscripts ^D and ^S denote demand and supply, respectively. If the demand and supply of attributes are responsive to price changes, then equations (7a) and (7b) should be estimated simultaneously. Rosen suggested a two-stage procedure whereby the first step is to regress the observed product price $P(Z)$ with all its attributes:

$$P(Z) = F(Z_1, \dots, Z_n) \quad (8)$$

The resulting marginal implicit prices, $P'_i(Z_i) = \delta P_i(Z) / \delta Z_i$, evaluated at each individual observation's level of Z , are then entered as endogenous variables in the second-stage simultaneous estimation of equations (7a) to (7c) (see Ohsfeldt and Smith, 1984). The second-stage estimation will include socio-economic variables such as income, age, etc. (Rosen, 1974; Freeman, 1974). The second-stage estimation assumes that sub-markets with similar socio-economic characteristics will respond similarly to any given set of property values irrespective of property neighbourhoods. By incorporating the implicit prices of property attributes derived from equation (8), the second-stage hedonic function can, thus, be specified as:

$$P_i = f\{P'_i(Z_i), Y_i, A_i, \dots\} \quad (9)$$

where P_i is equilibrium price; $P'_i(Z_i)$ is as defined above; Y_i is consumer per capita income and A_i is consumer age. Applying Rosen's framework to real estate modelling poses problems of reproduction of structural estimates of the second-stage price function; endogeneity of the quantity of the characteristic and its marginal implicit price in the price function; and non identifiability of the demand and supply functions (Brown and Rosen, 1982; Freeman, 1993; Palmquist, 1984; Pollak and Wachter, 1975; and Epplé, 1987).

To overcome these problems, this study has resorted to specifying the hedonic relationship as follows. First, the reduced-form hedonic model similar to the ones developed by Freeman (1974), Milon et al. (1984), and Palmquist (1989) is used. In general, the reduced-form model, assuming an equilibrium market condition, is expressed as follows (see Milon et al., 1984; Palmquist, 1989):

$$P = P(z_1, z_2, \dots, z_n) \quad (10)$$

where P is the price of property parcel and $z = (z_1, z_2, \dots, z_n)$ is a vector of n attributes of the property such as plot size, residential type, distance from town, and ethnic composition. Assuming a free-entry market, the buyers are unable to influence the price schedule although the prices paid will depend on the attributes of particular parcels. Similarly, the sellers cannot influence the price schedule, unless the attributes can be changed, for example, by erecting building or making improvements on the parcels. Given these assumptions, Equation (10) is used to estimate the implicit price (calculated at the mean value) of each of the parcel attributes under the *ceteris paribus* assumption (Xu et al, 1993).

Therefore, the implicit prices represent the equilibrium per unit prices of various attributes of property parcels,

identified by the levels of z , resulting from the interaction of buyers and sellers in the property market (Palmquist, 1989). This choice of specification also avoids simultaneity in the hedonic relationship which often leads to puzzling estimation problems.

Price-contour sub-markets are applied in this study to create general demand-supply shifters that somehow allow the separation of hedonic price functions for the heterogeneous residential markets. Although no specific identification is made for any particular demand-supply shifters, price-contour sub-markets provide a pervasive representation of the various interactions of urban neighbourhood elements underlying the variations in residential property prices across the study area.

Equation (10) is a general hedonic price function for both linear and non-linear representations of equation (7a through (9). The implicit price functions may be increasing, decreasing or constant depending on the functional form of $P(Z)$. The measurement of the implicit prices of the different attributes of the hedonic model raises questions about the correct model specification. The hedonic theory gives no indication about the best functional form to use. Thus, the method used to identify correct model specification is often assessed purely empirically.

In practice, the correct model specification can be determined by two approaches. A pragmatic approach simply identifies which set of results produces the best fit (collectively indicated by R^2 or adjusted R^2 , F-value, standard error of estimate, sum squared error) and provides the most consistent and plausible models. At a more complex level, Box-Cox transformations may be used, which incorporate statistical tests to determine the correct functional form (Box and Cox, 1964).

Re-writing Equation (10), gives the general specification for the first-equation hedonic model as follows:

$$P(Z)^{\lambda_1} = \beta Z^{\lambda_2} + \varepsilon \quad (11)$$

where Z is as defined earlier, ε is a normal independently distributed error term, and $P(Z)^{\lambda_1}$ and Z^{λ_2} are transformations of the forms:

$$\begin{aligned} P^{\lambda_1} &= (P^{\lambda_1} - 1)/\lambda_1 & \lambda_1 &\neq 0, P > 0 \\ \ln P & & \lambda_1 &= 0, P > 0 \end{aligned}$$

$$\begin{aligned} Z^{\lambda_2} &= (Z^{\lambda_2} - 1)/\lambda_2 & \lambda_2 &\neq 0, Z > 0 \\ \ln Z & & \lambda_2 &= 0, Z > 0 \end{aligned}$$

Different values of λ may be chosen for the dependent and continuous explanatory variables, but, to avoid the estimation becoming cumbersome, it is assumed that λ is equal for all variables (see Greene, 1990, p. 351). The

search procedure would not be efficient if there are more than two or three λ 's (Maddala, 1977). Therefore, the same value of λ is chosen for the above model, that is, $\lambda_1 = \lambda_2 = \lambda$.

In general, the Box-Cox transformation indicates whether the model to which the data will best fit is a linear or non-linear model. The Box-Cox transformation, however, has one major caveat: the optimum equation (one with the smallest sum squared errors, mean squared errors, or root mean squared errors) may not produce a model that can be easily used for estimating the implicit prices of property attributes (Milon et al., 1984). This occurs when the optimum equation results in the Box-Cox parameter, say λ_1 , such that $0 > \lambda_1 > 1$, for the dependent variable. Furthermore, parameter estimates tend not to be stable, that is, they are susceptible to the inclusion of other variables in the regression equation.

Therefore, the choice of model for estimating these implicit prices is normally confined to the special cases of Box-Cox functions. In this context, if $\lambda_1 = \lambda_2 = 1$, Equation (11) is linear; if $\lambda_1 = \lambda_2 = 0$, it is double log; if $\lambda_1 = 1$ and $\lambda_2 = 0$, it is linear-log; and if $\lambda_1 = 0$ and $\lambda_2 = 1$, it is log-linear. In these special cases, the choice of best function is determined primarily by the standard statistical tests, two of which are the likelihood ratio test (Griffith et al., 1993; Maddala, 1992) and the Box-Cox test for model equivalence (Greene, 1990).

In applying the best-fit criterion, models with different forms of the dependent variables (e.g. log-log and linear-linear) cannot be compared (Kennedy, 1992). Box and Cox (1964) have proposed a procedure to enable this comparison – test for model equivalence – and this can be found in Griffiths et al. (1993, pp. 344-347).

A fundamental problem in the hedonic modelling is misspecification. Like almost any other statistical techniques, misspecification is difficult to avoid despite the availability of a number of corrective measures. The implicit assumption that a model is correctly specified would probably never be realistic (Doran, 1989, p. 6). Therefore, it is imperative that the models used to estimate the implicit prices of properties are checked to ascertain whether they are correctly specified before use. The RESET (Ramsey, 1969) is normally used to examine misspecification while test of 'model equivalence' is used to choose among alternative specifications (Judge et al., 1985; Griffiths et al., 1993).

4.2 Geographic Information System as a Spatial Tool

We use the GIS database to do a number of tasks. First, we use GIS map to interactively measure the straight-line distance of particular residential sub-markets

from the central business district (City Square, Johor Bahru). Second, we use GIS map to show the spatial patterns of the factors determining investment returns for residential properties. The main purpose is to make certain observations and, thus, conclusions pertaining to the possible spatial relationship between residential price and its urban neighbourhood characteristics. Third, we use GIS to depict the resulting hedonic modelling, in particular, to show the distribution of the predicted residential prices in the study area.

GIS spatial functions are also used to create price contour based sub-markets for the hedonic modelling. The purpose is to subdivide the study area into a number of market segments reflecting the differences in the selling prices of residential properties in the study area. Using the function available in Arc View 3.1, residential sub-markets are interpolated using the irregular distance weight (IDW) technique.

5. Data and Analysis Procedure

5.1 Analysis Framework

The general steps in this study are as follows. First, sample out the single-storey terrace residential prices in Johor Bahru. Second, perform contour interpolation of selling prices of the properties to create residential “sub-markets”. Third, generalize these contours further to form major residential sub-markets with the purpose of reducing the number micro locations into three general sub-markets classes – low-price, medium-price, and high-price sub-markets. Fourth, specify the hedonic functions and regress the actual selling prices of individual properties against the neighbourhood variables and other property characteristics such as the physical, locational, socio-demographic, and market factors. [Figures 3 through 10 show the GIS maps used to generate data for the variables included in the hedonic models.] Sixth, estimate the implicit prices of these characteristics and test their similarity among the different sub-markets. Finally, the predictive performance of the hedonic models is evaluated by comparing the actual selling prices of residential properties with the regressed values within the GIS-generated price-contour sub-markets.

5.2 The Study Area and Data

5.2.1 *The Sample Area and Data*

The study area is Johor Bahru - Malaysia’s third largest capital city with a population of about 1.2 million - situated on the southern tip of Peninsular Malaysia. Serving as the main regional centre, it represents a fairly typical Malaysian city with the breakdown of employment sectors close to that of the national average.

The 1996 census records that about 25% of the working population are employed in the banking, finance and business service sector, a figure slightly higher than the Malaysian average of 22%. The main business centres are located in the vicinity of the CBD, approximately 3 km in radius and is bounded by various urban-fringe service centres. This area contains mainly multi-storey modern residential, office, and commercial buildings.

Various residential “sub-markets” are assumed to exist across the study area, from the CBD’s centre-point known as the City Square, extending as far as 30 km radius from it. There about 150 residential neighbourhoods in this area, at the time of study

A sample of 800 individual single-storey terraced units in 42 neighbourhoods (individual housing estates) is derived from the Property Market Report (1997-2000). This data set is used to create aggregate variables for the 42 neighbourhoods as the sample points. This means, on average, 19 individual residential units in each neighbourhood have been used to form aggregate neighbourhood variables. For each price observation within a given neighbourhood, there is a set of price determinants recorded for that neighbourhood. This serves a two-pronged purpose. First, to enable regression of prices against their determinants. Second, to enable mapping of prices and these determinants across the study area. As many as 200 other out-sample individual residential units are used for predictive purposes.

The hedonic price models are estimated through regression, in which the dependent variable is neighbourhood’s median sale price of single-storey terrace residences. The independent variables are specified according to Table 1 and some of them are expressed as dummy variables. The analysis, to some extent, is constrained by the availability of data on these attributes. A total of eighteen independent variables, set out in Table 1, describe the urban neighbourhood characteristics of the study area. The details of these variables are now considered.

5.2.2 *The Variables Considered*

5.2.2.1 *Physical Variables*

The physical attributes of residential properties within a particular area are an important element characterising urban residential sub-markets in Malaysia. The basic physical attributes of a residential product are its plot size and floor size. In Malaysia, residential developments almost always comprise mixed housing types with various land and floor sizes. For example, clusters of residential units are built in the size categories of 22 ft. x 70 ft.; 25 ft. x 65 ft.; etc. and are spatially randomly

distributed across any given geographic areas.

Plot size and floor size directly influence the selling price of a residential; the larger the size the higher is the price. One study found that 72% variation in the selling price of residential products is explained by plot and floor size (Ami and Associates, 2001). These two elements differ from one location to another. In Malaysia, the standard land size of landed properties in the housing schemes is generally in the range of 1,200 – 3,000 sq. ft. depending on the residential types. However, expensive residential properties (mainly bungalows) can be up to 45,000 sq. ft. in land area. Because there is a wide range of land size across urban locations, it can be conveniently used as an element characterising urban sub-markets with respect to residential property prices.

Floor size does not generally characterise sub-markets, although there could possibly be significant variation

in this factor, especially between older and more recent housing schemes across a region. Floor size is considered more of a physical product characteristic. Because of this, variation in the selling prices across residential sub-markets reflects the influence of different floor sizes on the selling prices of housing products. However, to avoid multicollinearity with plot size, this variable is excluded from the models.

Other core attributes of a residential product also influence its rental and market price. These are such as building design and layout. The cladding and standard of the exterior and structure are also important to portray the image of the product, and to the subsequent repair and maintenance expenditure. The type, design and quality of car porch, fence, automatic gate, etc. are extra features that contribute to product appeal and, thus, influence demand for residential properties. Since these attributes are so varied in terms of number and quality,

Table 1: List of Variables Used and Their Expected Results

Variable definition	Label	Source	Expected sign ^a
Dependent:			
Neighbourhood's median per unit residential price	PRICE	Property Market Report	
Neighbourhood physical characteristics:			
Neighbourhood's median lot size	LOTSIZE	Property Market Report	-
Neighbourhood's areal size	NEIG_SI	GISmap ^b	-
Neighbourhood locational characteristics:			
Neighbourhood distance from the CBD	DCBD	GISmap ^b	-
Neighbourhood within 1-km from expressway	WAYBUF	GISmap ^b	+
Neighbourhood within ½-km from railway	RAILBUF	GISmap ^b	?(+)
Neighbourhood within 1-km from trunk road	TRUBUF	GISmap ^b	+
Neighbourhood distance form nearest town	DNEAR	GISmap ^b	+
Neighbourhood within 2 ½-km from industrial centre	DFAC	GISmap ^b	?(+)
Neighbourhood's socio-demographic characteristics:			
% of Indian	INDIAN	JBCC, CJBMC	+
% of Chinese	CHINESE	JBCC, CJBMC	+
% of Malays	MALAY	JBCC, CJBMC	+
% of other races	O_RACE ^c	JBCC, CJBMC	+
Three-year average no. of violent crime	VIOCRI	City Police Dept.	-
Three-year average no. of property crime	PROCRI	City Police Dept.	?(-)
Sectoral population	SECPop	Dept. of Statistics	+
Market characteristics:			
Low-price sub-market	SUB_1	GISmap ^b	Control group
Medium-price sub-market	SUB_2	GISmap ^b	+
High-price sub-market	SUB_3	GISmap ^b	+

Note: ^a Question mark (?) indicates uncertain coefficient sign of the variable involved, but the sign in the bracket is the more possible regression outcome. ^b A detailed procedure of using these maps for generating spatial information used in the hedonic models can be obtained from the author on request. ^c Non-citizens are used as control group. JBCC = Johor Bahru City Council. CJBMC = Central Johor Bahru Municipality Council.

it is not possible to include all of them in the hedonic models. Therefore, they are assumed to represent a composite characteristic manifested in the residential type - single-storey, double-storey, semi-detached, and bungalows – within a given sub-market. *Neighbourhood size* is considered in the modelling because it is assumed that a larger neighbourhood supplies a greater number of housing stock and, thus, affect the selling prices of residential units, at least in that area. More supply of housing stock would possibly help lower the selling prices of the property.

5.2.2.2 Locational Variables

Across particular geographic areas, the key relationship determining property prices could be distance from the town, proximity to the highway, inter-city railways, links with commuter train and bus networks, etc. Straszheim (1975), for example, noted that variation in housing characteristics and prices by location is a fundamental characteristic of the urban residential market.

This study considers the straight-line *radial distance (km) from the central business district* (City Square, Johor Bahru) as a spatial variable in the hedonic model. In addition, the *straight-line distance (km) from the nearest local centre* is also calculated as a location variable. The local centre is included on the assumption that residential price is also influenced by micro location, i.e., the heart of the locality where there is predominant occupation of areas for commercial, industrial, financial, government residential, public facilities, and other private institutions. Because the study area is quite large, about 350 km², the influence of local market forces on property prices could be captured if consideration is given to satellite towns. Functioning in the same way as the central business district, local centres render various services to the local communities and, thus, form the immediate local markets for the local population.

Some other location elements such as *expressways, connecting (trunk) roads, inter-city railways* are also present in the study area and they could exert positive effects on residential prices due their connectivity and accessibility functions for the population or negative effects due to noise, dust, fume, and/or safety. To account for the possible influences of these transport networks on residential prices, residential properties located within half-km buffer from the expressways and railways are included in the hedonic model (Figure 3). Neighbourhoods located within one-km buffer from the *connecting (trunk) roads* are also included in the hedonic model. Besides, neighbourhoods located within 2 1/2-km radius from nearest industrial centre are included in the hedonic model as well (Figure 4).

5.2.2.3 Socio-Demographic Variables

Some aspects of the socio-demography play significant roles in shaping up urban sub-markets characteristics. Areas with a high crime rate may suppress residential prices due to unfavourable demand situations. The composition of ethnic groups (*% Indian, %Chinese, and % Malays*) within a particular sub-market may be associated with higher or lower residential prices. In the same way, sub-markets with a higher average income of the population can be associated with higher prices of residential products. The *size of sectoral population*, which create demand competition, may affect residential prices, negatively and, therefore, it is an important socio-demographic variable.

Based on the data obtained from the Johor Bahru City Council and Central Johor Bahru Municipality Council, we use ArcView 3.1 to assign spatial composition of ethnic group, number of crime committed, and sectoral population data to each neighbourhood (Figures 6-9). More refined information such as population composition by sex, age group, occupation, and income is not possible and due to lack of data sources, thus, cannot be incorporated into the hedonic models.

5.2.2.4 Market Variables

Residential prices are a result of various interacting factors such as product quality, accessibility, landscape, influence of local centre, etc. and they are manifested in the formation of arbitrary residential sub-market areas. In particular, Barlowe (1986, pp. 292-293) hypothesizes that smaller markets operate in different areas and deal with different types of properties.

In this study, market variables are simply defined as spatial boundaries within which residential areas are sub divided on the basis of profiles of residential prices. Due to the complexity of defining each element describing a market, GIS mapping capability is used to delineate these boundaries and, thus, the aerial demarcations of residential sub-markets.

6. Results and Discussion

The descriptive statistics of the sample variables are presented in Table 2. The single-storey terraced residential market in the study area has a price range between RM 101,000 and 250,500 per unit, with the standard deviation of RM 32,746. The mean per square foot price (land and building) of the properties is about RM1,163 with the standard deviation of about RM 222. The range of plot and floor size in the sample has been controlled to a considerably narrow range to avoid too much variation in the sample that can lead to less accurate price predictions.

The population of the study area witnesses the Chinese to slightly outnumber the Malays, but both racial groups are rather evenly distributed across the sub-region. The level of crime in the study area is quite controllable, if the size of Johor Bahru population is considered.

A number of models are tested using the enter procedure in the Statistical Package for Social Sciences (SPSS). In this procedure, all the seventeen independent variables are entered in the hedonic regression. Box-Cox transformation is undertaken to determine the most appropriate function. Besides, the explanatory power, model significance, plausibility of coefficient sign, and coefficient's magnitude are also assessed as a pragmatic approach to evaluate model's quality.

On the basis of Box-Cox transformation, two functional forms – linear and log-log – are found to be producing the best results, on the basis of R^2 , adjusted R^2 , F-value, standard error of estimate, and sum squared error. [Table 3 summarizes the main results for both specifications.] However, the Box-cox test of model equivalence as mentioned earlier found the log-log model to fit the sample data better. Notwithstanding this, the log-log specification gives slightly poorer explanatory power and some insignificant results compared to the linear specification. In contrast, the liner function provides the added advantage of ease of interpretation with regard to attribute prices. Nevertheless, for comparative purposes,

both linear and log-log functions are reported in this section.

From Table 3, we can make two important observations. First, much better regression results have been obtained by sub-dividing the study area into sub-markets, whereby these sub-markets are demarcated-based on GIS-generated residential price contours. Second, overall, the regressions using per unit of residential price as the dependent variable have produced better results compared to those using per sq. ft. price. Part of the explanation is that, the use of per sq. ft. price has reduced the necessary variation in the data set to explain the vast differences of residential characteristics across the study area.

Further analysis on the correlation matrix (not reported here, though), show that there is no serious problem of multicollinearity among the independent variables, thus, no variable modification is deemed necessary. Table 4 shows that the regression coefficients across the hedonic price models in terms of magnitudes and signs. The models explain about 61-63 percent (for per sq. ft. hedonic models) and about 80-82 percent (for per unit of residential hedonic models) of variation in residential selling prices.

A few variables have plausible signs and magnitudes by comparing Table 1 and Table 4. The constant represents the influence of all attributes not included in the regression

Table 2: Descriptive statistics of the sample variables.

Variables	N*	Minimum	Maximum	Mean	Std. Deviation
PRICE (RM)	42	101,000	250,500	162,535.7143	32,746.1398
PRICESQ (RM)	42	744.68	1,700.68	1,162.5017	221.6651
LOTSIZE (Sq. m.)	42	100.00	188.00	140.5714	16.1866
NEIG_SI (acre)	42	8.21	2738.31	366.1555	517.7462
DCBD (km)	42	2.24	18.38	10.2676	4.4267
WAYBUF (km)	42	0	1	.55	.50
RAILBUF (km)	41	0	1	.24	.43
TRUBUF (km)	42	0	1	.93	.26
DNEAR (km)	42	.01	3.40	1.6996	1.0074
DFAC (km)	42	.59	15.00	4.0052	4.1404
INDIAN (%)	42	3	23	10.52	7.51
CHINESE (%)	42	3	82	49.62	17.92
MALAY (%)	42	15	94	39.86	16.47
O_RACE (%)	42	2	51	7.21	10.20
VIOCRI (3-year average)	42	10	558	239.00	123.59
PROCRI (3-year average)	42	400	2,481	1,316.43	594.97
SECPop (person)	42	81,204	200,890	160,070.50	45,980.96
SUB_1 (RM)	42	0	1	.24	.43
SUB_2 (RM)	42	0	1	.45	.50
SUB_3 (RM)	42	0	1	4.76E-02	.22

* N repr □

Table 3: Basic Results (Dep. Selling Price)

	Without Market Segmentation			
	Per. Sq. Ft. Price		Per Unit of Residential Price	
	Linear	Log-log	Linear	Log-log
R ²	0.609	0.608	0.625	0.626
Adj. R ²	0.398	0.373	0.424	0.402
F-value	2.892	2.584	3.100	2.795
Std. error of estimate (SEE)	173.9668	6.713E-02	24,999.80	6.713-02
Sum Squared Error (SSE)	786,875.4	0.113	1.6E+10	0.113
No. of significant variables (at least 10% level):	5	4	6	4
No. insignificant variables	10	11	9	11
Total no. of variables	15	15	15	15

	With Market Segmentation			
	Per. Sq. Ft. Price		Per Unit of Residential Price	
	Linear	Log-log	Linear	Log-log
R ²	0.793	0.778	0.822	0.789
Adj. R ²	0.655	0.630	0.704	0.648
F-value	5.748	5.262	6.938	5.598
Std. error of estimate (SEE)	131.7249	5.153E-02	17,924.73	5.153E-02
Sum Squared Error (SSE)	416434.6	6.374E-02	7.7E+09	6.374E-02
No. of significant variables (at least 10% level):	7	5	7	4
No. insignificant variables	10	12	10	13
Total no. of variables	17	17	17	17

Table 4: Regression Results Including Sub-market Variables*

Variables	Per Sq Ft Price		Per Unit of Residential Price	
	Linear	Log-log	Linear	Log-log
	INTERCEPT	2251.188***	4.923***	131,688.346***
LOTSIZE (Sq. m.)	-8.156***	-0.984***	-9.139***	-7.738E-03
NEIG_SI (acre)	-3.655E-02	-7.738E-03	-2,950.457	-0.130
DCBD (km)	-26.097***	-0.130**	-3.806** *	-2.111E-02**
WAYBUF (km)	-43.086	-2.111E-02	-5,228.124	3.142E-02
RAILBUF (km)	111.967**	3.142E-02	16,885.088**	3.779E-02
TRUBUF (km)	129.060	3.779E-02	16,797.057	2.026E-04
DNEAR (km)	11.530	2.026E-04	2,115.454	2.372E-02
DFAC (km)	7.900	2.372E-02	1,292.228	-2.460E-02
INDIAN (%)	-0.566	-2.460E-02	-85.663	3.780E-02
CHINESE (%)	0.875	3.780E-02	186.573	7.811E-02
MALAY (%)	0.694	7.811E-02	117.479	2.723E-02
ORACE (%)	4.188**	2.723E-02	566.318	4.081E-02
VIOCRI (3-year average)	0.281	4.081E-02	36.103**	2.189E-02
PROCRI (3-year average)	-5.226E-03	2.189E-02	5.707	-8.625E-03
SECPop (person)	-4.962E-04	-8.625E-03	-3.550E-02	0.102
SUB_2 (RM)	242.374***	0.102***	37,482.331***	0.134***
SUB_3 (RM)	397.676***	0.134***	63,249.316***	1.632E-02***

* The reason for reporting only regression incorporating sub-market variables is to save space. Regression results without sub-market variables are, however, used to compare models' predictive capability (see Table 5). ** Significant at 10% level. *** Significant at 5% level and better.

equations and is the base to which other variables are added. In this study, the constant term in the linear per unit residential price model represents the 1997 “modal” single-storey terraced residential units located within the Johor Bahru’s outer city areas (housing estates within 20 km radius from the city).

An examination of the regression coefficients points to the importance of property sub-markets in explaining the variation in prices across the study area. Sub-markets, distance from the CBD, and plot size are among the most important factors affecting residential prices across the study area. For example, a single-storey terrace residential in the medium-price or high-price sub-market could have fetched RM 240-398 more of per sq. ft. selling price than that in the low-price sub-markets, based on the linear model. There also could have been a reduction of approximately RM 26 of per sq. ft. price of a residential unit located each km away from the CBD, based on the linear model. In the same way, a bigger plot size may have reduced per sq. ft. selling price of a single-storey terraced residential by RM 8.15.

The effects of public infrastructure (as proxied by neighbourhood distance from expressways, railways, and connecting roads), are quite noticeable, but are not significant to warrant a conclusive argument for their importance. The hypothesized effect of industrial locations on residential selling prices is as expected. Contrary to some belief that proximity to industrial areas tends to reduce property prices due to perceived environmental problems associated with such areas, industrial locations could have exerted a positive effect on property prices due to their symbiotic existence with other public and private facilities in the vicinity areas.

As a multi-racial society and with the government’s policies of promoting ethnic co-existence, the distribution of ethnic groups has a negligible (but positive) effect on the single-storey terraced residential sub-markets in the study area. Nevertheless, in some places, the existence of non-citizens could have boosted property prices as evidenced by the per sq. ft. price linear model.

Violent crime and/or property crime could have reduced selling prices of residential properties, but their effects were not found to be significant. Furthermore, the magnitudes of effects of these factors on property prices are very small. Effective safety and security controls by the police force, people’s perception of non-seriousness of crime levels in Johor Bahru, overwhelming desire to live in the urban areas, and simple ignorance of the crime conditions could have been the main reasons for this phenomenon.

Finally, there is evidence of lower residential property

prices in areas where there is a larger population, although the regression results are not significant. Coupled with the results for neighbourhood size, it can be said that, a larger supply of residential stock could naturally help bring down property prices in a particular area, *ceteris paribus*.

In order to test the structural stability of the models, the classical Chow test is undertaken (Chow, 1960). The data are randomly split into three samples and the hedonic models are re-estimated using all the samples. The calculated *F* statistic is 1.86 as against the tabulated critical value of 1.57 at 5% level. So, the null hypothesis is rejected and the conclusion is that, residential sub-markets do exist across the study area. Such a finding is expected based on the theory of housing market and empirical evidence from many previous studies.

A number of other variables are found to be insignificant but are not excluded from the models by assumption that they are theoretically pertinent urban neighbourhood variables (see Table 4).

One critical question of the usefulness of GIS-hedonic approach applied in this study is whether it is able to improve the predictive capability of the hedonic model – the very essence of any statistical based property value modelling.

Table 5 shows that properly sub dividing the residential market into sub-markets has led to improvement in the predictive capability of the hedonic models. Prediction results ‘with’ and ‘without’ sub-markets for both linear and log-log functions are contrasting. Dividing the residential market into sub-markets have resulted in more proportion (62-79%) of residential properties being predicted below $\pm 10\%$ margin of error. On the other hand, aggregating the residential market into one market has resulted in more proportion (up to 55%) of residential properties being predicted within $\pm 20\%$ margin of error.

Table 5 also shows that the log-log functions have better predictive ability compared to the linear functions. This is quite interesting especially because earlier discussion has disclosed that the log-log functions are less superior than the linear functions in terms of variable significance and, thus, are less desirable for explanatory purposes. This finding suggests that both linear and log-log functions may be used ‘back-to-back’ in any model-based property valuations due to their respective strengths.

Another interesting observation from Table 5 is that, the per unit residential price models (total price models) perform much better than the per sq. ft. residential price models. However, this has been so done with a caveat,

i.e. the property types and characteristic differences, especially in the plot size have been controlled within a reasonable range.

7. Conclusions

The main objective of this study is to examine the usefulness of GIS-based price-contour technique in creating spatial dummy variables to segment residential property market into sub-markets and to examine the contributions of neighbourhood characteristics to residential prices. The results emphasise on the importance of properly delineating residential sub-markets to ensure more accurate prediction of property prices. In addition, these results indicate the importance of the central business district as a regional residential centre. In other words, the distance from Johor Bahru city centre is found to be a major contributor to the determination of residential property prices.

The model explains about 80-82 per cent of variation in price with the consideration of sub-markets. From a

valuation point of view this level of accuracy is quite acceptable (Watkin, 2001). The analysis quantifies the impact that different urban neighbourhood characteristics have on residential prices and, as such, is an improvement on the “rules of thumb” used by valuers to adjust comparable evidence before applying them to the subject property.

The classical Chow test has indicated that the hedonic prices of urban neighbourhood characteristics are not likely to be uniform across geographic areas and, thus, justified the use of the technique proposed in this study. This study has found that using GIS to demarcate residential sub-market has resulted in a better predictive capability of the hedonic models. Clearly, this is a research area that needs to be pursued in future.

Acknowledgement

I would like to thank my Research Officer, Nik Adib b. Nik Din for his assistance in preparing the data and maps for the analysis.

Table 5: Predictive Capability of the Hedonic Models

	Without market segmentation variable			
	Per. Sq. Ft. Price		Per Unit of Residential price	
	Linear	Log-log	Linear	Log-log
Mean absolute percentage error	25.44	10.30	10.50	10.30
Below \pm 10%	19.05	57.14	54.76	57.14
\pm 10% - 20%	26.19	33.33	40.48	33.33
Above \pm 20%	54.76	9.52	4.76	9.52
	With market segmentation variable			
	Per. Sq. Ft. Price		Per Unit of Residential price	
	Linear	Log-log	Linear	Log-log
Mean absolute percentage error	8.13	6.78	8.84	6.79
Below \pm 10%	71.43	78.57	61.90	76.19
\pm 10% - 20%	19.05	16.67	28.57	19.05
Above \pm 20%	9.52	4.76	9.53	4.76

References

- Ami and Associates (2001). *Market Study for a Proposed Mixed Development on a 75-acre (30.36 hectare) Land in the Mukim of Kinta, Perak*. Report prepared for Tekad Intisari Sdn. Bhd. For submission to Perak State Economic Planning Unit.
- Barlowe, R. (1986). *Land Resource Economics: The Economics of Real Estate*, 4th ed., Englewood Cliffs, New Jersey: Prentice-Hall.
- Basu, S. and Thibodeau, T.G. (1998). Analysis of Spatial Autocorrelation in Residential Prices. *Journal of Real Estate Finance and Economics*, 17(1): 61-85.
- Bingham, R.D. and Zhang, Z. (1997). Poverty and Economic Morphology of Ohio, Central-City Sub-marketss. *Urban Affairs Review*.
- Bourassa, S. C., Hamelink, F., Hoesli, M. and MacGregor, B. (1999). Defining Housing Submarkets. *Journal of Housing Economics*, 8: 160-183.

- Box, G.E.P. and Cox, D.R. (1964). An Analysis of Transformation. *Journal of the Royal Statistical Society*, 26(2), Series B: 211-51.
- Brown, J.N. and Rosen, H. (1982). On the Estimation of Structural Hedonic Price Models, *Econometrica*, 50: 765-68.
- Can, A. (1990). Measurement of Sub-markets Dynamics in Urban Residential Prices. *Economic Geography*, 6: 254-272.
- Chen, I. (1994). Modeling Housing and Demographic Diversity at Census Tract Versus Block Group Levels of Aggregation. *Journal of the Urban and Regional Information Systems Association*, 6(1): 11-20.
- Chen, L. (1994). *Spatial Analysis of Housing Market Using Hedonic Based on GIS*. Massachusetts Institute of Technology: Unpublished Ph.D Thesis.
- Chow.G.C. (1960). Test of Equality Between Sets of Coefficients in Two Linear Regressions, *Econometrica* 285 591-605.
- Ding, C. and Knapp, G.J. (2002). Property Values in Inner City Sub-marketss: The Effects of Homeownership, Housing Investment, and Economic Development <http://www.smartgrowth.umd.edu/research/neighborhoodstability-816.pdf>
- Doran, H. E. (1989). *Applied Regression Analysis in Econometrics*. New York and Basel: Marcel Dekker, Inc.
- Dubin, R., Pace, R.K., and Thibodeau, T.G. (1999). Spatial Autoregression Techniques for Real Estate Data. *Journal of Real Estate Literature*, 7: 79-95.
- Dunse, N., Leishman, C. and Watkins, C. (2002). Testing for the existence of residential sub-markets: A comparison of evidence from two cities. *Urban Studies*, 39(3): 483-506.
- Epple, D. (1987). Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products. *Journal of Political Economy*, 95(February): 59-80.
- Freeman, A.M. (1974). On Estimating Air Pollution Control Benefits from Land Value Studies. *Journal of Environmental Economics and Management*, 1: 74-83.
- Freeman, A.M. (1979). Hedonic Prices, Property Values and Measuring Environmental Benefits: A Survey of the Issues. *Scandinavian Journal of Economics*, 81 (June) 154-173.
- Freeman, A.M. (1993). *The Measurement of Environmental and Resource Values*. Washington, D.C.: Resources for the Future.
- Furuseth, O. (2000). Charlotte Neighborhood Quality of Life Index. *Urban Quality Indicators – newsletter*, 19 (Fall): 8-10. [For further information, contact Urban Quality Communications, 1816 Pontiac Trail, Ann Arbor, Michigan 48105, United States.]
- Goodman, A. (1981). Housing Sub-markets within Urban Areas: Definitions and Evidence. *Regional Science Perspectives*, (9)1: 175-186.
- Greene, W.H. (1990). *Econometric Analysis*. London and New York: Mcmillan Publishing Company.
- Griffiths, E.G., Hill, R.C. and Judge, G.G. (1993). *Learning and Practising Econometrics*. New York: John Wiley & Sons, Inc.
- Hamid, Abdul, bin Hj. Mar Iman (2001). *Incorporating a Geographic Information System in the Modelling of Farm Property Values*. Unpublished Ph.D. thesis, Lincoln University, New Zealand.
- Jones, C., Leishman, C., and Watkins, C. (2001). Housing Market Processes, Urban Housing Submarkets and Planning Policy. Paper presented at *RICS Cutting Edge Conference 2001*, Oxford, United Kingdom.
- Judge, G. G., et al. (1985). *The Theory and Practice of Econometrics*. New York: John Wiley and Sons, Inc.
- Kain, J.F. and Quigley, J .M. (1970). Evaluating the Quality of the Residential Living Environment. *Environmental and Planning*, 2: 23-32.
- Kennedy, P. (1992). *A Guide to Econometrics*, 3rd ed., Oxford (UK) and Cambridge (USA): Basil Blackwell Ltd.
- King, D.A. and Sinden, J.A. (1988). Influence of Soil Conservation on Farm Land Values. *Land Economics*, 64(3): 242- 255.
- Koutsoyiannis, A. (1986). *Theory of Econometrics*, 2nd. ed., London: Macmillan Education Ltd.
- Lake, I.R., Lovett, A.A., Batemen, I.J. and Day, B. (2005). Using GIS and Large scale Digital Data to Implement Hedonic Pricing Studies, *International Journal of Geographic Information Science*, 14(6):521-541.

- Maddala, G.S. (1977). *Econometrics*. New York: McGraw-Hill Book Company.
- Maddala, G.S. (1992). *Introduction to Econometrics* (2nd ed.). Singapore: Maxwell Macmillan International.
- Megbolugbe, I.F., Hock-Smith, M.C., and Linneman, P.D. (1996). Understanding Sub-markets Dynamics: A Review of the Contributions of William G. Grigsby. *Urban Studies*, 33(10): 1779-1796.
- Milon, J.W., Gressel, J., and Mulkey, D. (1984). Hedonic Amenity Valuation and Functional Form Specification. *Land Economics*, 60 (4): 378-87.
- Oshfeldt, R.L. and Smith, B.A. (1984). Estimating the Demand for Heterogeneous Goods. *Review of Economics and Statistics*, 67: 165-71.
- Pace, R.K. and 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.
- Palmquist, R.B. (1984). Estimating Demand for the Characteristics of Housing. *The Review of Economics and Statistics*, 66(3): 394-404.
- Palmquist, R.B. (1989). Land as a Differentiated Factor of Production: A Hedonic Model and its Implications for Welfare Measurement. *Land Economics*, 65(1): 23-28.
- Pollak, R. and Wachter, M. (1975). The Relevance of the Residentialhold Production Function and Its Implication for the Allocation of Time. *Journal of Political Economy*, 83(April): 255-277.
- Property Market Reports* (1995-2001) (various years). Valuation and Property Services Division, Ministry of Finance, Kuala Lumpur.
- Ramsey, J. B. (1969). Tests for Specification Errors in Classical Linear Least-Squares Regression Analysis. *Journal of the Royal Statistical Society, Series B* (Methodological), 31: 350-371.
- Ripley, B.D. (1981). *Spatial Statistics*, New York: John Wiley.
- Rosen, S., (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of Political Economy*, 82(1): 34-55.
- Rothenberg, J. et al. (1991). *The Maze of Urban Housing Markets: Theory, Evidence, and Policy*. Chicago: University of Chicago Press.
- Schwartz, A. (1999). New York City and Subsidized Housing: Impacts and Lessons of City's \$5 Billion Capital Budget Housing Plan. *Housing Policy Debate*, 10(4): 839-878.
- Simons, R.A., Quercia, R. , and Maric, I. (1998). The Value Impact of New Residential Construction and Sub-markets Disinvestment on Residential Sales Prices. *Journal of Real Estate Research*, 15(2): 147-162.
- Straszheim, M. (1975). *An Econometric Analysis of the Urban Housing Markets*, New York, NY: National Bureau of Economic Research.
- Tse, R.Y.C. and Love, P.E.D. (2000). Measuring Residential Property Values in Hong Kong. *Property Management*, 18(5): 366-374.
- Tse, R.Y.C. (2002). Estimating Sub-markets Effects in Residential Prices: Towards a New Hedonic Model Approach. *Urban Studies*, 39(7): 1165-1180.
- Watkins, C. (2001). The definition and identification of housing submarkets. *Environment and Planning A*, 33: 2235-2253.
- Wiltshaw, D.G. (1996). Spatial autocorrelation and valuation accuracy: a neglected methodological problem. *Journal of Property Research*, 13: 275-286.
- Xu, F., Mittelhammer, R.C., and Barkley, P.W. (1993). Measuring the Contribution of Site Characteristics to the Value of Agricultural Land. *Land Economics*, 69(4): 356-69.

1 km Buffer From Trunk Road and Highway
and 1/2 km Buffer From Railway

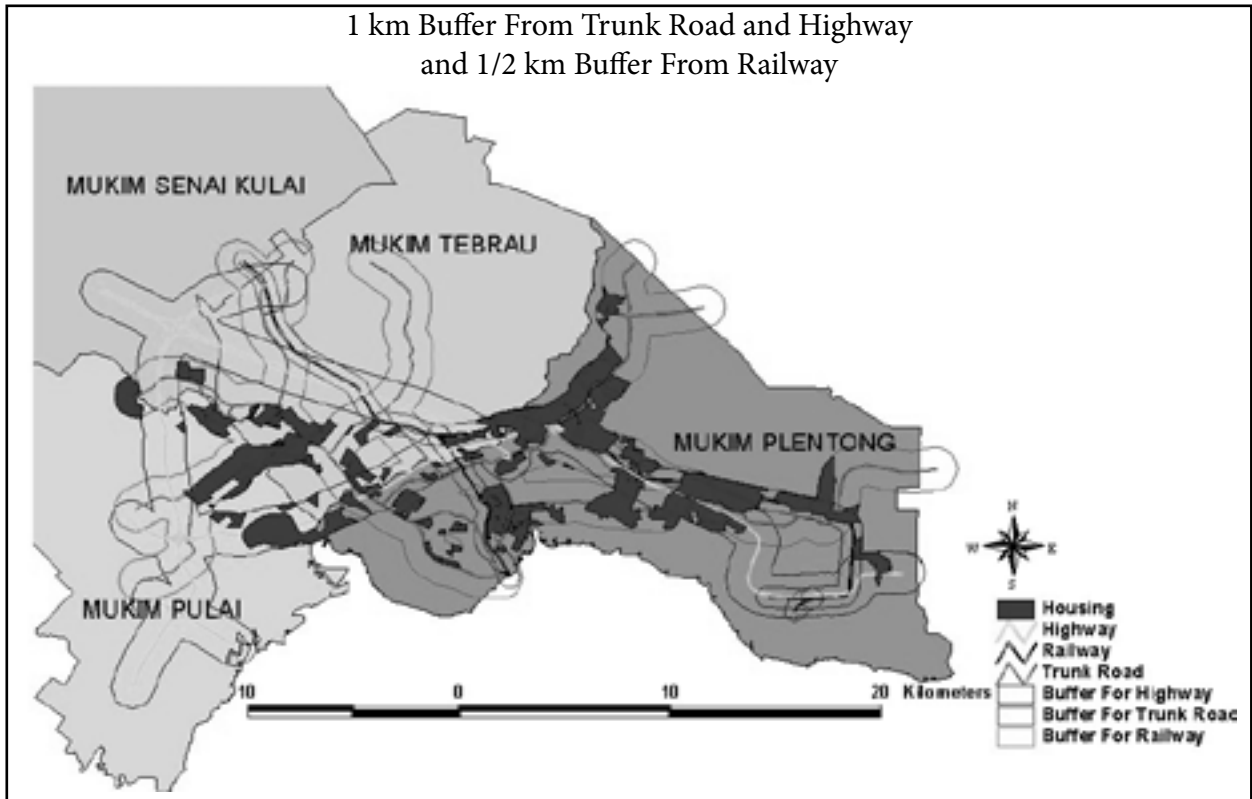


Figure 3

3 km Buffer From Manufacturing/Industrial Site

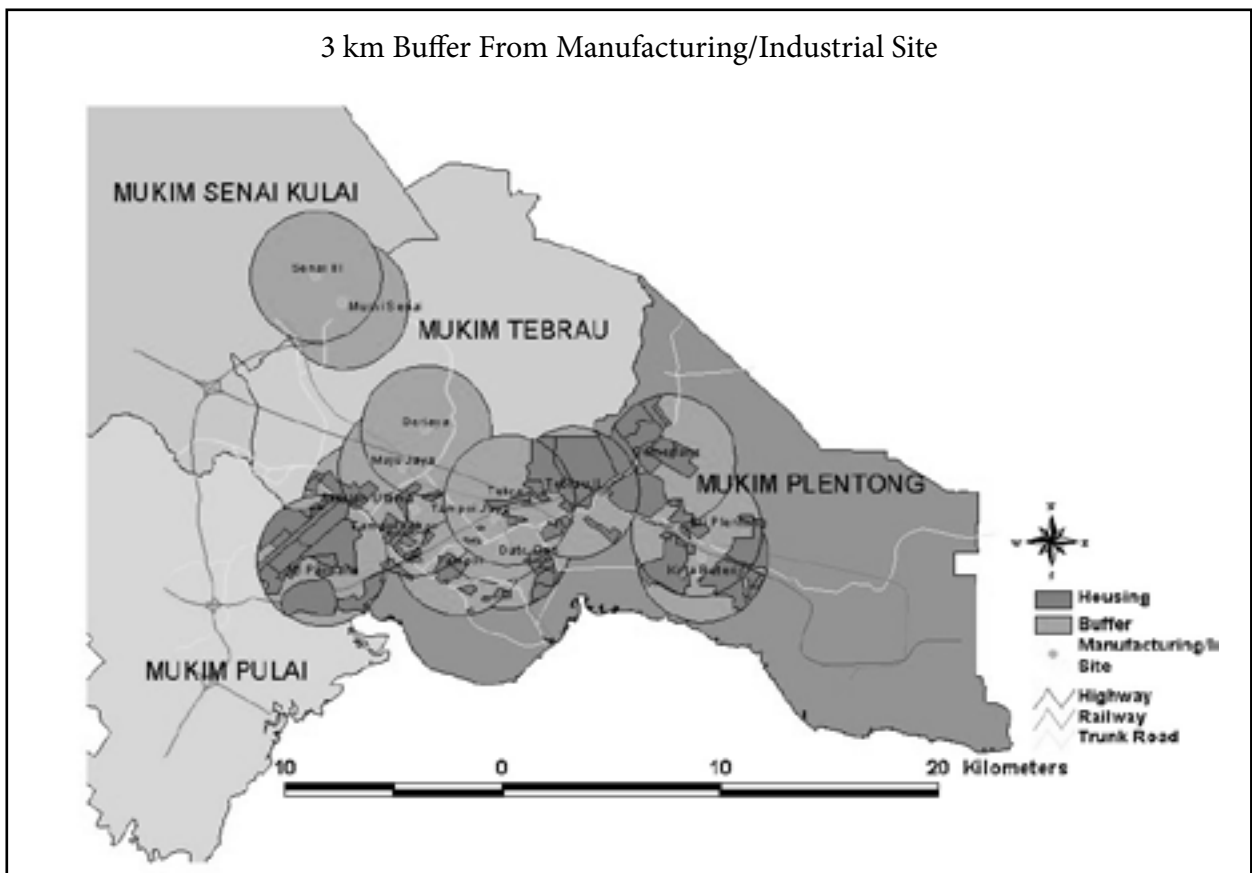


Figure 4

Interpolated Contours of Ethnic Composition For Indian

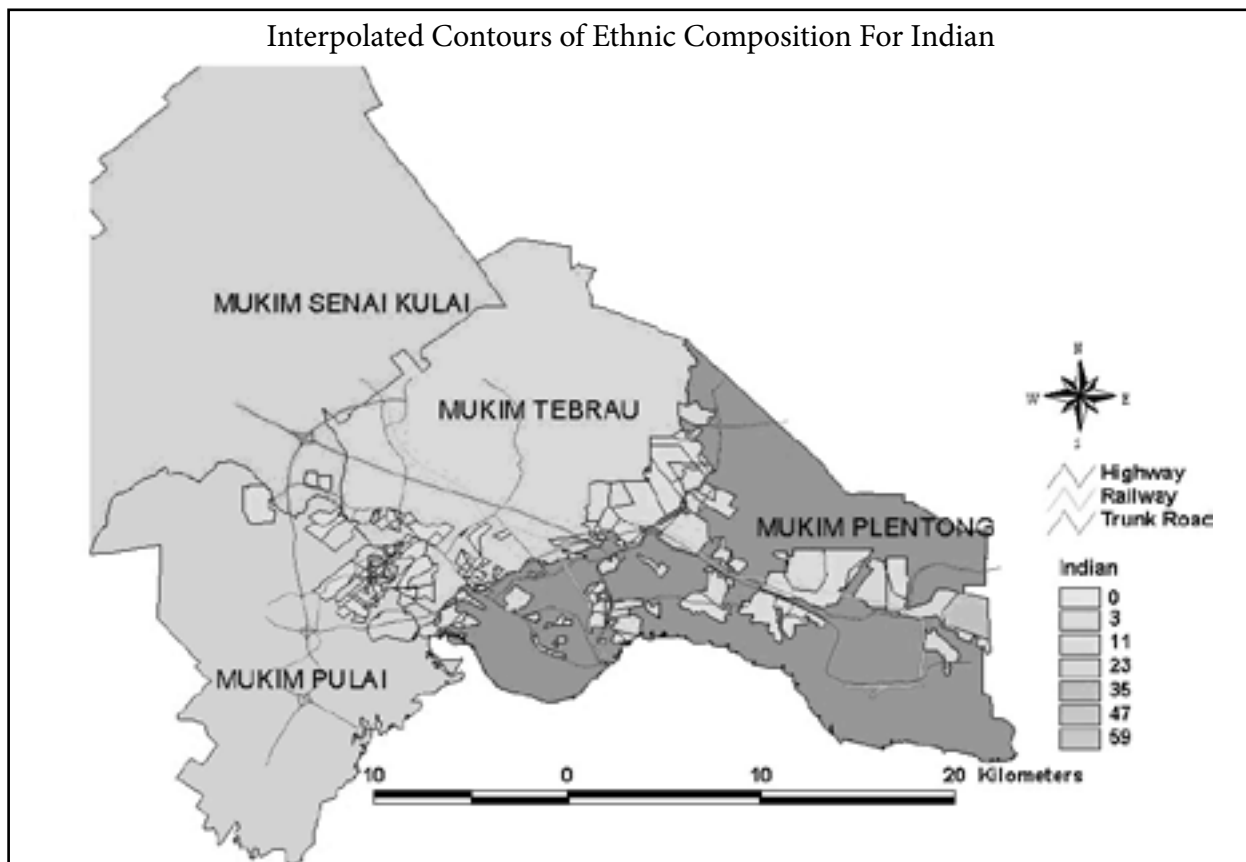


Figure 5

Interpolated Contours of Ethnic Composition For Chinese

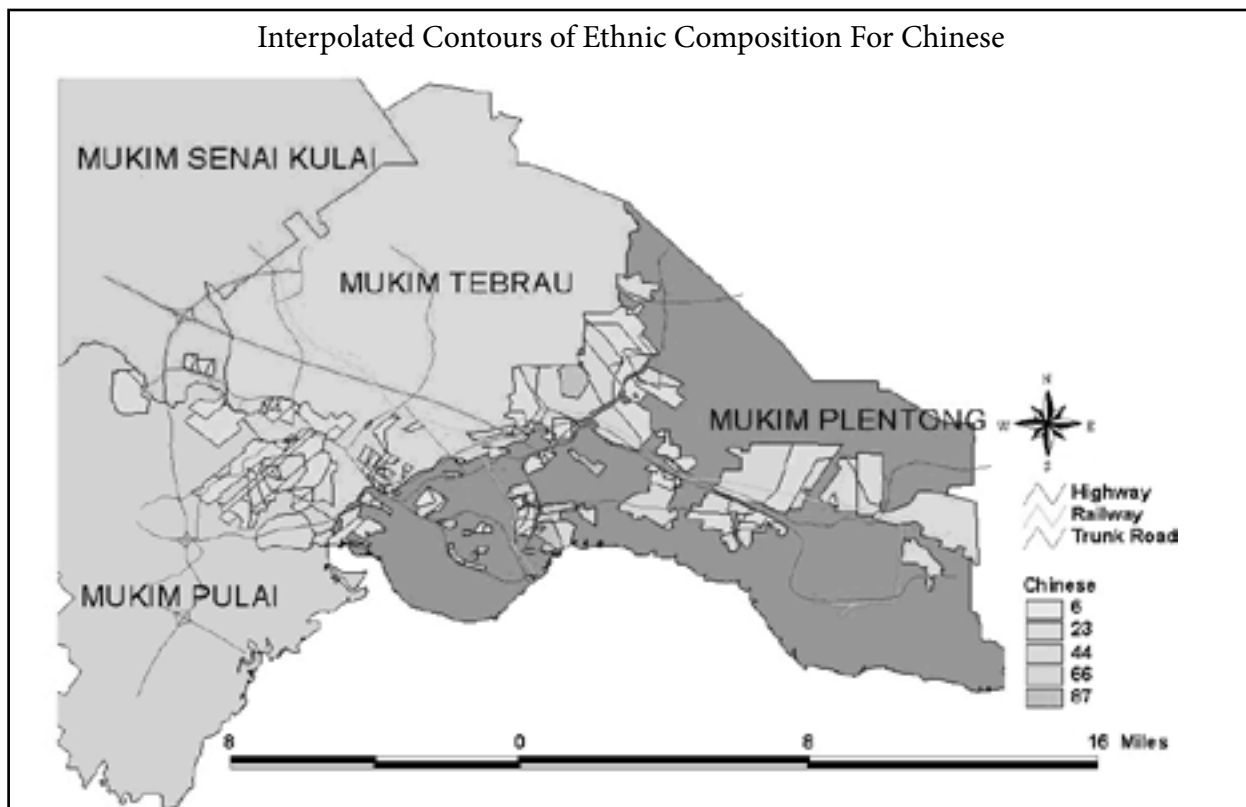


Figure 6

Interpolated Contours of Ethnic Composition For Malay

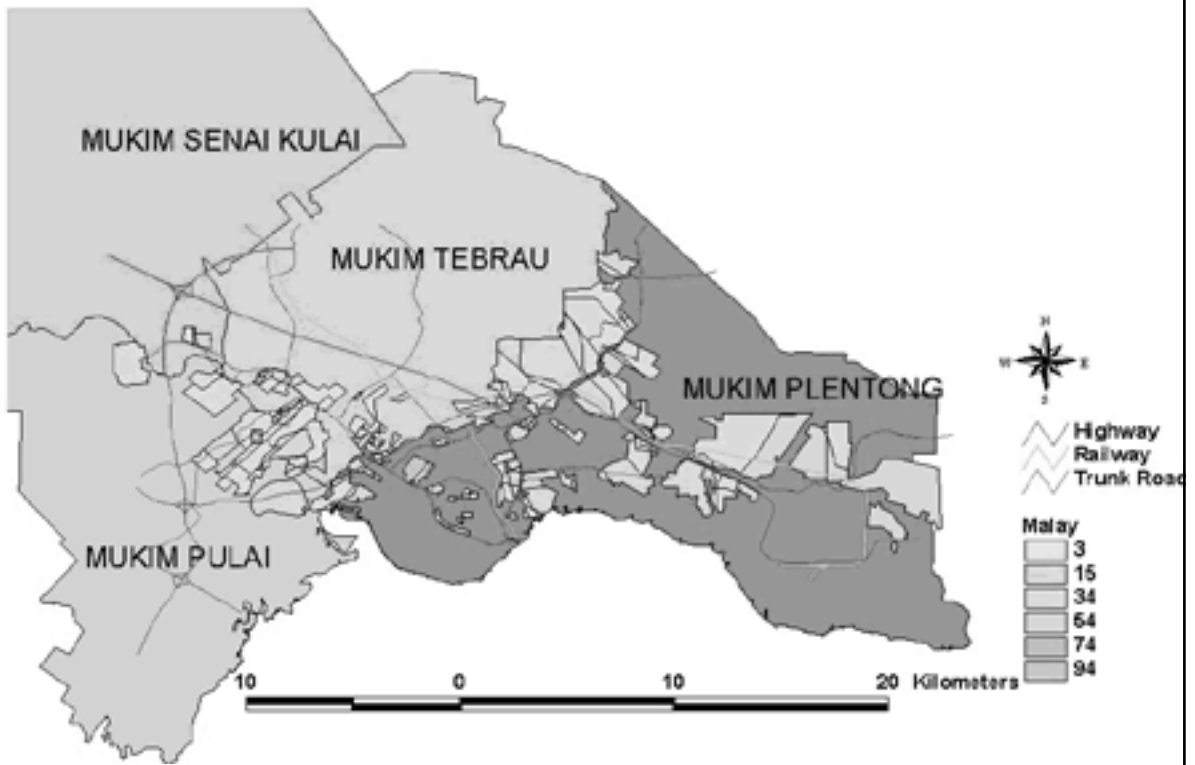


Figure 7

Distribution of Interpolated Contours of Crime Rates (Property Crime)

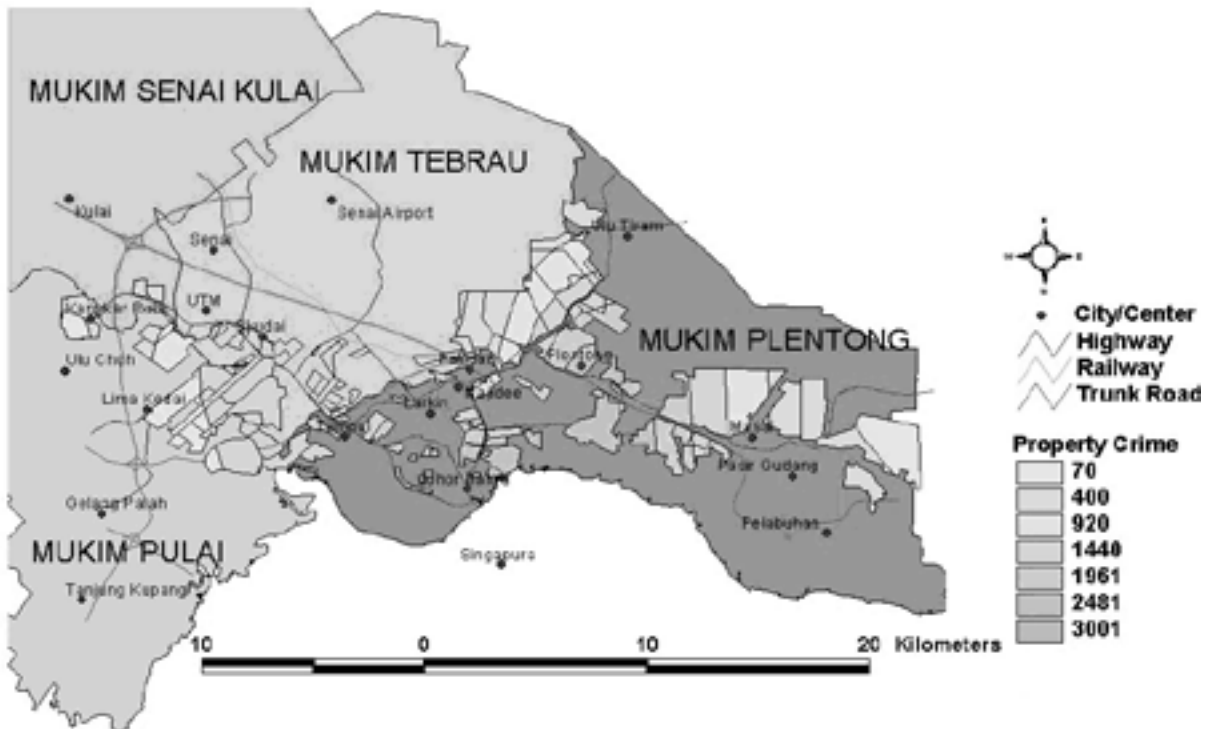


Figure 8

Distribution of Interpolated Contours of Crime Rates
(Violent Crime)

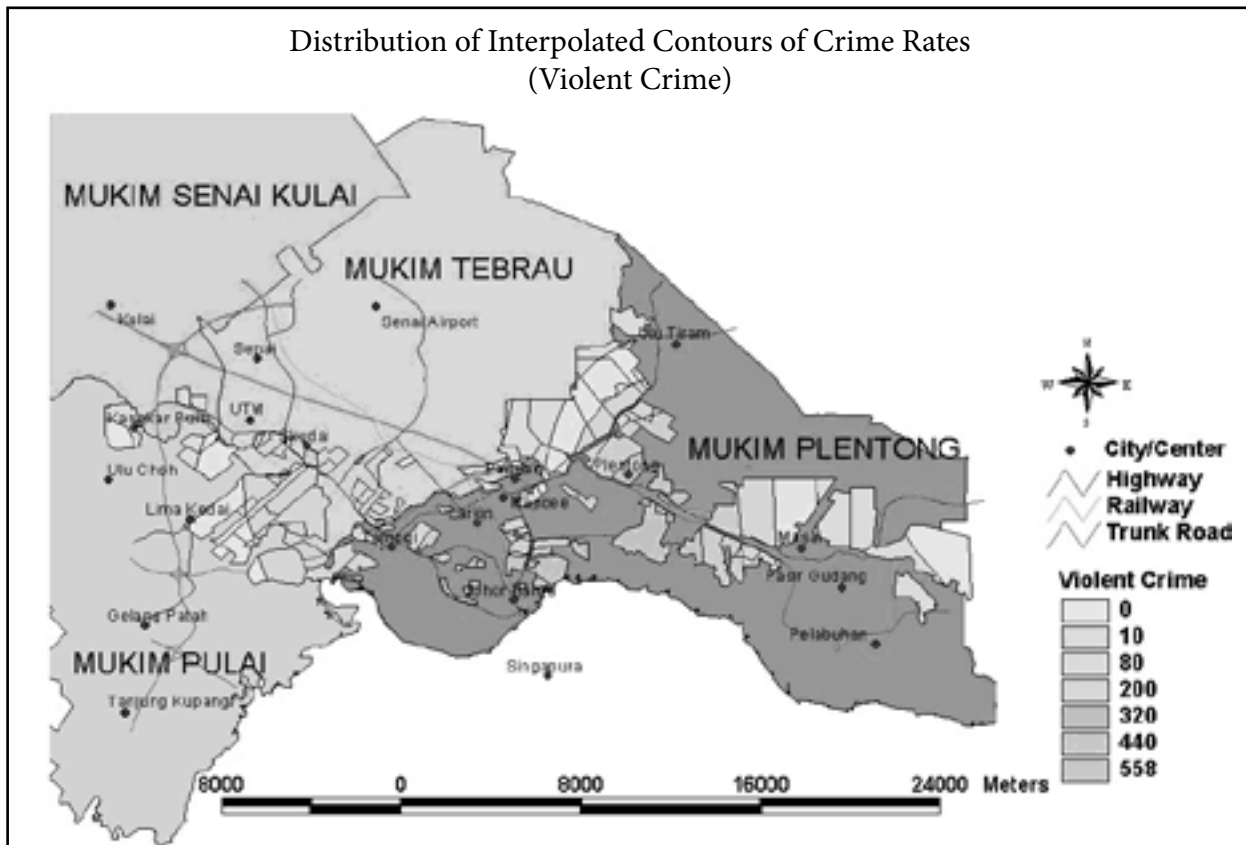


Figure 9

Spatial Distribution of Prices of Single Storey
Terraced Residential in Johor Bahru

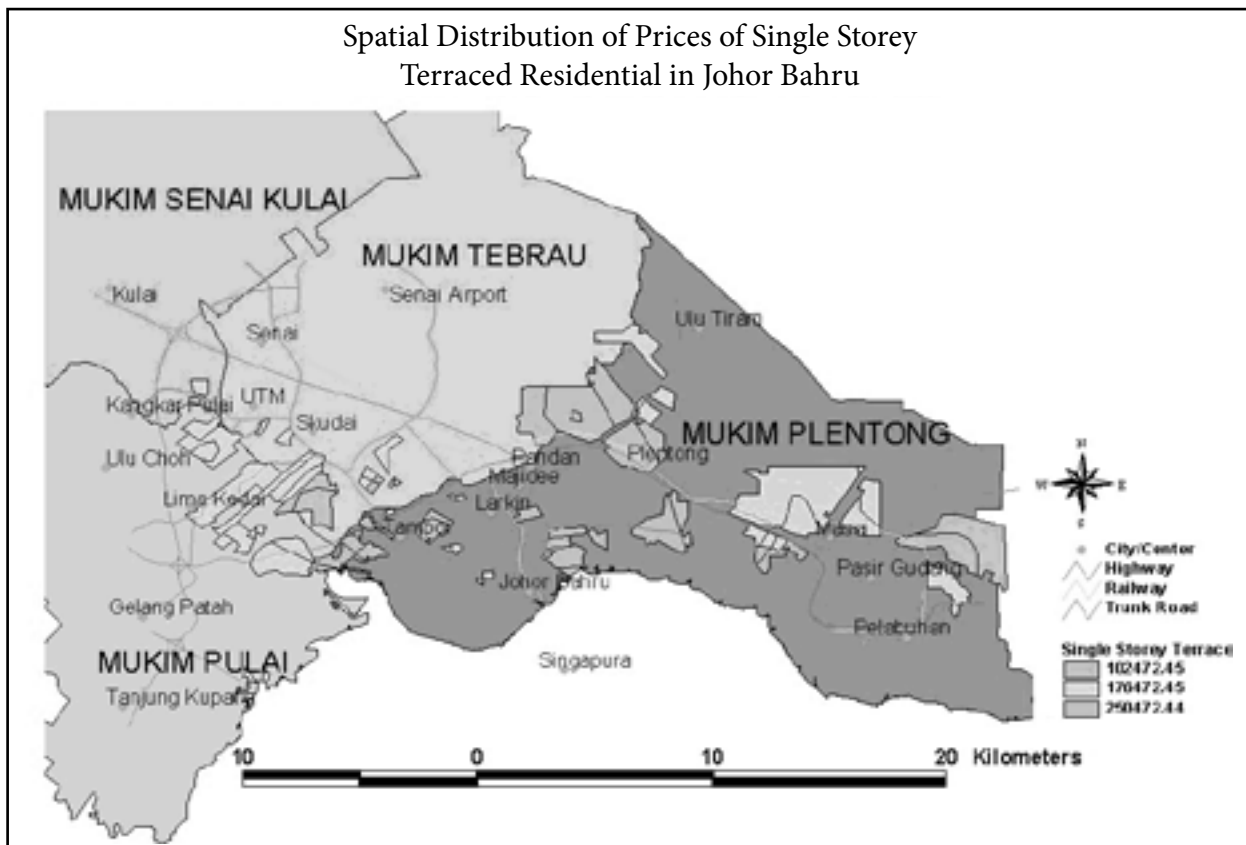


Figure 10